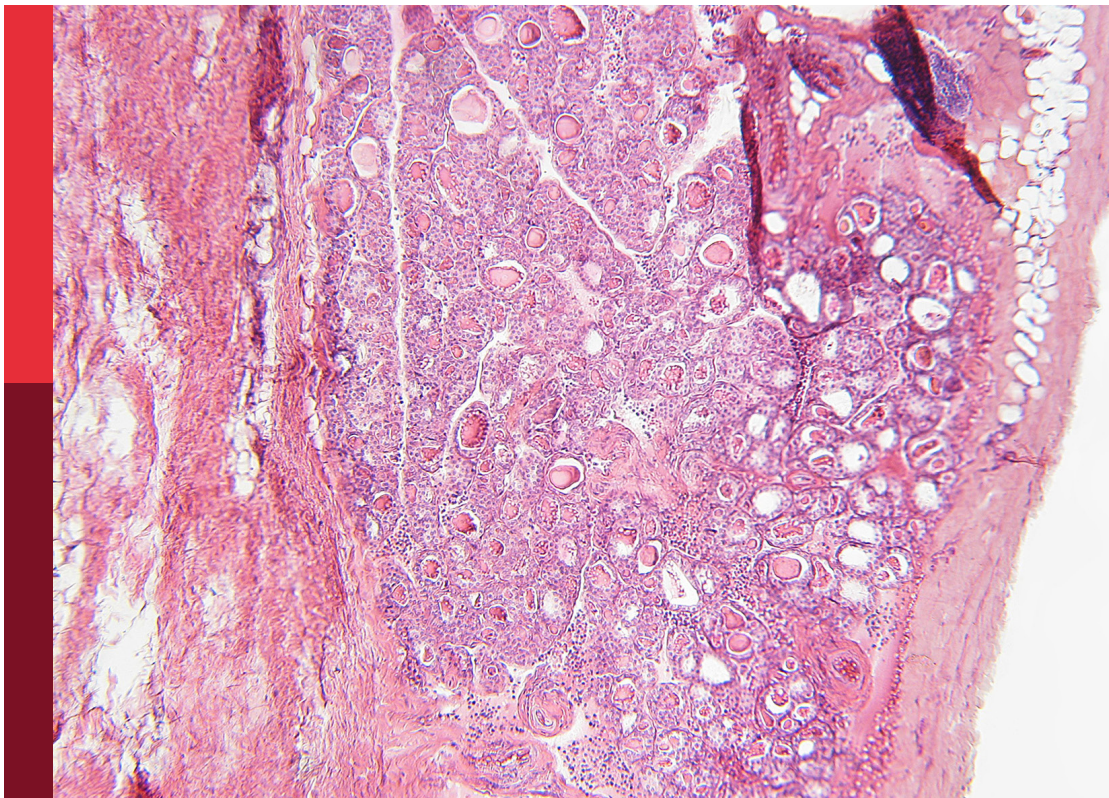
frontiers | Research Topics

Systems epidemiology of diabetes

Edited by
Changwei Li, Jin Liu, Tao Huang and Jie Li

Published in
Frontiers in Endocrinology



ISSN 1664-8714
ISBN 978-2-8325-2237-0
DOI 10.3389/978-2-8325-2237-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Systems epidemiology of diabetes

Topic editors

Changwei Li — Tulane University School of Public Health and Tropical Medicine, United States

Jin Liu — Duke-NUS Medical School, Singapore

Tao Huang — Peking University, China

Jie Li — Guangdong Provincial People's Hospital, China

Citation

Li, C., Liu, J., Huang, T., Li, J., eds. (2023). *Systems epidemiology of diabetes*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2237-0

Table of contents

05 Transition of Lipid Accumulation Product Status and the Risk of Type 2 Diabetes Mellitus in Middle-Aged and Older Chinese: A National Cohort Study
Jinyue Yu, Qian Yi, Leying Hou, Ge Chen, Yaojia Shen, Yuan Song, Yimin Zhu and Peige Song

15 Long-Term Physical Activity Participation and Subsequent Incident Type 2 Diabetes Mellitus: A Population-Based Cohort Study
Chenglong Li, Yanjun Ma, Rong Hua, Fanfan Zheng and Wuxiang Xie

26 Association Between the Severity of Diabetic Retinopathy and Optical Coherence Tomography Angiography Metrics
Binxin Xu, Jiahui Chen, Shaohua Zhang, Shengli Shen, Xuan Lan, Zhineng Chen, Zhiqiang Yan and Bingxiang Xu

38 Association Between Serum Furin and Fasting Glucose: A Cross-Sectional Study in Chinese Adults
Yan He, Hanyun Zhu, Min Zhang, Jing Li, Shengqi Ma, Yin Lu, Linan Chen, Mingzhi Zhang and Hao Peng

44 Identification of Shared Genes and Pathways in Periodontitis and Type 2 Diabetes by Bioinformatics Analysis
Junho Kang, Eun Jung Kwon, Mihyang Ha, Hansong Lee, Yeuni Yu, Ji Wan Kang, Yeongjoo Kim, Eun Young Lee, Ji-Young Joo, Hye Jin Heo, Eun Kyoung Kim, Tae Woo Kim, Yun Hak Kim and Hae Ryoun Park

53 Coffee Types and Type 2 Diabetes Mellitus: Large-Scale Cross-Phenotype Association Study and Mendelian Randomization Analysis
Xinpei Wang, Jinzhu Jia and Tao Huang

64 Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic Etiology and Causal Relationship
Yanna Chi, Xinpei Wang, Jinzhu Jia and Tao Huang

80 Adjuvant Probiotics of *Lactobacillus salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *Bifidobacterium animalis* subsp. *lactis* CP-9 Attenuate Glycemic Levels and Inflammatory Cytokines in Patients With Type 1 Diabetes Mellitus
Chung-Hsing Wang, Hung-Rong Yen, Wen-Li Lu, Hsieh-Hsun Ho, Wen-Yang Lin, Yi-Wei Kuo, Yen-Yu Huang, Shin-Yu Tsai and Hung-Chih Lin

91 Shared Genetic Basis and Causal Relationship Between Television Watching, Breakfast Skipping and Type 2 Diabetes: Evidence From a Comprehensive Genetic Analysis
Dongze Chen, Hanyu Wu, Xinpei Wang, Tao Huang and Jinzhu Jia

104 The Global Burden of Type 2 Diabetes Attributable to Tobacco: A Secondary Analysis From the Global Burden of Disease Study 2019
Jianjun Bai, Fang Shi, Yudiyang Ma, Donghui Yang, Chuanhua Yu and Jinhong Cao

116 Physical activity, sedentary behavior, and the risk of type 2 diabetes: A two-sample Mendelian Randomization analysis in the European population
Ming-Gang Deng, Han-Tao Cui, Yong-Bing Lan, Jia-Qi Nie, Yue-Hui Liang and Chen Chai

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 26 November 2021
doi: 10.3389/fendo.2021.770200

Transition of Lipid Accumulation Product Status and the Risk of Type 2 Diabetes Mellitus in Middle-Aged and Older Chinese: A National Cohort Study

Jinyue Yu[1,2†], Qian Yi[1†], Leying Hou[1], Ge Chen[3], Yaojia Shen[1], Yuan Song[4], Yimin Zhu[5] and Peige Song[1*]*

[1] School of Public Health and Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2] Institute of Child Health, University College London, London, United Kingdom, [3] Medical Research and Biometrics Center, National Center for Cardiovascular Diseases, Beijing, China, [4] School of Nursing, Henan University of Traditional Chinese Medicine, Zhengzhou, China, [5] Department of Epidemiology and Biostatistics, School of Public Health, Zhejiang University School of Medicine, Hangzhou, China

OPEN ACCESS

Edited by:
Tao Huang,
Peking University, China

Reviewed by:
Hualiang Lin,
Sun Yat-sen University, China
Zilong Zhang,
Sun Yat-sen University, China

***Correspondence:**
Peige Song
peigesong@zju.edu.cn
Yimin Zhu
zhuym@zju.edu.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *03 September 2021*
Accepted: *28 October 2021*
Published: *26 November 2021*

Citation:
Yu J, Yi Q, Hou L, Chen G, Shen Y, Song Y, Zhu Y and Song P (2021) Transition of Lipid Accumulation Product Status and the Risk of Type 2 Diabetes Mellitus in Middle-Aged and Older Chinese: A National Cohort Study. Front. Endocrinol. 12:770200. doi: 10.3389/fendo.2021.770200

Background: Lipid accumulation product (LAP), a product of waist circumference (WC) and fasting triglycerides (TG), is a measure of lipid accumulation and an effective predictor of metabolic syndrome. This study aimed to evaluate the associations of LAP and its longitudinal transitions with type 2 diabetes mellitus (T2DM) among middle-aged and older Chinese.

Methods: Data were extracted from the China Health and Retirement Longitudinal Study (2011, 2013, 2015, and 2018). LAP was defined as (WC-65) ×TG for men, and (WC-58) ×TG for women. Participants were classified into high- and low-LAP groups at baseline, and subsequently into four transition patterns during 2011-2015: maintained-high, maintained-low, high-to-low, and low-to-high LAP. The longitudinal transition patterns of LAP on the development of T2DM were assessed by multivariable Cox frailty models.

Results: Overall, 7397 participants were included for analysis, among whom 849 (11.5%) developed T2DM between 2011 and 2018. Women with high-LAP levels at baseline presented a higher risk of T2DM (hazard ratios [HR]=1.37, 95% confidence interval [CI]: 1.07-1.77), while no significant association was found in men. Compared with women with maintained-low LAP pattern, those with transition patterns of low-to-high LAP and maintained-high LAP were at higher risk of T2DM (HR =1.99 and 1.98, both $P<0.05$); however, for men, the significantly positive association was only observed in maintained-high LAP transition pattern (HR=1.53, 95% CI: 1.04-2.23).

Conclusions: Elevated LAP levels and the transition patterns of maintained-high LAP and low-to-high LAP are significant risk factors for T2DM in women. Preventions are needed to combat T2DM at an early dyslipidemic stage.

Keywords: lipid accumulation product (LAP), type 2 diabetes mellitus, risk factors, transition, China

INTRODUCTION

Diabetes mellitus (DM) is a disease of abnormal glucose metabolism that affects multiple organ systems. Globally, DM has become a high-risk factor for morbidity and mortality (1). According to the International Diabetes Federation, the number of people affected by DM globally in 2019 was 463 million, among whom an estimated 232 million people had undiagnosed DM (2). Type 2 diabetes mellitus (T2DM) accounts for over 90% of patients with DM, bringing profound psychological and physical distress to patients and putting a huge burden on health care systems (3). Evidence shows that the incidence and prevalence of T2DM vary by geographical region, with 79% of patients living in low-to-middle-income countries (2). The overall prevalence of T2DM in China increased sharply from 1.3% in 1980-1989 to 4.5% in 1990-1999, 6.8% in 2000-2009, and 8.7% in 2010-2014. Previous research estimated that by 2025, the prevalence of T2DM in China would increase to 12.5% (4), becoming a serious epidemic that produces considerable socioeconomic pressures on both individuals and the society.

A 2017 Lancet review shows that 60% of patients with T2DM are obese (body-mass index [BMI] ≥30 kg/m^2) (3). In China, a population-based survey with 5071 subjects aged 40 years or older in Shanghai demonstrated that obese people had a significantly higher risk of T2DM (5). Apart from obesity, individuals with normal weight but increased visceral adipose tissue content are also characterized by the presence of insulin resistance and impaired glucose tolerance, which are strongly associated with the development of T2DM (4, 6). In the current era when the association between high BMI and the morbidity of T2DM has been well demonstrated, research interests may need to focus more on the measurement of visceral adiposity, especially when visceral adiposity may represent a higher risk of T2DM (7).

Traditional techniques, such as magnetic resonance imaging (MRI) and computed tomography (CT), are currently the gold standard for measuring visceral adiposity (8). However, those are not suitable for routine clinical practice due to high costs, intensive labor, and the hazard posed by the use of radiation. Therefore, an alternative continuous index of visceral adiposity, the lipid accumulation product (LAP), has been proposed. Based on a combination of waist circumference (WC) and fasting triglycerides (TG), LAP could accurately reflect visceral adiposity and be easily obtained (9). A bulk of studies found that compared with conventional obesity indices, such as BMI, waist-to-hip ratio (WHR) and waist-to-height ratio (WHtR), LAP presented a better predictive ability in metabolic syndrome, DM, and impaired fasting glucose (10–12).

Previous studies have reported the efficacy of LAP in identifying visceral adiposity and insulin resistance (8, 13). However, there is a lack of studies investigating the association of LAP and the risk of T2DM among middle-aged and older Chinese (5). Moreover, whether the dynamic transition of LAP across years is associated with the development of T2DM is still poorly understood. Therefore, this study aimed to evaluate the effect of LAP and its transition on the development of T2DM in middle to older aged adults based on a national Chinese cohort study.

MATERIAL AND METHODS

Data and Sample

Data were extracted from the China Health and Retirement Longitudinal Study (CHARLS, available at http://charls.pku.edu.cn/en). CHARLS is a national survey in middle to older aged adults in China (aged 45 years and above) that attempts to provide a wealth of information ranging from socio-economic status to health conditions. The survey informs scientific research and priority setting related to the middle-aged and older adult population in China (14, 15).

The national baseline survey of CHARLS was fielded in 2011-2012. A total of 10257 households and 17708 individuals from 150 counties/districts and 450 villages/urban communities were involved across the country. The geographic regions of investigated provinces are presented in **Figure S1** and **Table S1**. The sampling strategy utilized for the CHARLS Survey involved multiple steps. First, all counties/districts were stratified according to region, urban/rural setting and economic conditions, and 150 counties/districts were randomly selected. Then, three primary sampling units (PSU) were randomly selected from each county, namely rural administrative villages and urban communities. Finally, the "CHARLS-GIS" software was used to map at least 24 families within each PSU, in which residents aged 45 years or older were interviewed with his/her spouse through a face-to-face personal interview in June 2011 and were followed up every two years. To date, CHARLS has been conducted for five rounds from 2011 to 2018. Blood tests were conducted in 2011 and 2015.

In total, 17708 participants were successfully interviewed in four rounds of the CHARLS Survey (2011, 2013, 2015 and 2018), with a family response rate of 80.5%. Individuals who were aged ≥ 45 years, with complete blood sample and biomarker data, and non-diabetic at baseline were included (n = 9485). After excluding individuals with no follow-up response (n = 531), age < 45 years (n = 194), and incomplete data of LAP at baseline (n = 1363), a total of 7397 subjects were included in the final analysis (**Figure S2**). The demographic, socio-economic, geographical and behavioral characteristics of the included and excluded subjects are outlined in **Table S2**.

This study was approved by the ethics review committee of Peking University and carried out by the National School for Development (China Centre for Economic Research) of Peking University. All participants signed written informed consent.

Measurements

From 2011 to 2018, trained interviewers collected information on demographics, geographic location, socio-economic status, health-related behaviors, and medical history through structured questionnaires. Anthropometry data were gathered following a standard protocol from the World Health Organization (16). Body weight was measured to the nearest 0.1 kg (in light clothes and without shoes) using a digital scale (Omron, HN-286). Height was measured to the nearest 0.1 cm without shoes on a stadiometer (Seca Corporation, 213). WC was horizontally measured at the middle point of the line between the lower rib and the upper iliac crest without wearing a coat. Blood pressure (BP) was measured at the right arm of the participants using an

electronic sphygmomanometer (Omron, hem-7200) at intervals of 45 seconds. Systolic (SBP) and diastolic blood pressures (DBP) were measured three times in the seated position after 10 minutes of rest by use of a sphygmomanometer.

Venous blood samples were collected by professional nurses after fasting for at least 12 hours at night. The whole blood count was performed immediately at the survey sites. The whole blood samples were then stored at 4 degrees Celsius, and the remaining samples were transported to the central Laboratory in Beijing Youanmen Centre for Clinical Laboratory of Capital Medical University for further laboratory analysis (14, 15). The levels of blood glucose, total cholesterol (TC), TG, low density lipoprotein (LDL-C) and high density lipoprotein (HDL-C) were tested by enzyme colorimetry. The Glycosylated hemoglobin (HbA1c) levels were measured by boronated-affinity high-performance liquid chromatography.

Definitions of Covariates

The tertiles of the natural logarithm of per capita expenditures (ln [PCE]) were used as indicators of family wealth, with the bottom, middle and top tertiles representing poor, middle, and rich status, respectively (17, 18). Participants' residence was classified into North China, Northeast China, East China, South-central China, Southwest China and Northwest China (see **Figure S1** and **Table S1** for more details). Hypertension was defined by an SBP≥140 mmHg, and/or a DBP≥90 mmHg, and/or a self-reported physician diagnosis, and/or currently with antihypertensive drugs, and/or under other related therapeutic measures (18, 19). Participants were categorized as normal, overweight or obese based on BMI, with cut-off points at 24 kg/m^2 and 28 kg/m^2 (20).

Definitions of LAP and T2DM

LAP was defined as (WC (cm)-65) × (TG concentration (mmol/l)) for men, and (WC (cm)-58) × (TG concentration (mmol/l)) for women (9). This formula included the minimum WC values (65 and 58 cm for men and women, respectively) used to define a gender-specific starting point in the Third National Health and Nutrition Examination (NHANES III) (9). In the present study, Participants were classified into high- and low-LAP groups at baseline using cut-off points obtained with the baseline receiver operating characteristics analysis, and then subsequently into four transition patterns during follow-up (2011-2015): maintained-high, maintained-low, high-to-low, and low-to-high LAP. The cut-off points of LAP for the diagnosis of T2DM were 18.86 (51.3% specificity and 67.7% sensitivity) for men and 40.53 (70.2% specificity and 51.1% sensitivity) for women (**Figure S3**).

The diagnosis of T2DM was established using criteria from the American Diabetes Association: fasting blood glucose ≥ 126 mg/dL (7.0 mmol/L), and/or random blood glucose ≥ 200 mg/dL (11.1 mmol/L), and/or HbA1c ≥ 6.5%, and/or self-reported diagnosis, and/or currently under hypoglycemic therapy (21).

Statistical Analysis

Statistical analysis was conducted using SAS statistical software (version 9.4; SAS Institute Inc., Cary, NC, USA). Differences in continuous and categorical variables across groups were assessed by Mann-Whitney test and Chi-square test. Person-years were calculated from the date of baseline interview and physical examination (CHARLS 2011) until the occurrence of T2DM events or death or the time when he/she was censored or the end of follow-up (CHARLS 2018), whichever came first. The cumulative incidence rates of T2DM were calculated by the Kaplan-Meier method. Cox frailty models with random effect and adjustments, were used to explore associations between LAP level and new onset T2DM by gender and setting (rural and urban). Model 1 adjusted for age. Model 2 adjusted for education, marital status, ln (PCE), region, hypertension, smoking, drinking, and general obesity based on Model 1. Model 3 adjusted for TC, LDL-C and HDL-C based on Model 2. In addition, the effects of four transition patterns of LAP on T2DM were assessed by multivariable Cox frailty models. A multiple imputation method was used to impute missing data of baseline characteristics. A total of 945 subjects were imputed for missing data (**Table S3**).

RESULTS

Demographic Characteristics

The baseline characteristics of the included participants stratified by gender, setting (rural and urban) and LAP transition pattern are shown in **Table 1** and **Table S4**. In total, 7397 participants (men=3447, women=3950) were involved in cohort analysis. No significant gender differences were observed between rural and urban areas. More women than men were illiterate in both rural and urban settings; the proportion of people with a low education level in rural areas was significantly higher than that in urban settings. The baseline LAP score was 25.46 (interquartile range [IQR]: 14.39-44.41). Totally, 4202 (56.8%) individuals were with low-LAP (men=1734, women=2468) while 3195 (43.2%) were identified as high LAP at baseline (men=1713, women=1482). Significant differences in LAP status were observed between genders in both rural and urban settings. The proportion of general obesity and hypertension was significantly higher in women than in men. Higher plasma HDL and TG were more observed in women compared to men in both rural and urban settings (P<0.05).

Association Between Baseline LAP and T2DM

Figure 1 shows the cumulative incidence of T2DM for low- and high-baseline LAP status stratified by setting and gender, from 2011 to 2018. Overall, the cumulative incidence of T2DM was highest in women with high-LAP at baseline (17.95%). Association between baseline LAP status and the risk of new-onset T2DM is shown in **Table 2** and **Table S5**. Positive associations between incident T2DM and high-LAP status at baseline were observed in both genders, with a crude hazard ratios (HR) of 1.78 (95% confidence interval [CI]: 1.43-2.22) for men and =2.00 (95%CI: 1.68-2.39) for women. After full adjustments, the association was still significant in women (HR =1.37, 95%CI: 1.07-1.77) but insignificant in men (HR = 1.29,

TABLE 1 | Demographic, socioeconomic and geographic characteristics of the included participants at baseline (CHARLS 2011).

Characteristics	Overall (n = 7,397)	Male		P^a	Female		P^b
		Rural (n = 2,353)	Urban (n = 1,094)		Rural (n = 2,626)	Urban (n = 1,324)	
Age group				0.131			0.156
45-49 years	1437(19.4%)	369(15.7%)	182(16.6%)		566(21.5%)	320(24.2%)	
50-59 years	2641(35.7%)	819(34.8%)	416(38.1%)		942(35.9%)	464(35.0%)	
60-69 years	2182(29.5%)	748(31.8%)	326(29.8%)		759(28.9%)	349(26.4%)	
≥70 years	1137(15.4%)	417(17.7%)	170(15.5%)		359(13.7%)	191(14.4%)	
Education				<0.001			<0.001
Illiterate	2173(29.4%)	393(16.7%)	74(6.7%)		1318(50.2%)	388(29.3%)	
Literate	1403(19.0%)	496(21.1%)	203(18.6%)		488(18.6%)	216(16.3%)	
Primary education	1661(22.5%)	699(29.7%)	259(23.7%)		450(17.1%)	253(19.1%)	
Middle school	2159(29.1%)	764(32.5%)	558(51.0%)		370(14.1%)	467(35.3%)	
Marital status				<0.001			0.035
Married or cohabiting	6518(88.1%)	2108(89.6%)	1035(94.6%)		2266(86.3%)	1109(83.8%)	
Single	879(11.9%)	245(10.4%)	59(5.4%)		360(13.7%)	215(16.2%)	
Ln(PCE)				<0.001			<0.001
Bottom tertile	2164(33.3%)	801(37.9%)	223(23.3%)		887(38.4%)	253(22.6%)	
Middle tertile	2166(33.3%)	704(33.3%)	326(34.0%)		749(32.5%)	387(34.6%)	
Top tertile	2166(33.3%)	607(28.7%)	409(42.7%)		672(29.1%)	478(42.8%)	
Region of China				<0.001			0.001
North	950(12.9%)	342(14.5%)	120(11.0%)		328(12.4%)	160(12.0%)	
Northeast	503(6.8%)	136(5.8%)	92(8.4%)		167(6.4%)	108(8.2%)	
East	2221(30.0%)	678(28.8%)	345(31.5%)		793(30.2%)	405(30.6%)	
South Central	1747(23.6%)	486(20.7%)	294(26.9%)		614(23.4%)	353(26.7%)	
Southwest	1366(18.5%)	475(20.2%)	194(17.7%)		474(18.1%)	223(16.8%)	
Northwest	610(8.2%)	236(10.0%)	49(4.5%)		250(9.5%)	75(5.7%)	
General Obesity[c]				<0.001			<0.001
Normal	3657(50.5%)	1452(62.9%)	497(46.2%)		1226(47.8%)	482(37.3%)	
Overweight	2748(37.9%)	715(30.9%)	461(42.8%)		1007(39.2%)	565(43.8%)	
Obese	840(11.6%)	144(6.2%)	118(11.0%)		334(13.0%)	244(18.9%)	
Hypertension[c]				0.009			0.011
Normal	4394(59.5%)	1473(62.8%)	634(58.1%)		1559(59.5%)	728(55.2%)	
Hypertension	2986(40.5%)	874(37.2%)	458(41.9%)		1063(40.5%)	591(44.8%)	
Smoking				<0.001			0.030
No Smoking	4506(61.1%)	538(23.0%)	323(29.7%)		2441(93.1%)	1204(91.1%)	
Smoking	2869(38.9%)	1805(77.0%)	766(70.3%)		181(6.9%)	117(8.9%)	
Drinking[c]				0.859			0.007
No Drinking	5089(68.8%)	1046(44.5%)	483(44.1%)		2345(89.3%)	1215(92.0%)	
Drinking	2304(31.2%)	1306(55.5%)	611(55.9%)		281(10.7%)	106(8.0%)	
TC				0.291			0.489
≤200mg/dL	4577(61.9%)	1602(68.1%)	725(66.3%)		1506(57.3%)	744(56.2%)	
>200mg/dL	2820(38.1%)	751(31.9%)	369(33.7%)		1120(42.7%)	580(43.8%)	
HDL				<0.001			<0.001
≥50mg/dL	3730(50.4%)	1179(50.1%)	437(39.9%)		1460(55.6%)	654(49.4%)	
<50mg/dL	3667(49.6%)	1174(49.9%)	657(60.1%)		1166(44.4%)	670(50.6%)	
LDL				0.001			0.253
≤100mg/dL	3214(43.5%)	1199(51.0%)	489(44.7%)		1031(39.3%)	495(37.4%)	
>100mg/dL	4183(56.5%)	1154(49.0%)	605(55.3%)		1595(60.7%)	829(62.6%)	
TG				<0.001			0.012
≤150mg/dL	5667(76.6%)	1925(81.8%)	824(75.3%)		1973(75.1%)	945(71.4%)	
>150mg/dL	1730(23.4%)	428(18.2%)	270(24.7%)		653(24.9%)	379(28.6%)	
LAP SCORE				<0.001			<0.001
Median (IQR)	25.46(14.39-44.41)	16.90(9.67-30.91)	23.56(12.89-41.87)		30.10(18.47-49.10)	35.59(20.91-57.75)	
LAP				<0.001			<0.001
Low LAP	4202(56.8%)	1303(55.4%)	431(39.4%)		1719(65.5%)	749(56.6%)	
High LAP	3195(43.2%)	1050(44.6%)	663(60.6%)		907(34.5%)	575(43.4%)	

Data were presented as n (%) or median with interquartile range (IQR); [a]comparison between rural males and urban males; [b]comparison between rural females and urban females; [c]data for some participants were missing; PCE, per capita expenditures; LAP, lipid accumulation product.

95% CI: 0.96-1.74). When further stratified by setting, fully adjusted models revealed significant associations between baseline high-LAP and T2DM only in rural women, with HR being 1.60 (95% CI: 1.18-2.17).

Association Between LAP Transition and T2DM

In terms of LAP transitions, it was found that the cumulative incidence of T2DM was the highest in participants with

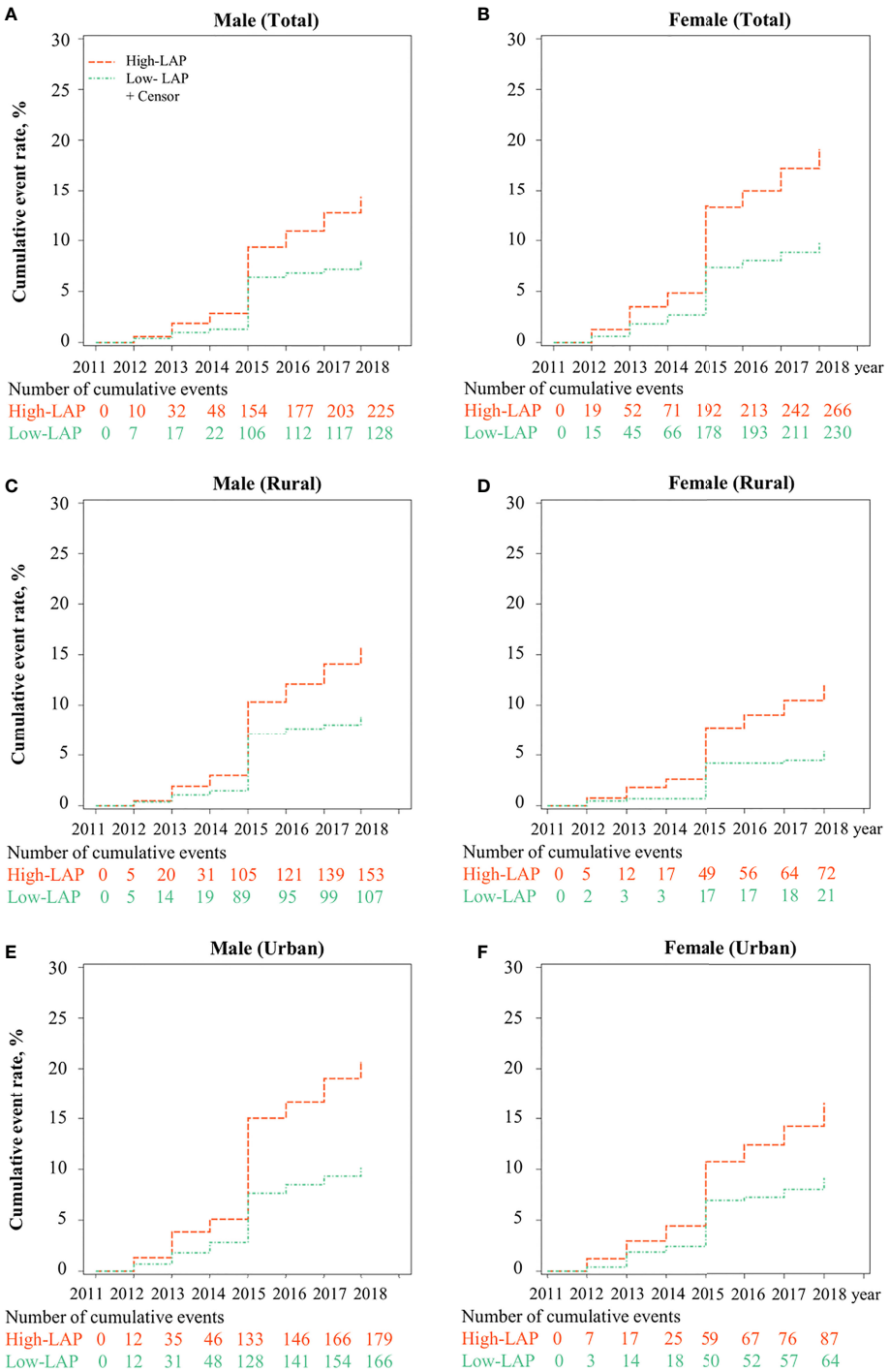


FIGURE 1 | Cumulative incidence of T2DM for LAP phenotypes (Low- and High-) stratified by urban/rural settings and sex from CHARLS 2011 to 2018. *LAP, lipid accumulation product; The cumulative event rate of T2DM was significantly different between low-LAP and high-LAP in* **(A)** *Male*, **(B)** *Female*, **(C)** *Male in rural*, **(D)** *Female in rural*, **(E)** *Male in urban, and* **(F)** *Female in urban*.

maintained-high (**Figure 2**). **Figure 3** reports the risk of T2DM by LAP transition pattern stratified by gender after adjusting for age, education, region, hypertension, smoking, drinking, general obesity, TC, LDL-C and HDL-C. Overall, compared to individual whose LAP level was maintained low from baseline to follow-up, a 1.5 to 2.0-fold risk of T2DM was observed in people with maintained-high LAP pattern (HR_{men} = 1.53, 95 CI%: 1.04-2.23, HR_{women} = 1.98, 95% CI: 1.43-2.75, $P<0.05$), and a two-fold risk

TABLE 2 | Hazard ratios for T2DM by LAP phenotypes in middle-aged and older Chinese, CHARLS 2011-2018.

Model	LAP group for Male		LAP group for Female	
	Low LAP (n = 1,734)	High LAP (n = 1,713)	Low LAP (n = 2,468)	High LAP (n = 1,482)
New-onset T2DM	128 (7.38)	225 (13.13)	230 (9.32)	266 (17.95)
Overall				
Unadjusted	1 (reference)	1.78 (1.43, 2.22)[a]	1 (reference)	2.00 (1.68, 2.39)[a]
Model 1	1 (reference)	1.83 (1.47, 2.28)[a]	1 (reference)	1.98 (1.66, 2.36)[a]
Model 2	1 (reference)	1.39 (1.05, 1.85)[a]	1 (reference)	1.51 (1.20, 1.90)[a]
Model 3	1 (reference)	1.29 (0.96, 1.74)	1 (reference)	1.37 (1.07, 1.77)[a]
Rural				
Unadjusted	1 (reference)	1.76 (1.37, 2.25)[a]	1 (reference)	2.12 (1.71, 2.62)[a]
Model 1	1 (reference)	1.82 (1.41, 2.33)[a]	1 (reference)	2.11 (1.71, 2.62)[a]
Model 2	1 (reference)	1.29 (0.94, 1.78)	1 (reference)	1.67 (1.27, 2.20)[a]
Model 3	1 (reference)	1.20 (0.86, 1.68)	1 (reference)	1.60 (1.18, 2.17)[a]
Urban				
Unadjusted	1 (reference)	2.27 (1.39, 3.69)[a]	1 (reference)	1.84 (1.34, 2.55)[a]
Model 1	1 (reference)	2.30 (1.41, 3.75)[a]	1 (reference)	1.73 (1.25, 2.39)[a]
Model 2	1 (reference)	2.06 (1.11, 3.84)[a]	1 (reference)	1.26 (0.82, 1.93)
Model 3	1 (reference)	1.85 (0.96, 3.59)	1 (reference)	1.02 (0.63, 1.63)

[a]*P < 0.05.Data were presented as n (%) or hazard ratios (95% CI); Associations between LAP and the risk of T2DM were assessed using multivariable Cox frailty models with random intercepts to account for clustering of participants by city; T2DM, type 2 diabetes mellitus; LAP, lipid accumulation product.*

Model 1 was adjusted for age. Model 2 was adjusted for education, marital status, Ln(PCE), region, hypertension, smoking, drinking, and general obesity based on Model 1. Model 3 was adjusted for TC, LDL-c and HDL-c based on Model 2.

of T2DM was observed in women whose LAP level was transferred from low to high (HR = 1.99, 95% CI: 1.46-2.71). Comparatively, male participants with three transition patterns (low-to-high, high-to-low, and maintained-high) showed no significant difference in the risk of T2DM in both urban and rural settings, compared to those who have maintained-low LAP pattern during the follow-up. Similar results were found in **Tables S6, S7** for data with multiple imputations.

DISCUSSION

In this prospective cohort study, we confirmed that higher LAP level was a significant risk factor for T2DM, especially in women. Women with high-LAP status at baseline presented a higher risk of T2DM. When assessing the relationship between LAP transitions and T2DM, women with maintained-high LAP pattern and low-to-high transition pattern were at an almost two-fold risk of T2DM compared to maintained-low LAP group. However, no such association was generally found in men, except for the maintained-high LAP pattern.

A number of studies have reported the rapid increase in the prevalence of DM in the Chinese adult population (22–24). The disease burden of DM in China is likely the result of population aging, urbanization, unhealthy diets, reduced exercise, and the consequent epidemic of obesity (25). Although visceral fat has been found to be independently associated with insulin resistance and could be used to estimate the risk of T2DM, the measurement for body visceral adiposity tissue volumes is not easy to conduct. MRI and CT have been considered the gold standard for visceral adiposity measurements, but both are not suitable for large epidemiological studies due to high costs and inconvenience. Comparatively, LAP can be easily measured in large-scale epidemiological studies and has been suggested as a useful surrogate marker of visceral adiposity. Previous studies have demonstrated that LAP has a similar or greater capacity to predict T2DM when compared with common fatness indices, such as BMI, WC, WHR and WHtR (10–12). For example, a six-year cohort study showed that LAP performed as good as BMI, WHR, and WHtR when predicting T2DM (26). Findings from this study were consistent with previous results, which indicated the positive effects of LAP on the prediction of DM in both genders (9, 26–28). Furthermore, in a recent cohort study examining the potential of LAP as an indicator for the development of T2DM among 15717 Chinese individuals (28), results showed a significantly increased risk of T2DM development in high LAP groups.

As for the effects of LAP by gender and by rural/urban setting, this study observed that the cumulative incidence of T2DM in women with high-LAP status was the highest. Similar results were found in a large population-based study in Japan (29), which reported the odds ratio for DM in subjects with high LAP was higher in women (OR=19.09) than in men (OR=7.40). A recent meta-analysis demonstrated a higher prevalence of T2DM in rural Chinese women than that in men (30). The difference in T2DM prevalence between genders was presumably explained by two reasons. The first might be the difference in sex steroid hormones and glucose homeostasis (31). Since the fat mass in healthy women is higher than that in men, the circulating free fatty acids and intramuscular fat content would consequently be higher (32), which could therefore induce insulin resistance in women. Another explanation might be the lack of physical activity in women compared to men. A regional study in China showed that rural men presented a significantly lower risk of developing metabolic syndrome than rural women and the low risks were observed in those who were less sitting and engaged in more vigorous physical activity (33). Moreover, randomized clinical trials have shown that interventions involving exercise were significantly associated with a reduced

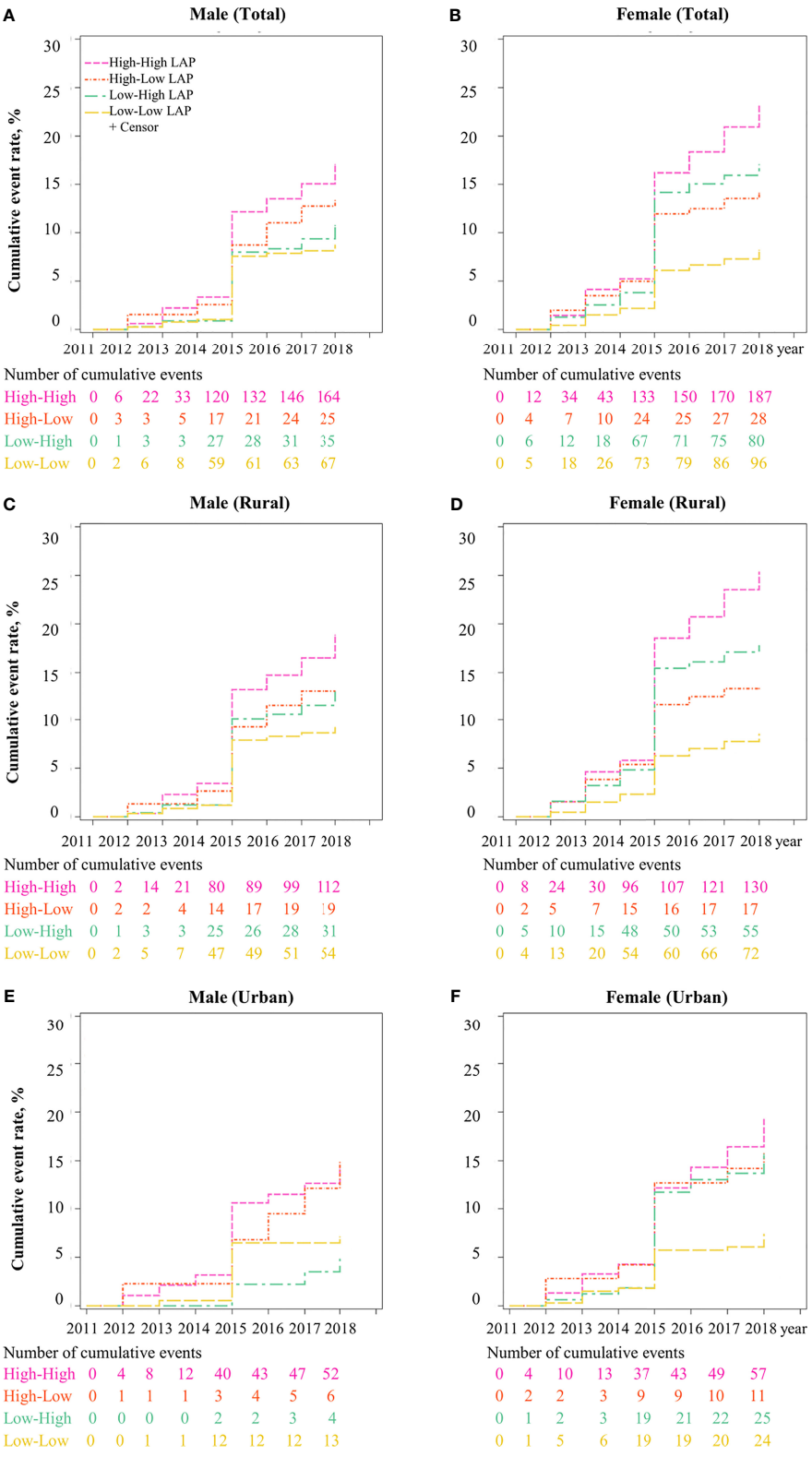


FIGURE 2 | Cumulative incidence of T2DM for LAP transitions from CHARLS 2011 to 2018. Note: LAP, lipid accumulation product; The cumulative event rate of T2DM was significantly different across four LAP transitions (high to high-, high to low-, low to high-, low to low-) in **(A)** Male, **(B)** Female, **(C)** Male in rural, **(D)** Female in rural, **(E)** Male in urban, and **(F)** Female in urban.

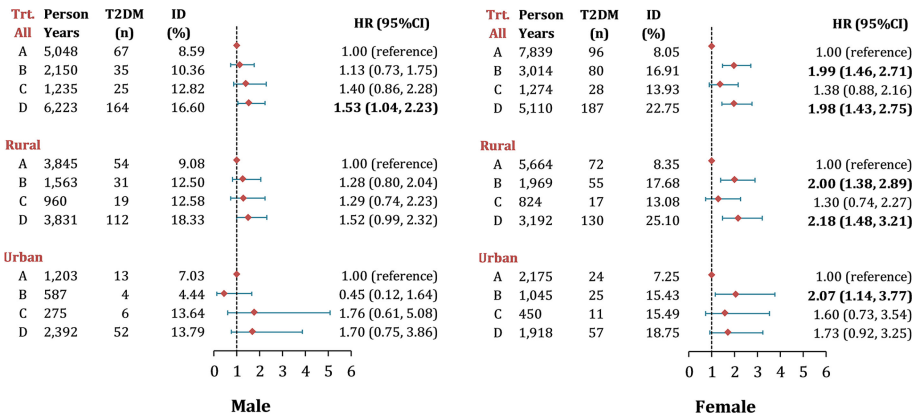


FIGURE 3 | Risk of new-onset T2DM by different LAP transitions in middle-aged and older Chinese. Note: Data were presented as n (%) and hazard ratios (95% CI), adjusted for age, education, region, hypertension, smoking, drinking, general obesity, TC, LDL-c level and HDL-c level; Hazard ratios for T2DM by LAP transitions were calculated using multivariable Cox frailty models with random effect, by which means clustering of participants was accounted for; P< 0.05 were highlighted in bold. ID, incidence density; LAP, lipid accumulation product; Trt., transition types during follow-up, the definition from group A to D were listed as following: Group A, maintain Low LAP during follow-up; Group B, Low LAP at baseline turned to High LAP at follow-up; Group C, High LAP at baseline turned to Low LAP at follow-up; Group D, maintain High LAP during follow-up.

risk of DM among people with prediabetes (34–36). Therefore, it is important that future research investigates the lifestyles of women in rural China, public health measures are also needed to mitigate the consequences of new cases of T2DM in this group.

Interestingly, for female participants, when assessing the impact of LAP transition on T2DM, the low-to-high LAP group had a similar risk for developing T2DM compared to the maintained-high LAP group, both groups were significantly higher than the maintained-low LAP group; comparatively, no significant difference was found between high-to-low LAP group and maintained-low LAP group. These findings may partly be explained by the "vicious circle" - the dyslipidemia-insulin resistance-hyperinsulinemia circle - in the development of T2DM (30). In this vicious circle, the consistently high level of TG, which leads to a maintained-high LAP, contributes to T2DM by competing with glucose to enter the cell, decreasing the activity of insulin receptors on fat cells, and preventing insulin from combining with receptors; moreover, the high LAP could also attribute to the decreased level of HDL, which can negatively influence the β cell's function in the pancreas and reduce insulin sensitivity (31). Conversely, insulin resistance could result in the increasing of TG and decreasing of HDL. Given those, an increase in LAP from a low to a high level may be associated with the development of T2DM. Meanwhile, reducing and maintaining LAP at a lower level would be vital for the prevention of T2DM. Hence, further intervention studies could consider ways of shifting LAP levels as a way of preventing T2DM, especially for women.

This study is the first to analyze the impact of LAP and its transition on incident T2DM in middle to older aged population. Compared to previous studies, the sample of this study was relatively large, and representative based on participants from multiple regions in China. One of the limitations in the present study is that, due to the original design of the survey, blood sample data were only available in 2011 and 2015, which brought sudden increases in incident T2DM in 2015. Moreover, several confounding factors, such as family history and diet were not accounted for in our multivariable analyses due to the absence of relative information in the CHARLS dataset. Additionally, considering over half of the participants were excluded from this study, selection bias may be introduced.

This study demonstrated that high LAP was associated with the development of T2DM among Chinese women over 45 years old. The maintained high LAP and transition of LAP from low to high levels were confirmed as risk factors for T2DM. Future efforts aimed at preventing T2DM should be made to explore ways of decreasing LAP levels.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: http://charls.pku.edu.cn/pages/data/111/en.html.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Review Committee of Peking University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conception or design: PS. Acquisition, analysis, or interpretation of data: GC, QY, and JY. Drafting the work or revising: JY, QY, YaS, YuS, LH, and YZ. Final approval of the manuscript: PS, YaS, YuS, LH, YZ, JY, and QY. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

All authors thank the China Health and Retirement Longitudinal Study (CHARLS) for providing data. We are grateful for those who designed, conducted and participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2021.770200/full#supplementary-material

REFERENCES

1. Roglic G, Unwin N. Mortality Attributable to Diabetes: Estimates for the Year 2010. *Diabetes Res Clin Pract* (2010) 87:15–9. doi: 10.1016/j.diabres.2009.10.006
2. I.D. Federation and IDF Diabetes Atlas. *International Diabetes Federation*. Brussels, Belgium: International Diabetes Federation (2019).
3. Chatterjee S, Khunti K, Davies MJ. Type 2 Diabetes. *Lancet* (2017) 389:2239–51. doi: 10.1016/S0140-6736(17)30058-2
4. Yuan H, Li X, Wan G, Sun L, Zhu X, Che F, et al. Type 2 Diabetes Epidemic in East Asia: A 35–Year Systematic Trend Analysis. *Oncotarget* (2018) 9:6718. doi: 10.18632/oncotarget.22961
5. He Y-H, Jiang G-X, Yang Y, Huang H-E, Li R, Li X-Y, et al. Obesity and Its Associations With Hypertension and Type 2 Diabetes Among Chinese Adults Age 40 Years and Over. *Nutrition* (2009) 25:1143–9. doi: 10.1016/j.nut.2009.04.003
6. Ganguly S, Ray L, Kuruvila S, Nanda SK, Ravichandran K. Lipid Accumulation Product Index as Visceral Obesity Indicator in Psoriasis: A Case–Control Study. *Indian J Dermatol* (2018) 63:136. doi: 10.4103/ijd.IJD_315_16
7. Bozorgmanesh, Mohammadreza, Farzad Hadaegh, Fereidoun Azizi. Predictive Performance of the Visceral Adiposity Index for a Visceral Adiposity-Related Risk: Type 2 Diabetes. *Lipids Health Disease* (2011) 10:1–9. doi: 10.1186/1476-511X-10-88
8. Du T, Yu X, Zhang J, Sun X. Lipid Accumulation Product and Visceral Adiposity Index are Effective Markers for Identifying the Metabolically Obese Normal-Weight Phenotype. *Acta Diabetologica* (2015) 52:855–63. doi: 10.1007/s00592-015-0715-2
9. Kahn HS. The Lipid Accumulation Product Is Better Than BMI for Identifying Diabetes: A Population-Based Comparison. *Diabetes Care* (2006) 29:151–3. doi: 10.2337/diacare.29.01.06.dc05-1805
10. Song J, Chen X, Jiang Y, Mi J, Zhang Y, Zhao Y, et al. Association and Interaction Analysis of Lipid Accumulation Product With Impaired Fasting Glucose Risk: A Cross-Sectional Survey. *J Diabetes Res* (2019) 2019:1–10. doi: 10.1155/2019/9014698
11. Ray L, Ravichandran K, Nanda SK. Comparison of Lipid Accumulation Product Index With Body Mass Index and Waist Circumference as a Predictor of Metabolic Syndrome in Indian Population. *Metab Syndrome related Disord* (2018) 16:240–5. doi: 10.1089/met.2017.0119
12. Tian T, Pei H, Chen Z, Hailili G, Wang S, Sun Y, et al. Comparison of Lipid Accumulation Product and Body Mass Index as Indicators of Diabetes Diagnosis Among 215,651 Chinese Adults. *PeerJ* (2020) 8:e8483. doi: 10.7717/peerj.8483
13. Xia C, Li R, Zhang S, Gong L, Ren W, Wang Z, et al. Lipid Accumulation Product Is a Powerful Index for Recognizing Insulin Resistance in Non-Diabetic Individuals. *Eur J Clin Nutr* (2012) 66:1035–8. doi: 10.1038/ejcn.2012.83
14. Zhang LF, Pei Q, Yang GP, Zhao YC, Mu YF, Huang Q, et al. The Effect of IGF2BP2 Gene Polymorphisms on Pioglitazone Response in Chinese Type 2 Diabetes Patients. [Erratum Appears in Pharmacology. 2014;94(3-4):198]. *Pharmacology* (2014) 94:115–22. doi: 10.1159/000363414
15. Zhao Y, Strauss J, Yang G, Giles J, Hu P, Hu Y, et al. *China Health and Retirement Longitudinal Study–2011–2012 National Baseline Users' Guide*. Beijing: National School of Development, Peking University (2013) p. 1–56.
16. W.H. Organization. *Physical Status: The Use of and Interpretation of Anthropometry, Report of a WHO Expert Committee*. Geneva, Switzerland: World Health Organization (1995).
17. Deaton A. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington: The World Bank (1997).
18. Song P, Wang H, Xia W, Chang X, Wang M, An L. Prevalence and Correlates of Hyperuricemia in the Middle-Aged and Older Adults in China. *Sci Rep* (2018) 8:4314. doi: 10.1038/s41598-018-22570-9
19. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. Evidence-Based Guideline for the Management of High Blood Pressure in Adults: Report From the Panel Members Appointed to the Eighth Joint National Committee (JNC 8). *JAMA* (2014) 311:507–20. doi: 10.1001/jama.2013.284427
20. Chen C, Lu FC. The Guidelines for Prevention and Control of Overweight and Obesity in Chinese Adults. *Biomed Environ Sci BES* (2004) 17 Suppl:1–36.
21. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* (2014) 37:S81–S90. doi: 10.2337/dc14-er03
22. Pan X-R, Yang W-Y, Li G-W, Liu J. Prevalence of Diabetes and Its Risk Factors in China, 1994. *Diabetes Care* (1997) 20:1664–9. doi: 10.2337/diacare.20.11.1664
23. Gu D, Reynolds K, Duan X, Xin X, Chen J, Wu X, et al. Prevalence of Diabetes and Impaired Fasting Glucose in the Chinese Adult Population: International Collaborative Study of Cardiovascular Disease in Asia (InterASIA). *Diabetologia* (2003) 46:1190–8. doi: 10.1007/s00125-003-1167-8
24. Jia W, Pang C, Chen L, Bao Y, Lu J, Lu H, et al. Epidemiological Characteristics of Diabetes Mellitus and Impaired Glucose Regulation in a Chinese Adult Population: The Shanghai Diabetes Studies, a Cross-Sectional 3-Year Follow-Up Study in Shanghai Urban Communities. *Diabetologia* (2007) 50:286–92. doi: 10.1007/s00125-006-0503-1
25. Yang W, Lu J, Weng J, Jia W, Ji L, Xiao J, et al. Prevalence of Diabetes Among Men and Women in China. *New Engl J Med* (2010) 362:1090–101. doi: 10.1056/NEJMoa0908292
26. Bozorgmanesh M, Hadaegh F, Azizi F. Diabetes Prediction, Lipid Accumulation Product, and Adiposity Measures; 6-Year Follow-Up: Tehran Lipid and Glucose Study. *Lipids Health Dis* (2010) 9:45. doi: 10.1186/1476-511X-9-45
27. Nusrianto R, Ayundini G, Kristanti M, Astrella C, Amalina N, Riyadina W, et al. Visceral Adiposity Index and Lipid Accumulation Product as a Predictor of Type 2 Diabetes Mellitus: The Bogor Cohort Study of Non-Communicable Diseases Risk Factors. *Diabetes Res Clin Pract* (2019) 155:107798. doi: 10.1016/j.diabres.2019.107798
28. Xu M, Huang M, Qiang D, Gu J, Li Y, Pan Y, et al. Hypertriglyceridemic Waist Phenotype and Lipid Accumulation Product: Two Comprehensive Obese Indicators of Waist Circumference and Triglyceride to Predict Type 2 Diabetes Mellitus in Chinese Population. *J Diabetes Research* (2020) 2020:1–12. doi: 10.1155/2020/9157430
29. Wakabayashi I, Daimon T. A Strong Association Between Lipid Accumulation Product and Diabetes Mellitus in Japanese Women and Men. *J Atheroscl Thromb* (2013) 21(3):20628. doi: 10.5551/jat.20628
30. Liu X, Li Y, Li L, Zhang L, Ren Y, Zhou H, et al. Prevalence, Awareness, Treatment, Control of Type 2 Diabetes Mellitus and Risk Factors in Chinese Rural Population: The RuralDiab Study. *Sci Rep* (2016) 6:1–9. doi: 10.1038/srep31426
31. Blaak E. Sex Differences in the Control of Glucose Homeostasis. *Curr Opin Clin Nutr Metab Care* (2008) 11:500–4. doi: 10.1097/MCO.0b013e32830467d3
32. Fuente-Martín E, Argente-Arizón P, Ros P, Argente J, Chowen JA. Sex Differences in Adipose Tissue: It Is Not Only a Question of Quantity and Distribution. *Adipocyte* (2013) 2:128–34. doi: 10.4161/adip.24075
33. Xiao J, Shen C, Chu MJ, Gao YX, Xu GF, Huang JP, et al. Physical Activity and Sedentary Behavior Associated With Components of Metabolic Syndrome Among People in Rural China. *PloS One* (2016) 11:e0147062. doi: 10.1371/journal.pone.0147062
34. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, et al. Prevention of Type 2 Diabetes Mellitus by Changes in

Lifestyle Among Subjects With Impaired Glucose Tolerance. *New Engl J Med* (2001) 344:1343–50. doi: 10.1056/NEJM200105033441801
35. Group DPPR. Reduction in the Incidence of Type 2 Diabetes With Lifestyle Intervention or Metformin. *New Engl J Med* (2002) 346:393–403. doi: 10.1056/NEJMoa012512
36. Li G, Zhang P, Wang J, Gregg EW, Yang W, Gong Q, et al. The Long-Term Effect of Lifestyle Interventions to Prevent Diabetes in the China Da Qing Diabetes Prevention Study: A 20-Year Follow-Up Study. *Lancet* (2008) 371:1783–9. doi: 10.1016/S0140-6736(08)60766-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 30 November 2021
doi: 10.3389/fendo.2021.769549

Long-Term Physical Activity Participation and Subsequent Incident Type 2 Diabetes Mellitus: A Population-Based Cohort Study

Chenglong Li[1,2,3], *Yanjun Ma*[1,2,3], *Rong Hua*[1,2,3], *Fanfan Zheng*[4*] and *Wuxiang Xie*[1,2,3*]

[1] Peking University Clinical Research Institute, Peking University First Hospital, Beijing, China, [2] Peking University Clinical Research Institute Heart and Vascular Health Research Center at Peking University Shougang Hospital, Beijing, China, [3] Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing, China, [4] School of Nursing, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

OPEN ACCESS

Edited by:
Tao Huang,
Peking University, China

Reviewed by:
Lirong Liang,
Capital Medical University, China
Yutong Samuel Cai,
University of Leicester,
United Kingdom
Changwei Li,
Tulane University School of Public Health and Tropical Medicine,
United States

***Correspondence:**
Fanfan Zheng
fanfan@bjmu.edu.cn
Wuxiang Xie
xiewuxiang@hsc.pku.edu.cn

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: 02 September 2021
Accepted: 02 November 2021
Published: 30 November 2021

Citation:
Li C, Ma Y, Hua R, Zheng F and Xie W (2021) Long-Term Physical Activity Participation and Subsequent Incident Type 2 Diabetes Mellitus: A Population-Based Cohort Study. Front. Endocrinol. 12:769549. doi: 10.3389/fendo.2021.769549

Background: Uncertainty remains concerning association between long-term physical activity and incident type 2 diabetes mellitus (DM). We intended to evaluate physical activity participation over a 6-year span and assess association with subsequent 10-year incident DM risk, as well as examine mediation role by obesity.

Methods: A total of 9757 community-dwelling adults aged ≥ 50 years in England were included in the population-based cohort. Physical activity participation, including trajectories and cumulative participation were assessed using weighted ***Z*** score over a 6-year span from wave 1 (2002–2003) to wave 4 (2008–2009). Incident DM recorded over a 10-year span from wave 4 (2008–2009) to wave 9 (2018–2019) was outcome.

Results: 5 distinct activity trajectories were identified, including persistently low (N=3037, incident DM=282), initially low then improving (1868, 90), initially high then declining (325, 20), persistently moderate (2489, 170), and persistently high (2038, 108). Compared with persistently low, participants of initially low then improving, persistently moderate and high were associated with lower incident DM risk, with multivariable-adjusted hazard ratios (HR) of 0.41 (95% confidence interval [CI]: 0.32 to 0.53, $P<0.001$), 0.70 (95% CI: 0.56 to 0.89, $P=0.004$) and 0.49 (95% CI: 0.37 to 0.65, $P<0.001$), respectively. Elevated cumulative activity was also associated with lower DM risk, with each quintile increment in cumulative weighted ***Z*** score corresponding to HR of 0.76 (95% CI: 0.71 to 0.82, $P<0.001$). Mediation analysis found that body mass index, waist circumference and change in body mass index mediate 10% ($P<0.001$), 17% ($P<0.001$) and 9% ($P<0.001$) of the observed association between activity and incident DM, respectively.

Conclusions: For middle aged and older adults, both gradually improved and persistently active participation in physical activity were associated with subsequent lower risk of incident DM, with obesity playing a potential mediator. Strategies focusing on improving and maintaining active participation in physical activity might be beneficial from DM prevention perspective.

Keywords: prediction and prevention of type 2 diabetes, clinical science, epidemiology, exercise, obesity

INTRODUCTION

According to data from the Global Burden of Disease, the global annual incident type 2 diabetes mellitus (DM) has exceeded 20 million in the year 2017, increased by 30.5% compared with the year 2007 (1). The key role by physical activity in the prevention of DM has been well-established by previous evidence, from both interventional and observational studies (2–8). However, considering that physical activity could experience considerable alterations in long-term and the relatively low incidence of DM, longer period was needed for both comprehensive physical activity evaluation and assessment of its association with subsequent DM risk reduction. So evidence from interventional studies could not evaluate the long-term trajectories of physical activity and its long-term impact on DM risk reduction, due to limited follow-up period (3–6). And most observational studies only evaluated physical activity at single timepoint of study entry, without accounting for the longitudinal change and cumulative activity participation in a long-term period (7, 8). More importantly, few studies have incorporated the role of obesity into consideration when evaluating the association between physical activity and incident DM. As reported by previous studies, the association between changes in physical activity and incident DM could be attenuated, when further adjusting for obesity in follow-up (5, 9). So it was possible that obesity could play as a mediator between physical activity and incident DM, which remained an unanswered question.

Therefore, we intended to investigate the long-term physical activity participation during a 6-year span based on a life-span approach, by evaluating activity trajectories and cumulative participation, and assess its association with subsequent 10-year risk of incident DM among community-dwelling adults aged ≥ 50 years. We also aimed at exploring whether obesity, assessed by body mass index (BMI) and waist circumference, played the mediating role between long-term physical activity and incident DM.

MATERIALS AND METHODS

Study Population

The English Longitudinal Study of Ageing (ELSA) is an ongoing prospective and nationally representative cohort of community-dwelling adults aged ≥ 50 years in England. Details concerning objective, design, and method of the study have been published (10). The ELSA was approved by the London Multicentre Research Ethics Committee (MREC/01/2/91), with informed consent obtained for all participants. We used survey data from waves 1 (2002–2003) to 4 (2008–2009) to evaluate physical activity participation, and waves 4 (2008–2009) to 9 (2018–2019) to assess incident DM. Participants were excluded if they had been diagnosed of DM by wave 4. Study timeline was presented in **Supplemental Figure 1**.

Physical Activity Assessment

Biennial evaluation of frequency in participating mild, moderate and vigorous intensity physical activities was performed using three questions, differentiating from work related activities. Answers included 4 categories: 1) hardly ever or never; 2) one to three times per month; 3) once per week; 4) more than once per week. Further details were presented in **Supplemental Methods**.

A three-stage approach was used to assess physical activity participation. Firstly, we assigned a score of 1, 2, 3 according to frequency of participating in activities of mild, moderate and vigorous intensity, respectively. A score of 1 was assigned for hardly ever or never, 2 assigned for one to three times per month, 3 assigned for at least once per week. Score assignment was the same for the three types of activity (11). Secondly, for each of three types of activity, we calculated standardized ***Z*** score by subtracting corresponding mean and dividing by standard deviation (SD) of assigned scores at baseline. An activity ***Z*** score of 1 at a certain wave indicated that the particular score of physical activity at the wave was 1 SD higher than the average score of physical activity at baseline. Similar approach can be found for handling cognitive test scores by previous studies (12–14). Thirdly, to account for differences between mild, moderate and vigorous intensity activities, a weighted global activity ***Z*** score was summarized. The weights were selected based on metabolic equivalent of tasks (MET) estimates by 2011 Compendium of Physical Activities (15). After calculation, MET weights of 2.3, 4.4, and 7.2 were assigned to mild, moderate, and vigorous intensity activities ***Z*** score, respectively. Detailed information for weights calculation was presented in **Supplemental Table 1**, which was consistent with previous studies (16).

Obesity Assessment

The ELSA performed physical measurement at regular intervals, with standardized protocols applied, as described in **Supplemental Methods**. Measurement was conducted by

trained nurses for all participants in wave 0, 2, 4, 6, 8 (the year 1998, 2004, 2008, 2012, 2016). We used data in wave 4 to assess obesity, including BMI and waist circumference.

Ascertainment of DM and Prediabetes

According to the 2014 American Diabetes Association guidelines (ADA), DM was defined as an fasting plasma glucose ≥ 126 mg/dL (7.0 mmol/L) or an HbA1c level ≥ 48 mmol/mol (6.5%), a self-reported physician diagnosis of DM or current use of glucose-lowering therapy (17–19). Among participants without DM, we defined prediabetes as an fasting plasma glucose in the range 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) or an HbA1c level in the range 38.8 to 46.4 mmol/mol (5.7 to 6.4%) (17). The ELSA also collected information about the age being newly diagnosed of DM by physicians since last visit. Time to incident DM was calculated as length between age at wave 4 and age of firstly being diagnosed.

Covariates

Covariates included demographic and clinical variables assessed at wave 1. Demographic variables included sex, age (years), ethnicity, educational background (high-level education or not), cohabitation status (living alone or not), current smoking (yes or no), alcohol consumption (at least once per week). Clinical variables included overweight status, depressive symptoms, hypertension, prediabetes status, stroke, cardiovascular diseases, cancer, and chronic lung diseases. Considering that elder participants could experience decline in physical activity participation due to deteriorated mobility, we further adjusted for the mobility status.

We defined high-level education as above senior level of high school or 12 or more years of education. Mobility status was defined as whether participants reported any difficulties in daily living activities during waves 1 to 4. These activities included eating, bathing, dressing, getting in/out of bed and walking across a room, which could reflect the general mobility status. Depressive symptoms were defined based on an 8-item version of the Center for Epidemiologic Studies Depression Scale (CESD-8), with scores ≥4 regarded as having depressive symptoms (13, 20). Overweight status was defined by BMI (≥25 or <25 kg/m^2) according to World Health Organization (21). Personal history of other diseases was derived from records of self-reported physician diagnoses.

Statistical Analysis

For characteristics description, mean ± SD was used for continuous variables and numbers (percentage) for categorical variables. Overall differences between groups were tested using analysis of variance (ANOVA) or chi-square test.

We used weighted ***Z*** score of physical activity over a 6-year span from waves 1 to 4 to evaluate activity trajectories and cumulative participation. For physical activity trajectories, the group-based trajectory modeling (GBTM) was used to identify potential trajectories of participation in global, mild, moderate and vigorous activities, respectively. The GBTM used maximum likelihood estimation to identify participants sharing similar trajectory (22). It can handle data distributions including censored normal, Poisson and Bernoulli, and we used censored normal. Trajectory group of the highest probability was determined for each participant, which was then included in further multivariate analysis. We used SAS Proc Traj to fit GBTM models, with details described in **Supplemental Methods** (23). For cumulative participation, we used area under the curve (AUC) to calculate cumulative weighted physical activity ***Z*** score, based on Trapezoid rule (24). Details for AUC calculation were presented in **Figure 1**. Quintiles of cumulative weighted activity ***Z*** score were included in analysis as a numerical variable to perform linear trend test and mediation analysis.

Linear regression model was applied to assess associations of physical activity trajectories on subsequent BMI and waist circumference measured at wave 4. Cox regression was utilized to assess association of physical activity on risk of incident DM. Proportional hazard assumption was evaluated using weighted Schoenfeld residuals, with violated covariates included as model terms of interaction with time (25). To address ties in failure time, the method based on exact conditional probability under the proportional hazards assumption was applied (26). Hazard ratios (HR) and 95% Wald's confidence intervals (CI) were reported. Observations with missing values were excluded from analysis.

Mediation analysis was performed to evaluate the role of obesity in association between cumulative weighted activity ***Z*** score quintiles and incident DM. We included two mediators to represent the obesity status, including BMI and waist circumference measured at wave 4. We also assessed the mediating role of change in BMI from waves 0 to 4. As a result, three mediation models were built. Each mediation model comprised of two separate models:

1) the mediator model, denoted as:

$$\mathbf{X + C \rightarrow M}$$

where X represented quintiles of cumulative weighted global physical activity ***Z*** score (numerical) and C represented vector of covariates including sex, age, ethnicity, education, cohabitation status, mobility status, current smoking, alcohol consumption, depressive symptoms, hypertension, prediabetes status, stroke, cardiovascular diseases, chronic lung diseases and cancer, as well as BMI in wave 0 (when the mediator was change in BMI). M represented hypothesized mediators. After assessing normality distribution for BMI, change in BMI, and waist circumference, linear regression was used for fitting the mediator model;

2) the response model, denoted as:

$$\mathbf{X + M + C \rightarrow Y}$$

where new variable Y represented recorded incident DM from waves 4 to 9, with meaning for X, M and C remaining the same. Considering the incident DM of binary type, binary logistic regression was used for fitting the response model, with exponential transformed coefficient equivalent to odds ratio (OR). Quasi-Bayesian approximation method was used to

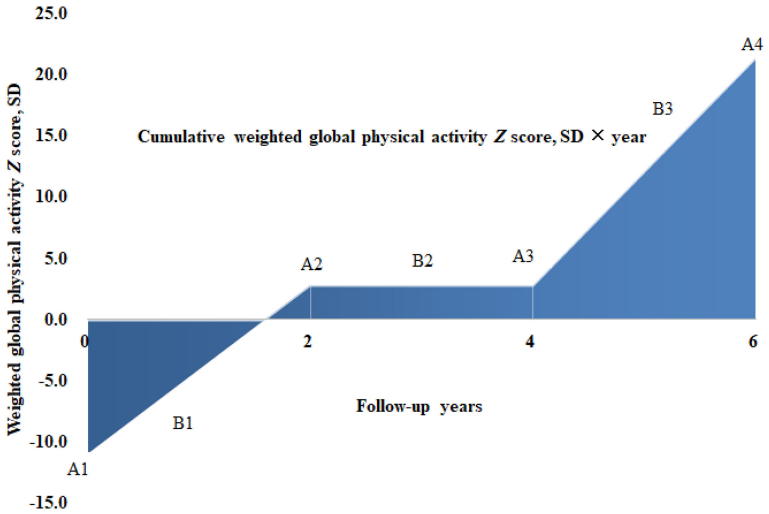


FIGURE 1 | Cumulative physical activity ***Z*** score calculation for 1 participant across 4 visits over a 6-year span in the ELSA. A1, A2, A3, A4 denotes weighted global physical activity ***Z*** score of 4 visits at wave 1 (year 2002), wave 2 (year 2004), wave 3 (year 2006) and wave 4 (year 2008), while B1, B2 and B3 indicates average activity ***Z*** score between consecutive visits of A1, A2, A3 and A4, respectively. According to trapezoid rule, the cumulative weighted activity ***Z*** score (SD × year), denoted as area under the curve in color blue, could be calculated as (B1 × 2y + B2 × 2y + B3 × 2y), with same approach applied to calculate cumulative mild, moderate and vigorous physical activity ***Z*** score over a 6-year span at the level of participants.

perform significance test for mediating effects, based on 2000 Monte Carlo replications (27).

Several sensitivity analyses were conducted. Firstly, we respectively identified trajectories of mild, moderate and vigorous activities, as well as calculated cumulative participation *Z* score. Then we evaluated associations between these sub-domain physical activity participation and outcomes. Secondly, we performed stratified analyses according to sex, age group (≥ 65 or <65 years), and overweight status (BMI < 25 or ≥ 25 kg/m^2). Thirdly, we excluded participants suffered any difficulties in activities of daily living during waves 1 to 4, and participants developed incident DM within two years after wave 4, to address reverse causation (28). Finally, we analyzed association between physical activity and incident DM recorded in 2 to 6 years after wave 4, to explore the potential time range of protective tole by physical activity on incident DM.

All statistical analyses were conducted using SAS 9.4 (SAS Institute, Cary, NC), and the package mediation in R language 3.6.2 (R Foundation, Vienna, Austria), with two-tailed alpha of 0.05 referred as statistically significant level.

RESULTS

Study Population

A total of 9757 diabetes-free participants from the ELSA were included for analysis. Detailed process of participant selection was shown in **Supplemental Figure 2**.

Among the included participants, 4243 (43.5%) participants were male, and the mean age was 58.9 ± 10.3 years. Differences were observed in most characteristics between activity trajectories, with details summarized in **Table 1**.

Long-Term Physical Activity Trajectories

As shown in **Figure 2**, 5 trajectories of global physical activity were identified, including: 1) persistently low trajectory (N=3037), with constantly low participation; 2) initially low then improving (N=1868), with initially low participation but turned to elevating afterwards; 3) initially high then declining (N=325), with initially high participation but turned to declining afterwards; 4) persistently moderate (N=2489), with constantly intermediate participation; 5) persistently high (N=2038), with highly active participation throughout the span.

Association of Physical Activity Participation on Obesity

As illustrated in **Table 2**, significant associations were observed of global physical activity trajectories and cumulative participation on subsequent obesity. Compared with the persistently low trajectory, participants in trajectories of initially low then improving, persistently moderate and high had significantly lower BMI of 1.121 kg/m^2 (95% CI: 0.777 to 1.465, P <0.001), 0.688 kg/m^2 (95% CI: 0.345 to 1.031, P < 0.001), and 1.441 kg/m^2 (95% CI: 1.070 to 1.813, P <0.001), respectively. And those with initially high then declining trajectory showed no significant differences compared with persistently low trajectory. Similar results were observed for waist circumference. Compared with the persistently low trajectory, participants in trajectories of initially low then improving, persistently moderate and high had significantly lower waist circumference of 3.183 cm (95% CI: 2.351 to 4.015, P <0.001), 2.367 cm (95% CI: 1.544 to 3.191, P < 0.001), and 4.740 cm (95% CI: 3.846 to 5.634, P <0.001), respectively.

As for cumulative physical activity participation, it was found that elevated participation was associated with both lower BMI

TABLE 1 | Baseline characteristics of study population according to physical activity trajectories.

Characteristics[a]	AllN=9757	Initially low then improvingN=1868	Persistently lowN=3037	Initially high then decliningN=325	Persistently moderateN=2489	Persistently highN=2038	P value[b]
Male (%)	4243 (43.5%)	923 (49.4%)	1202 (39.6%)	139 (42.8%)	988 (39.7%)	991 (48.6%)	<0.001
Age (years)	58.9 ± 10.3	52.4 ± 7.8	59.1 ± 11.7	66.1 ± 10.7	61.9 ± 9.5	59.3 ± 8.2	<0.001
White (%)	9445 (96.8%)	1787 (95.7%)	2888 (95.1%)	316 (97.2%)	2442 (98.1%)	2012 (98.7%)	<0.001
High level of education (%)	2444 (25.0%)	48 (2.6%)	222 (7.3%)	90 (27.7%)	940 (37.8%)	1144 (56.1%)	<0.001
Living alone (%)	1879 (19.3%)	46 (2.5%)	492 (16.2%)	106 (32.6%)	759 (30.5%)	476 (23.4%)	<0.001
Current smoking (%)	980 (10.0%)	29 (1.6%)	245 (8.1%)	56 (17.2%)	426 (17.1%)	224 (11.0%)	<0.001
Drinking ≥ once per week (%)	3800 (38.9%)	92 (4.9%)	498 (16.4%)	188 (57.8%)	1545 (62.1%)	1477 (72.5%)	<0.001
Hypertension (%)	3087 (31.6%)	64 (3.4%)	688 (22.7%)	192 (59.1%)	1272 (51.1%)	871 (42.7%)	<0.001
Prediabetes (%)	1028 (10.5%)	19 (1.0%)	177 (5.8%)	62 (19.1%)	412 (16.6%)	358 (17.6%)	<0.001
Cardiovascular disease (%)	508 (5.2%)	16 (0.9%)	171 (5.6%)	38 (11.7%)	180 (7.2%)	103 (5.1%)	<0.001
Chronic lung disease (%)	301 (3.1%)	5 (0.3%)	120 (4.0%)	20 (6.2%)	111 (4.5%)	45 (2.2%)	<0.001
Cancer (%)	306 (3.1%)	6 (0.3%)	68 (2.2%)	22 (6.8%)	114 (4.6%)	96 (4.7%)	<0.001
Body mass index (kg/m^2)	27.2 ± 4.4	27.2 ± 4.4	28.3 ± 5.1	27.6 ± 4.3	27.3 ± 4.4	26.5 ± 3.9	<0.001
Waist circumference (cm)	91.3 ± 12.7	92.6 ± 12.3	93.8 ± 13.5	92.3 ± 12.8	91.2 ± 12.6	89.9 ± 12.4	<0.001
Systolic blood pressure (mm Hg)	140.1 ± 19.4	137.3 ± 18.6	145.8 ± 21.3	142.8 ± 20.0	140.4 ± 19.3	136.6 ± 17.7	<0.001
Diastolic blood pressure (mm Hg)	77.5 ± 12.0	76.7 ± 12.2	77.9 ± 12.4	77.4 ± 11.4	77.8 ± 12.3	77.0 ± 11.6	0.193
Cumulative weighted global physical activity ***Z*** score (SD × year)	13.0 ± 56.7	-12.2 ± 28.2	-42.1 ± 17.8	16.6 ± 24.5	26.4 ± 19.8	101.2 ± 22.8	<0.001

[a]*Data are presented as mean ± SD, n (%).*
[b]*P value reported for differences between trajectory groups using analysis of variance, or chi-square test.*

and waist circumference, with each 10 units increment in cumulative weighted global activity *Z* score associated with decreased BMI of 0.109 kg/m^2 (95% CI: 0.085 to 0.133, P <0.001) and waist circumference of 0.348 cm (95% CI: 0.291 to 0.406, P <0.001).

Association Between Physical Activity Participation and Incident DM

During a 10-year span, 670 incident DM cases were reported, with results summarized in **Table 3**. Compared with the persistently low trajectory, participants in initially low then

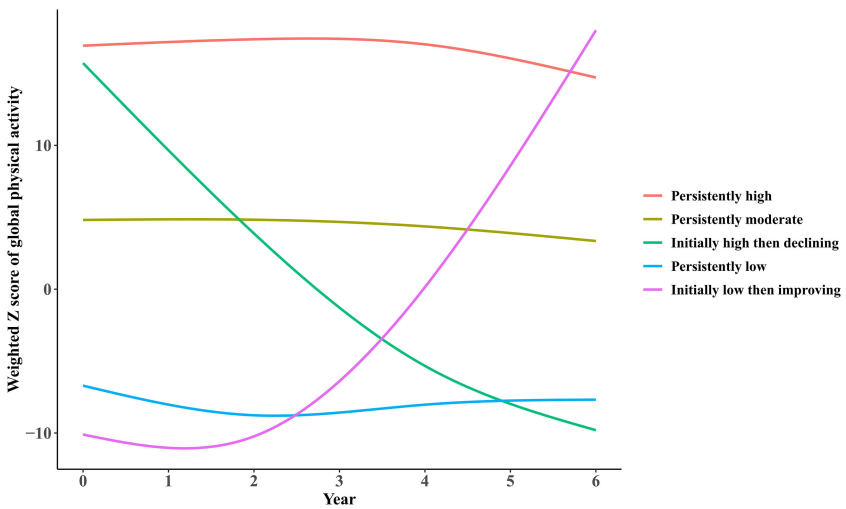


FIGURE 2 | Trajectories of participation in global physical activities by participants from the ELSA over a 6-year span.

TABLE 2 | Association of long-term physical activity participation on subsequent obesity assessed by BMI and waist circumference.

Physical activity participation	BMI (kg/m^2)		Waist circumference (cm)	
	β (95% CI)[a]	P value	β (95% CI)	P value
Global physical activity trajectories				
Persistently low	Reference		Reference	
Initially low then improving	-1.121 (-1.465, -0.777)	<0.001	-3.183 (-4.015, -2.351)	<0.001
Initially high then declining	-0.134 (-0.817, 0.549)	0.701	-1.487 (-3.101, 0.127)	0.071
Persistently moderate	-0.688 (-1.031, -0.345)	<0.001	-2.367 (-3.191, -1.544)	<0.001
Persistently high	-1.441 (-1.813, -1.070)	<0.001	-4.740 (-5.634, -3.846)	<0.001
Cumulative weighted global physical activity participation *Z* score (SD × year)				
Quintile 1	Reference		Reference	
Quintile 2	-0.810 (-1.179, -0.441)	<0.001	-2.620 (-3.506, -1.735)	<0.001
Quintile 3	-0.994 (-1.387, -0.601)	<0.001	-2.873 (-3.818, -1.927)	<0.001
Quintile 4	-1.396 (-1.786, -1.006)	<0.001	-4.221 (-5.161, -3.281)	<0.001
Quintile 5	-1.860 (-2.276, -1.444)	<0.001	-5.746 (-6.748, -4.745)	<0.001
Test for linear trend	-0.436 (-0.533, -0.339)	<0.001	-1.338 (-1.572, -1.104)	<0.001
Per 10 units increment	-0.109 (-0.133, -0.085)	<0.001	-0.348 (-0.406, -0.291)	<0.001

[a]*Adjusted covariates included sex, age, ethnicity, education, cohabitation status, mobility status, current smoking, alcohol consumption, depressive symptoms, overweight status, hypertension, prediabetes status, stroke, cardiovascular diseases, chronic lung diseases and cancer.*

improving, persistently moderate and high trajectories had significant lower DM risk, with HR of 0.41 (95% CI: 0.32 to 0.53, P <0.001), 0.70 (95% CI: 0.56 to 0.89, P =0.004) and 0.49 (95% CI: 0.37 to 0.65, P <0.001), respectively. By contrast, DM risk in initially high then declining trajectory showed no significant differences with persistently low trajectory. For cumulative activity participation, elevated cumulative weighted global activity ***Z*** score was also associated with lower DM risk, with each quintile and 10 units increment corresponding to HR of 0.76 (95% CI: 0.71 to 0.82, P <0.001) and 0.94 (95% CI: 0.92 to 0.95, P <0.001), respectively.

Mediation Analysis for Obesity

As shown in **Figures 3–5**, significant mediation roles by BMI, waist circumference, and change in BMI were identified in association between cumulative weighted global physical activity ***Z*** score quintiles and subsequent incident DM, after adjusting for other covariates. The total indirect effect through BMI (βi=-0.03, P <0.001), defined as product of indirect effect 1 (β1) and indirect effect 2 (β2) reached significance level, which mediated 10% (P <0.001) of overall association. Similarly, the total indirect effect through waist circumference (βi=-0.05, P <0.001) and change in BMI (βi=-0.02, P <0.001) were significant, mediating 17% (P <0.001) and 9% (P <0.001) of overall association, respectively.

Sensitivity Analyses

Similar with global physical activity, 5 trajectories were identified for both mild and moderate activities, as shown in **Supplemental Figures 3, 4**, respectively. For vigorous activities, a total of 4 trajectories were identified, shown in **Supplemental Figure 5**.

Associations of sub-domain physical activity participation on obesity, and incident DM risk were presented in **Supplemental Tables 2, 3**, respectively. For obesity, consistent results with

TABLE 3 | Association of long-term physical activity participation on subsequent incident DM.

Physical activity participation	Events/Total	Model 1[a]		Model 2[b]	
		HR (95% CI)	P value	HR (95% CI)	P value
Global physical activity trajectories					
Persistently low	282/3037	Reference		Reference	
Initially low then improving	90/1868	0.34 (0.26, 0.43)	<0.001	0.41 (0.32, 0.53)	<0.001
Initially high then declining	20/325	1.16 (0.73, 1.84)	0.531	0.75 (0.47, 1.20)	0.226
Persistently moderate	170/2489	1.04 (0.84, 1.29)	0.736	0.70 (0.56, 0.89)	0.004
Persistently high	108/2038	0.67 (0.52, 0.87)	0.003	0.49 (0.37, 0.65)	<0.001
Cumulative weighted global physical activity participation Z score (SD × year)					
Quintile 1	182/2015	Reference		Reference	
Quintile 2	136/1888	0.68 (0.54, 0.87)	0.002	0.57 (0.45, 0.73)	<0.001
Quintile 3	140/1951	0.89 (0.70, 1.12)	0.321	0.51 (0.40, 0.66)	<0.001
Quintile 4	107/1949	0.58 (0.45, 0.75)	<0.001	0.38 (0.29, 0.50)	<0.001
Quintile 5	105/1954	0.58 (0.44, 0.78)	<0.001	0.34 (0.25, 0.46)	<0.001
Test for linear trend	–	0.87 (0.82, 0.93)	<0.001	0.76 (0.71, 0.82)	<0.001
Per 10 units increment	–	0.97 (0.95, 0.98)	<0.001	0.94 (0.92, 0.95)	<0.001

[a]*Adjusted for age × time, sex, ethnicity and education.*

[b]*Additionally adjusted for cohabitation status, mobility status, current smoking, alcohol consumption, depressive symptoms, overweight status, hypertension, prediabetes status, stroke, cardiovascular diseases, chronic lung diseases and cancer.*

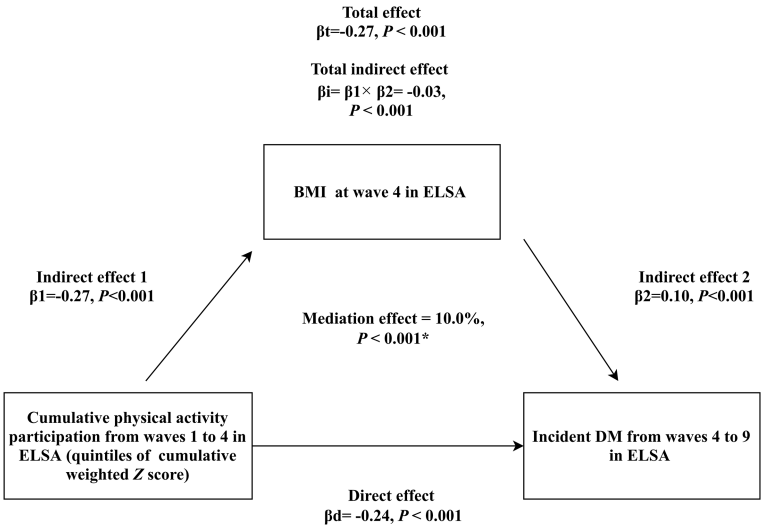


FIGURE 3 | Mediation analysis assessing the mediating role of BMI measured at wave 4 in the association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. βd: coefficient for direct association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. βi: coefficient for total indirect association of BMI. βt: coefficient for total association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. *: adjusted covariates included sex, age, ethnicity, education, cohabitation status, mobility status, current smoking, alcohol consumption, depressive symptoms, hypertension, prediabetes status, stroke, cardiovascular diseases, chronic lung diseases and cancer.

global activity participation were observed for all three mild, moderate and vigorous intensity activities. For incident DM, consistent results were mainly observed for moderate and vigorous intensity activities participation.

Stratified analyses by sex (**Supplemental Tables 4, 5**), age (**Supplemental Tables 6, 7**), and overweight status (**Supplemental Tables 8, 9**) showed that similar results were consistently observed. After further excluding participants with

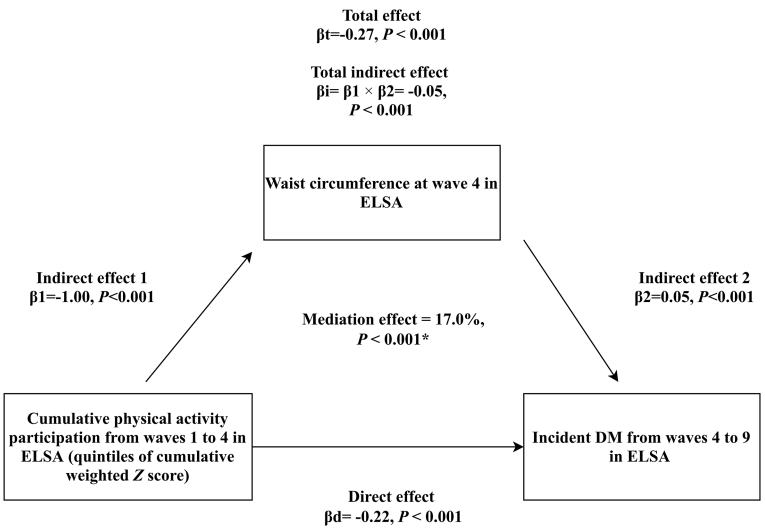


FIGURE 4 | Mediation analysis assessing the mediating role of waist circumference measured at wave 4 in the association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. βd: coefficient for direct association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. βi: coefficient for total indirect association of waist circumference. βt: coefficient for total association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. *: adjusted covariates included sex, age, ethnicity, education, cohabitation status, mobility status, current smoking, alcohol consumption, depressive symptoms, hypertension, prediabetes status, stroke, cardiovascular diseases, chronic lung diseases and cancer.

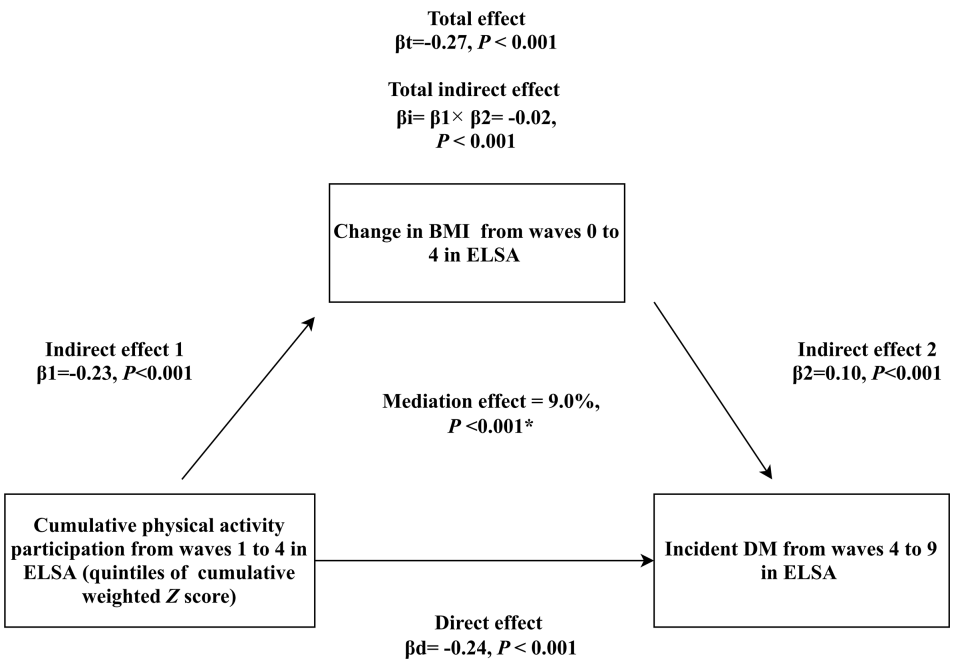


FIGURE 5 | Mediation analysis assessing the mediating role of change in BMI from waves 0 to 4 in the association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. βd: coefficient for direct association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. βi: coefficient for total indirect association of change in BMI. βt: coefficient for total association of cumulative weighted global physical activity ***Z*** score quintiles on subsequent incident DM. *: adjusted covariates included sex, age, ethnicity, education, cohabitation status, mobility status, current smoking, alcohol consumption, BMI in wave 0, depressive symptoms, hypertension, prediabetes status, stroke, cardiovascular diseases, chronic lung diseases and cancer.

difficulties in daily living activities or developed incident DM within two years after wave 4, our main results preserved, shown in **Supplemental Tables 10, 11**. For physical activity trajectories, significant associations with incident DM risk reduction in 2 to 6 years were consistently observed for the persistently moderate group, shown in **Supplemental Table 12**. The initially low then improving group was not associated with incident DM risk in 2 years but associated with lower incident DM risk in 4 to 6 years after wave 4, while no significant associations were observed for the initially high then declining and persistently high groups in this span.

DISCUSSION

Our study presented several major findings. Firstly, persistently active participation in physical activity was associated with a lower risk of subsequent incident DM. Secondly, for those with initially low activity, gradually improved participation was also associated with subsequent lower DM risk. These findings were encouraging, indicating that taking measures towards being physically active could never be too late for DM prevention. According to WHO recommendations, even for adults aged ≥ 65, health benefits could still be gained by maintaining adequate activity (29). And our findings further illustrated DM prevention benefits. Finally, both overall obesity assessed by BMI and central obesity by waist circumference significantly mediated association between activity and incident DM, with central obesity in particular. The implication was that obesity could be crucial to understand the link between activity and DM. According to previous studies, central obesity was as important as overall obesity in predicting DM risk, and our study further consolidated its significance by illustrating potential mediation role linking long-term activity with subsequent DM risk (30–32).

To our current knowledge, this is the first prospective study to investigate both trajectories and cumulative participation of physical activity over a 6-year span, and assess its association with subsequent 10-year incident DM, as well as the mediation role by obesity, with findings implicative of potential DM prevention significance from active physical participation.

Several previous studies also discussed association between long-term activity, obesity and DM. In a cohort study evaluating longitudinal trajectories of metabolic control from childhood to young adulthood in DM patients, researchers used GBTM approach to identify longitudinal trajectories of HbA1c and evaluated related factors (33). They found that elevated frequency of participating in physical activity was associated with lower probability of belonging to HbA1c increase trajectories, illustrating potential significance of being physically active (33). Another cohort study recruiting Chinese adults aged 20 to 80 years with impaired fasting glucose found that an inverse dose–response relationship between leisure-time physical activity and incident DM risk (34). The researchers also

found that 19.2% incident DM could be avoided if those inactive participants could have engaged in WHO recommendation levels of leisure-time physical activity (34). In another study investigating association between BMI trajectories and DM among women aged between 30 and 100 years, researchers found that leisure time physical activity attenuated the general deleterious effect of obesity on DM, especially for women of the consistently non-obese trajectory (35). Although these findings showed similarity with ours, few of them directly evaluated long-term physical activity participation and the longitudinal association of on subsequent incident DM. In another 5-year cohort study of older adults aged ≥ 50 years in China, researchers assessed association between onset of DM and long-term physical activity trajectories (36). They found that, participants with longer cumulative years in participating physical activity experienced a later onset of obesity and DM, compared to their sedentary counterparts. They also evaluated the association of activity on risk of onset of obesity and DM, but failed to observe significant results (36). Despite that, the study indicated the significance of long-term active participation in physical activity in delaying onset of DM and obesity, which was consistent with our findings. Another study including South Koreans over 18 years old also found that both sufficient baseline physical activity level and its temporal increase were associated with a lower risk of incident DM, compared with those with temporal decrease in activity (37). Regardless of these shared findings in protective role of long-term active activity, our findings concerning the mediation role by BMI and waist circumference still require verification by further prospective studies. Although studies reported the mediation role by central obesity in linking genetic predisposition with incident DM risk, whether same mediation role maintained for long-term physical activity remain unresolved (38).

Our study found consistent association of moderate and vigorous intensity activities on incident DM, demonstrating significance of regular moderate-to-vigorous activities participation for DM prevention. The findings were also consistent with recommendations by ADA concerning DM prevention (39).

Our study possessed several strengths. Firstly, we simultaneously assessed trajectories and cumulation of participation in physical activity over a 6-year span, and prospectively assessed its associations with subsequent 10-year incident DM risk. Such design enabled stronger capability of drawing casual conclusions. Secondly, we used the GBTM approach to explore all possible pattern of physical activity trajectories over a 6-year span, instead of merely assuming monotonic trajectories. The approach could efficiently incorporate repeated measurements of activity participation at multiple timepoints, thus could address limitations by only considering activity participation at single time-point. Thirdly, our study evaluated the mediation role by obesity, which enabled better understanding of the observed association between long-term physical activity and DM. Finally, our study population was nationally-representative of community-dwelling adults aged ≥ 50 years in England, with large sample-size and long follow-up period.

Several limitations also required attention. Firstly, we only considered frequency when evaluating physical activity participation, without accounting for duration. Besides, only reported frequency was used, thus the issue of reporting bias remains non-neglectable. Secondly, the majority of participants included was of White ethnicity, which restricted generalization to other ethnicities. Thirdly, the challenging issue of reverse causation remains. Chances were that some previous diagnosed diseases or pre-clinical changes in metabolic status could impact regular physical activity participation, and lead us to biased conclusions. Nevertheless, considering the long follow-up period of our study, influence by reverse causation on our findings could be limited (40). Fourthly, due to data restrictions, we could not account for the potential impact by dietary pattern on incident DM. Giving the significant role by diet in incident DM prevention, such missing consideration could leave our findings to potential bias (18). Finally, due to the nature of observational study, influence by unmeasured confounding factors on our results could not be eliminated, impeding further steps towards conclusive casual relationships (41).

In summary, we found that for middle aged and older adults, both gradually improved and persistently active participation in physical activity were associated with subsequent lower risk of incident DM, with obesity playing a potential mediator. Strategies focusing on improving and maintaining active participation in physical activity might be beneficial from DM prevention perspectives.

DATA AVAILABILITY STATEMENT

Original survey dataset from the ELSA are freely available to all bonafide researchers. Access to data can be obtained by visiting their websites (https://www.elsa-project.ac.uk/). The data can also be obtained on request (xiewuxiang@hsc.pku.edu.cn).

ETHICS STATEMENT

The ELSA was approved by the London Multicentre Research Ethics Committee (MREC/01/2/91). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CL contributed to formal analysis and writing original draft. YM and RH contributed to data curation and manuscript editing efforts. FZ and WX conceptualized the study design and funding acquisition, as well as manuscript reviewing and editing efforts. All authors had full access to the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. FZ and WX and are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have

been omitted. All authors contributed to the article and approved the submitted version.

FUNDING

The present study was supported by the National Natural Science Foundation of China (project no. 81974490) and 2019 Irma and Paul Milstein Program for Senior Health Research Project Award. The funders had no role in the study design; the collection, analysis, and interpretation of data; the writing of the manuscript; or the decision to submit the article for publication.

ACKNOWLEDGMENTS

We appreciate efforts made by the original data creators, depositors, copyright holders, the funders of the data collections and their contributions for access of data from the English Longitudinal Study of Ageing (waves 1–9).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2021.769549/full#supplementary-material

REFERENCES

1. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, Regional, and National Incidence, Prevalence, and Years Lived With Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* (2018) 392:1789–858. doi: 10.1016/S0140-6736(18)32279-7
2. Colberg SR, Sigal RJ, Yardley JE, Riddell MC, Dunstan DW, Dempsey PC, et al. Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association. *Diabetes Care* (2016) 39:2065–79. doi: 10.2337/dc16-1728
3. Zhang Y, Pan X-F, Chen J, Xia L, Cao A, 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* (2020) 63:21–33. doi: 10.1007/s00125-019-04985-9
4. Pan X-R, Li G-W, Hu Y-H, Wang J-X, Yang W-Y, An Z-X, et al. Effects of Diet and Exercise in Preventing NIDDM in People With Impaired Glucose Tolerance: The Da Qing IGT and Diabetes Study. *Diabetes Care* (1997) 20:537–44. doi: 10.2337/diacare.20.4.537
5. Kriska AM, Rockette-Wagner B, Edelstein SL, Bray GA, Delahanty LM, Hoskin MA, et al. The Impact of Physical Activity on the Prevention of Type 2 Diabetes: Evidence and Lessons Learned From the Diabetes Prevention Program, a Long-Standing Clinical Trial Incorporating Subjective and Objective Activity Measures. *Diabetes Care* (2021) 44:43–9. doi: 10.2337/dc20-1129
6. Harris T, Limb ES, Hosking F, Carey I, DeWilde S, Furness C, et al. Effect of Pedometer-Based Walking Interventions on Long-Term Health Outcomes: Prospective 4-Year Follow-Up of Two Randomised Controlled Trials Using Routine Primary Care Data. *PLoS Med* (2019) 16:e1002836. doi: 10.1371/journal.pmed.1002836
7. Cleven L, Krell-Roesch J, Nigg CR, Woll A. The Association Between Physical Activity With Incident Obesity, Coronary Heart Disease, Diabetes and Hypertension in Adults: A Systematic Review of Longitudinal Studies Published After 2012. *BMC Public Health* (2020) 20:726. doi: 10.1186/s12889-020-08715-4
8. Bennett DA, Du H, Bragg F, Guo Y, Wright N, Yang L, et al. Physical Activity, Sedentary Leisure-Time and Risk of Incident Type 2 Diabetes: A Prospective Study of 512 000 Chinese Adults. *BMJ Open Diabetes Res Care* (2019) 7:e000835. doi: 10.1136/bmjdrc-2019-000835
9. Hamman RF, Wing RR, Edelstein SL, Lachin JM, Bray GA, Delahanty L, et al. Effect of Weight Loss With Lifestyle Intervention on Risk of Diabetes. *Diabetes Care* (2006) 29:2102–7. doi: 10.2337/dc06-0560
10. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort Profile: The English Longitudinal Study of Ageing. *Int J Epidemiol* (2013) 42:1640–8. doi: 10.1093/ije/dys168
11. Hamer M. Physical Activity and Trajectories in Cognitive Function: English Longitudinal Study of Ageing. *Physiol Behav* (2018) 176:1570–3. doi: 10.1038/s41395-018-0061-4
12. Zheng F, Yan L, Yang Z, Zhong B, Xie W. HbA1c, Diabetes and Cognitive Decline: The English Longitudinal Study of Ageing. *Diabetologia* (2018) 61:839–48. doi: 10.1007/s00125-017-4541-7
13. Zheng F, Zhong B, Song X, Xie W. Persistent Depressive Symptoms and Cognitive Decline in Older Adults. *Br J Psychiatry* (2018) 213:638–44. doi: 10.1192/bjp.2018.155
14. Hua R, Ma Y, Li C, Zhong B, Xie W. Low Levels of Low-Density Lipoprotein Cholesterol and Cognitive Decline. *Sci Bull* (2021) 16:1684–90. doi: 10.1016/j.scib.2021.02.018
15. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DRJ, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: A Second Update of Codes and MET Values. *Med Sci Sports Exercise* (2011) 43:1575–81. doi: 10.1249/MSS.0b013e31821ece12
16. Pang Y, Lv J, Kartsonaki C, Yu C, Guo Y, Du H, et al. Association of Physical Activity With Risk of Hepatobiliary Diseases in China: A Prospective Cohort Study of 0.5 Million People. *Br J Sports Med* (2021) 55:1024–33. doi: 10.1136/bjsports-2020-102174
17. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* (2014) 37 Suppl 1:S81–90. doi: 10.2337/dc14-S081
18. Hu Y, Ding M, Sampson L, Willett WC, Manson JE, Wang M, et al. Intake of Whole Grain Foods and Risk of Type 2 Diabetes: Results From Three Prospective Cohort Studies. *BMJ (Clin Res ed)* (2020) 370:m2206. doi: 10.1136/bmj.m2206
19. Schneider ALC, Pankow JS, Heiss G, Selvin E. Validity and Reliability of Self-Reported Diabetes in the Atherosclerosis Risk in Communities Study. *Am J Epidemiol* (2012) 176:738–43. doi: 10.1093/aje/kws156
20. Li C, Miles T, Shen L, Shen Y, Liu T, Zhang M, et al. Early-Life Exposure to Severe Famine and Subsequent Risk of Depressive Symptoms in Late Adulthood: The China Health and Retirement Longitudinal Study. *Br J Psychiatry* (2018) 213:579–86. doi: 10.1192/bjp.2018.116
21. WHO. Obesity and Overweight. Available at: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.
22. Nagin D. *Group-Based Modeling of Development*. Cambridge, MA and London, England: Harvard University Press (2009). doi: 10.4159/9780674041318
23. Jones BL, Nagin DS. Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them. *Sociological Methods Res* (2007) 35:542–71. doi: 10.1177/0049124106292364
24. Jenkins LM, Garner CR, Kurian S, Higgins JP, Parrish TB, Sedaghat S, et al. Cumulative Blood Pressure Exposure, Basal Ganglia, and Thalamic Morphology in Midlife. *Hypertens (Dallas Tex: 1979)* (2020) 75:1289–95. doi: 10.1161/HYPERTENSIONAHA.120.14678
25. Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika* (1994) 81:515–26. doi: 10.1093/biomet/81.3.515
26. DeLong DM, Guirguis GH, So YC. Efficient Computation of Subset Selection Probabilities With Application to Cox Regression. *Biometrika* (1994) 81:607–11. doi: 10.2307/2337133
27. Imai K, Keele L, Tingley D, Yamamoto T. "Causal Mediation Analysis Using R," in: *Advances in Social Science Research Using R*. Volume 196. (2010). pp. 129–54. doi: 10.1007/978-1-4419-1764-5_8
28. Biswas A, Gilbert-Ouimet M, Mustard CA, Glazier RH, Smith PM. Combined Associations of Work and Leisure Time Physical Activity on Incident Diabetes Risk. *Am J Prev Med* (2021) 60:e149–58. doi: 10.1016/j.amepre.2020.09.017

29. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 Guidelines on Physical Activity and Sedentary Behaviour. *Br J Sports Med* (2020) 54:1451–62. doi: 10.1136/bjsports-2020-102955
30. Ohnishi H, Saitoh S, Takagi S, Katoh N, Chiba Y, Akasaka H, et al. Incidence of Type 2 Diabetes in Individuals With Central Obesity in a Rural Japanese Population: The Tanno and Sobetsu Study. *Diabetes Care* (2006) 29:1128–9. doi: 10.2337/diacare.2951128
31. Nyamdorj R, Qiao Q, Söderberg S, Pitkäniemi JM, Zimmet PZ, Shaw JE, et al. BMI Compared With Central Obesity Indicators as a Predictor of Diabetes Incidence in Mauritius. *Obes (Silver Spring Md)* (2009) 17:342–8. doi: 10.1038/oby.2008.503
32. Krishnan S, Rosenberg L, Djoussé L, Cupples LA, Palmer JR. Overall and Central Obesity and Risk of Type 2 Diabetes in U.S. Black Women. *Obes (Silver Spring Md)* (2007) 15:1860–6. doi: 10.1038/oby.2007.220
33. Schwandt A, Hermann JM, Rosenbauer J, Boettcher C, Dunstheimer D, Grulich-Henn J, et al. Longitudinal Trajectories of Metabolic Control From Childhood to Young Adulthood in Type 1 Diabetes From a Large German/Austrian Registry: A Group-Based Modeling Approach. *Diabetes Care* (2017) 40:309–16. doi: 10.2337/dc16-1625
34. Lao XQ, Deng H-B, Liu X, Chan T-C, Zhang Z, Chang L-Y, et al. Increased Leisure-Time Physical Activity Associated With Lower Onset of Diabetes in 44 828 Adults With Impaired Fasting Glucose: A Population-Based Prospective Cohort Study. *Br J Sports Med* (2019) 53:895–900. doi: 10.1136/bjsports-2017-098199
35. Werneck AO, Oyeyemi AL, Szwarcwald CL, Sardinha LB, Silva DR. Body Mass Index Trajectories and Noncommunicable Diseases in Women: The Role of Leisure Time Physical Activity. *Am J Hum Biol* (2021) 33:e23492. doi: 10.1002/ajhb.23492
36. Zhou P, Hughes AK, Grady SC, Fang L. Physical Activity and Chronic Diseases Among Older People in a Mid-Size City in China: A Longitudinal Investigation of Bipolar Effects. *BMC Public Health* (2018) 18:486. doi: 10.1186/s12889-018-5408-7
37. Lee J-Y, Ryu S, Sung K-C. Association of Baseline Level of Physical Activity and its Temporal Changes With Incident Hypertension and Diabetes Mellitus. *Eur J Prev Cardiol* (2018) 25:1065–73. doi: 10.1177/2047487318774419
38. Huang T, Qi Q, Zheng Y, Ley SH, Manson JE, Hu FB, et al. Genetic Predisposition to Central Obesity and Risk of Type 2 Diabetes: Two Independent Cohort Studies. *Diabetes Care* (2015) 38:1306–11. doi: 10.2337/dc14-3084
39. American Diabetes Association. 3. Prevention or Delay of Type 2 Diabetes: Standards of Medical Care in Diabetes-2020. *Diabetes Care* (2020) 43:S32–6. doi: 10.2337/dc20-S003
40. Tarp J, Hansen BH, Fagerland MW, Steene-Johannessen J, Anderssen SA, Ekelund U. Accelerometer-Measured Physical Activity and Sedentary Time in a Cohort of US Adults Followed for Up to 13 Years: The Influence of Removing Early Follow-Up on Associations With Mortality. *Int J Behav Nutr Phys Act* (2020) 17:39. doi: 10.1186/s12966-020-00945-4
41. Streeter AJ, Lin NX, Crathorne L, Haasova M, Hyde C, Melzer D, et al. Adjusting for Unmeasured Confounding in Non-Randomised Longitudinal Studies: A Methodological Review. *J Clin Epidemiol* (2017) 87:23–34. doi: 10.1016/j.jclinepi.2017.04.022

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 10 December 2021
doi: 10.3389/fendo.2021.777552

Association Between the Severity of Diabetic Retinopathy and Optical Coherence Tomography Angiography Metrics

Binxin Xu[1,2†], Jiahui Chen[3,4,5†], Shaohua Zhang[3,4,5], Shengli Shen[6], Xuan Lan[6], Zhineng Chen[7], Zhiqiang Yan[1,2*] and Bingxiang Xu[8,9,10*]*

[1] Human Phenome Institute, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Institute of Brain Science, School of Life Sciences, Fudan University, Shanghai, China, [2] Institute of Molecular Physiology, Shenzhen Bay Laboratory, Shenzhen, China, [3] Department of Ophthalmology and Vision Science, Eye and ENT Hospital of Fudan University, Shanghai, China, [4] NHC Key Laboratory of Myopia (Fudan University), Laboratory of Myopia, Chinese Academy of Medical Sciences, Shanghai, China, [5] Key Laboratory of Visual Impairment and Restoration of Shanghai, Shanghai, China, [6] Human Phenome Institute, Fudan University, Shanghai, China, [7] School of Computer Science, Fudan University, Shanghai, China, [8] School of Kinesiology, Shanghai University of Sport, Shanghai, China, [9] State Key Laboratory of Translational Medicine and Innovative Drug Development, Jiangsu Simcere Diagnostics Co., Ltd., Nanjing, China, [10] Nanjing Simcere Medical Laboratory Science Co., Ltd., Nanjing, China

OPEN ACCESS

Edited by:
Changwei Li,
Tulane University School of Public Health and Tropical Medicine, United States

Reviewed by:
Wuxiang Xie,
Peking University, United States
Hao Peng,
Soochow University, China

***Correspondence:**
Zhineng Chen
zhinchen@fudan.edu.cn
Zhiqiang Yan
zqyan@fudan.edu.cn
Bingxiang Xu
bingxiang.xu@hotmail.com

[†]These authors share first authorship

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: 15 September 2021
Accepted: 11 November 2021
Published: 10 December 2021

Citation:
Xu B, Chen J, Zhang S, Shen S, Lan X, Chen Z, Yan Z and Xu B (2021) Association Between the Severity of Diabetic Retinopathy and Optical Coherence Tomography Angiography Metrics. Front. Endocrinol. 12:777552. doi: 10.3389/fendo.2021.777552

Diabetic retinopathy, the most serious ocular complication of diabetes, imposes a serious economic burden on society. Automatic and objective assessment of vessel changes can effectively manage diabetic retinopathy and prevent blindness. Optical coherence tomography angiography (OCTA) metrics have been confirmed to be used to assess vessel changes. The accuracy and reliability of OCTA metrics are restricted by vessel segmentation methods. In this study, a multi-branch retinal vessel segmentation method is proposed, which is comparable to the segmentation results obtained from the manual segmentation, effectively extracting vessels in low contrast areas and improving the integrity of the extracted vessels. OCTA metrics based on the proposed segmentation method were validated to be reliable for further analysis of the relationship between OCTA metrics and diabetes and the severity of diabetic retinopathy. Changes in vessel morphology are influenced by systemic risk factors. However, there is a lack of analysis of the relationship between OCTA metrics and systemic risk factors. We conducted a cross-sectional study that included 362 eyes of 221 diabetic patients and 1,151 eyes of 587 healthy people. Eight systemic risk factors were confirmed to be closely related to diabetes. After controlling these systemic risk factors, significant OCTA metrics (such as vessel complexity index, vessel diameter index, and mean thickness of retinal nerve fiber layer centered in the macular) were found to be related to diabetic retinopathy and severe diabetic retinopathy. This study provides evidence to support the potential value of OCTA metrics as biomarkers of diabetic retinopathy.

Keywords: diabetic retinopathy, diabetes, optical coherence tomography angiography, retinal vessel segmentation, retinal vessel quantification, systemic risk factors

INTRODUCTION

Diabetic retinopathy (DR), is one of the most frequent causes of blindness in the working-age population (1). DR causes expression of vascular endothelial growth factor (2), inflammation (3), upregulation of renin–angiotensin (4), oxidative stress (5), activation of protein kinase C (6), the formation of sorbitol (7), and advanced glycation end-products (8), which lead to an increased permeability of retinal vessel, capillary non-perfusion area, microaneurysm formation, and retinal endothelial cell damage. DR also causes retinal neurodegeneration (9, 10) and predicts an increased risk of life-threatening systemic vascular complications (11).

Optical coherence tomography (OCT) is a new non-contact and non-invasive imaging technique. Combined with the optical coherence tomography angiography (OCTA), it allows a layered view of the vascular morphology and blood flow alterations of the retina and choroid, enabling a more in-depth study of specific retinal capillary layer lesions caused by DR (12). Several studies have reported that OCTA metrics were associated with the severity of DR, such as foveal avascular zone (FAZ) area and vessel density, indicating the potential value of OCTA for managing DR progress (13–15). Vessel segmentation is a necessary step for quantifying OCTA metrics (such as vessel density, vessel skeleton fractal dimension, and vessel skeleton density). However, the retinal vessel segmentation methods used in these studies have two glaring limitations: i) discontinuity of segmented vessels, and ii) poor and false segmentation of low-contrast vessel regions.

Most published OCTA-based retinal vessel segmentation methods can be divided into the following kinds: local adaptive thresholds, hessian filter, and non-local means filter. Reif et al. (16) combined the low-pass filter and the local adaptive threshold to achieve the segmentation of vascular. Chu et al. (17) used the global threshold to remove noise, the hessian matrix to enhance the image, and the adaptive threshold to segment blood vessels. The two methods previously mentioned have low complexity and can achieve available segmentation results, but their anti-noise ability is poor. The segmentation method implemented by Tang et al. (14) prevented the enhancement of noise by a non-local means filter and segmented vessels by using a phansalkar adaptive local thresholding. However, the method proposed by Tang et al. (14) has high computational complexity and is prone to introduce false-positive results. Due to the amount of noise present in the OCTA images, the poor quality of the capillaries, the requirement for continuity of the segmented vessels, and the existing vessel segmentation methods do not achieve satisfactory results.

In response to the above limitations, this study developed a multi-branch retinal vessel segmentation method (abbreviated as MRVSM) that does not enhance noise and can effectively extract low-contrast capillaries. MRVSM designed special methods and corresponding parameters to segment the large vessels and capillaries separately since the pixel intensities vary considerably between vessels.

Reviewing the previous studies on the relationship between OCTA metrics and the severity of DR, a comprehensive analysis of the effect of systemic risk factors on OCTA metrics is still lacking. Consideration of systemic risk factors is crucial for the reliable study of the relationship between the OCTA metrics and the severity of DR. In this study, we analyzed the relationship of OCTA metrics to 36 systemic risk factors. After controlling the identified systemic risk factors, we studied whether the OCTA metrics were associated with DR and severe DR. This study can identify eyes at high risk of DR progression, which facilitates early diagnosis and timely intervention for DR.

MATERIALS AND METHODS

Participants

The study followed the ethical standards set out in the 1964 Declaration of Helsinki and was approved by the Ethics Committee of Eye and ENT Hospital of Fudan University and the Ethics Committee of School of Life Sciences of Fudan University with informed consent forms.

Approximately 596 healthy participants (1,192 eyes) were eligible for the recruitment criteria. Among them, 100 healthy participants were recruited from the Eye and ENT Hospital of Fudan University and 496 healthy participants were recruited from the Human Phenome Institute of Fudan University. Approximately 240 diabetes participants (418 eyes) were recruited from the Eye and ENT Hospital of Fudan University. After the image quality control, the study excluded 97 eyes (reasons shown in **Supplementary Figure 1**). Finally, the study included 587 healthy participants (1,151 eyes) and 221 diabetes participants (362 eyes). Approximately 54 healthy participants and 68 diabetes participants from the Eye and ENT Hospital of Fudan University and 496 from the Human Phenome Institute of Fudan University underwent systemic test, including the fasting blood test, urine test, blood pressure test, pulse test, vision test, and measurements of demographic variables. All the participants measured for demographic variables and blood glucose and were identified as diabetes participants or healthy participants according to the diagnostic criteria proposed by the American Diabetes Association (18). The severity of DR was graded by ophthalmologists from the Eye and ENT Hospital of Fudan University according to the Diabetic Retinopathy Early Treatment Study (19).

The characteristics of the included population are presented in **Table 1**. Among the healthy participants (mean [SD] age, 38 [14] years), 54% were female and among the diabetes participants (mean [SD] age, 70 [4] years), 57% were female.

The recruitment criteria of healthy participants were: (1) between 20 and 85 years of age, (2) long-term residence in Shanghai, (3) no adverse lifestyle habits, (4) no major illnesses, and (5) self-reported healthy.

The recruitment criteria of diabetes participants were: (1) between 20 and 85 years of age, and (2) having diabetes.

The image quality control with the following exclusion criteria: (1) signal intensity less than 6, (2) presence of motion

TABLE 1 | The characteristics of study population.

Characteristics	Unit	Healthy participants	Diabetes participants
		Mean (SD) or %	Mean (SD) or %
Age	Year	38 (14)	69 (4)
Gender	Female	54%	57%
	Male	46%	43%
Eyes	Left eyes	50%	49%
	Right eyes	50%	51%
Severity of diabetic retinopathy	No DR		34%
	Mild		48%
	Moderate		11%
	Severe		7%

artifacts, tomography, and dislocation, (3) image center deviation from the macula,(4) inaccurate vessel delamination, and (5) low image resolution.

Measurement of Systemic Risk Factors

The systemic test included fasting blood test, urine test, blood pressure test, pulse test, vision test, and measurements of demographic variables. Indicators of fasting blood test were basophil count, basophil ratio, eosinophil count, eosinophil ratio, hematocrit, hemoglobin, lymphocyte count, lymphocyte ratio, mean corpuscular hemoglobin, mean cell hemoglobin concentration, monocyte count, mean platelet volume, plateletcrit, platelet distribution width, platelet larger cell ratio, platelet count, red cell distribution width coefficient of variation, neutrophilic granulocyte count, neutrophilic granulocyte ratio, creatinine, total cholesterol, glucose, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, total bilirubin, triglyceride, uric acid, and glycosylated hemoglobin. Indicators of blood pressure test were diastolic blood pressure and systolic blood pressure. The indicators of urine tests, pulse test, and vision test are urine specific gravity, pulse, and uncorrected visual acuity respectively. Demographic variables include age, sex, height, and weight. Height and weight were converted to body mass index (BMI) and uncorrected visual acuity was converted to logarithm of the minimum angle of resolution for analysis.

Data Collection Based on OCTA and OCT

Swept-source OCT (DRI OCT Triton; Topcon Inc, Tokyo, Japan) was used in our study. The DRI OCT uses infrared light at a wavelength of 1,050 nm as the light source and achieves a high A-scan rate of 100 kHz, resulting in high signal penetration and fast imaging. An OCT scan of a 3 mm x 3 mm field of view takes approximately 4 s. The lateral resolution is 20 μm and the axial resolution is 8 μm. DRI OCT uses an en face visualization technology to reconstruct the blood flow signal in three dimensions. The embedded software IMAGEnet 6 can divide the retinal into superficial capillary plexus (SCP) and deep capillary plexus (DCP). SCP is segmented from the inner limiting membrane (offset 2.6 μm) to the junction (offset 15.6 μm) of the inner plexiform layer and inner nuclear layer. DCP is 15.6 μm down from the junction of the inner plexiform layer and the inner core layer to 70.2 μm.

Both eyes of every participant have been performed with OCT scans centered on the macula and optic disc, OCTA imaging centered on the macula with 3 mm x 3 mm area, and color fundus photograph. Among the 100 healthy participants from the Eye and ENT Hospital of Fudan University, 37 of them underwent four repeated OCTA imaging within 10 min to analyze the reliability of the segmentation method proposed in this article. All the OCTA images of the participants were submitted to the technicians for quality control.

Segmentation and Quantification of the Retinal Vessels for OCTA Images

The OCTA images of SCP were exported in TIF format and automatically segmented. As depicted in **Figure 1**, MRVSM included a large vessel segmentation branch (LVSB), a capillary segmentation branch, and their fusion module. Considering that capillaries are difficult to distinguish from the background, we firstly segmented large vessels and then capillaries, where the segmentation methods are designed separately and their parameters adjusted accordingly. The benefit of this process is that it effectively reduces the segmentation error rate of capillaries and better suppresses noises.

Line detectors with different window sizes were used to distinguish between large vessels and capillaries, which was proposed in (20). The line detector computed the line strength of the target pixel which was defined as the largest average gray levels along the line going through the target pixel over the 12 directions (15° of the angular interval from 0°) in the set square window and was used to distinguish between vessel pixel and background pixel. The line detector was sensitive to different vessels according to the window size. In large vessel segmentation, the line detector with a window size of 25 x 25 pixels were first used to extract the structural features of large vessels. The linear combination of structural features extracted by the line detector and the intensity of vessels was binarized to roughly separate the large vessels and the backgrounds. The linear combination coefficients were 1/3 for structural features and 2/3 for intensity. The detected background intensity was halved to improve the contrast between the large vessels and the background. Vessel pixels and background ones were separated using K-means clustering according to their gray levels (21). The separation was simplified since the large vessels were enhanced.

In the capillary segmentation, we set the gray value of the detected large vessel pixels to 45 to eliminate the interference of large vessels. As the gray levels of the capillary pixels are close to the background pixels, pre-processing is required to enhance the

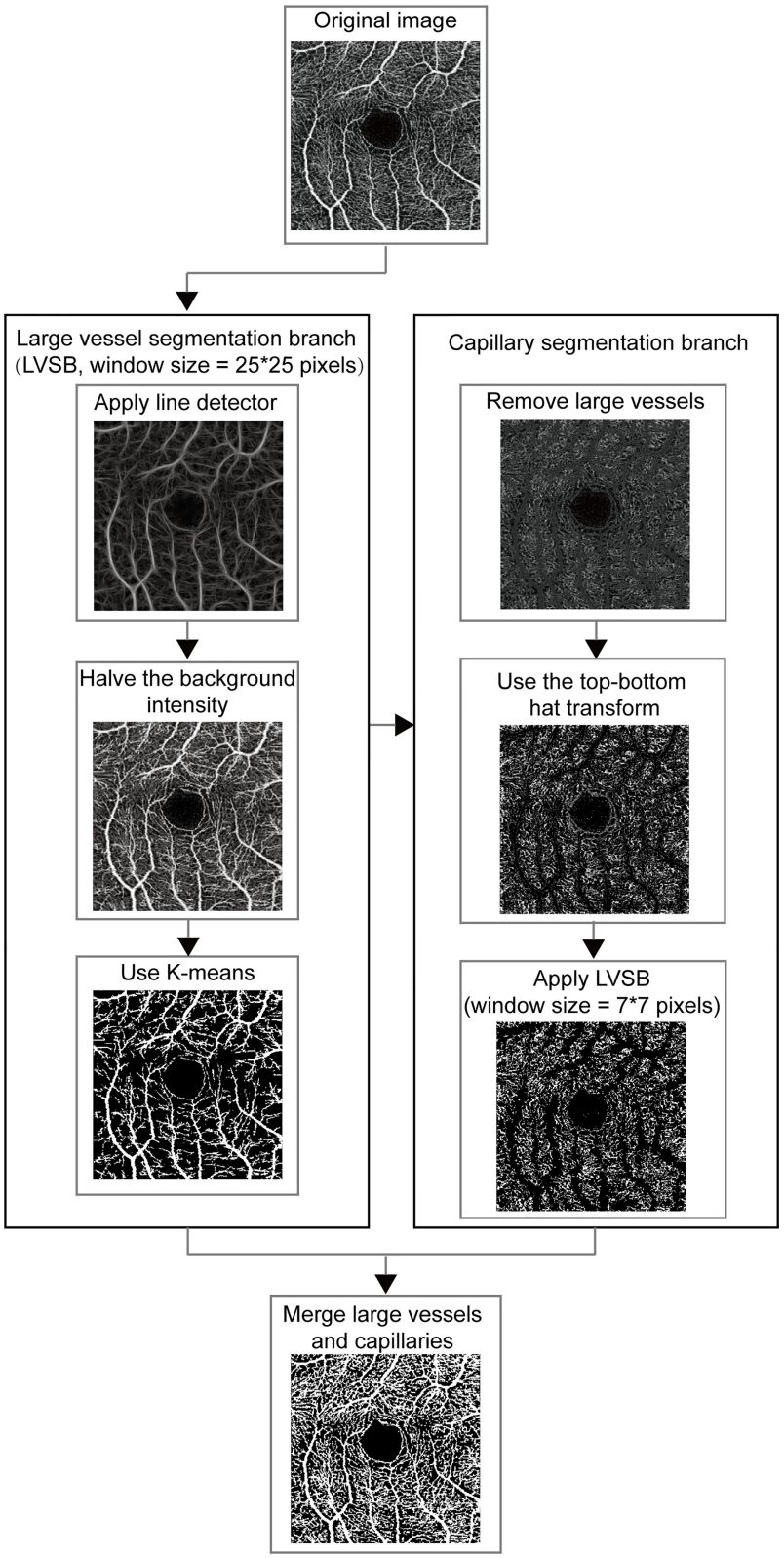


FIGURE 1 | The framework for vessel segmentation method on OCTA image. LVSB, large vessel segmentation branch.

capillaries. The top–bottom hat transform was first used to highlight capillary details and inhibit background which added the top hat transform effect and subtracted the bottom hat transform effect (22). Then we used the same process as LVSB, integrating line detector and k-means clustering to segment capillaries. In order to fit the capillary, the window size of the line detector in the capillary segmentation branch was smaller. It was empirically set to 7 x 7 pixels for capillary segmentation. In the fusion module, large vessels and capillaries, which are the foreground of the segmented images obtained from the two branches, were merged to generate the segmentation result.

Quantification of OCTA Metrics of Retinal Vessel

The vessels on the OCTA images were segmented by MRVSM automatically, FAZ was detected by the method in (23), and a non-perfusion zone was detected by the method in (24). Approximately 24 OCTA metrics were quantified based on the extracted blood vessels, FAZ, and non-perfusion zone. Among the OCTA metrics, vessel skeleton fractal dimension was defined in (16). Vessel density, vessel skeleton density, vessel perimeter index, vessel diameter index, and vessel complexity index were defined in (17). Big vessel density, big vessel skeleton density, small vessel density, small vessel skeleton density, non-perfusion area, FAZ circularity index, and FAZ area were defined in (24). Vessel density in 1 mm circle and vessel density in 2 mm circle were defined in (25). FAZ perimeter, FAZ major axis diameter, FAZ minor axis diameter, and FAZ orientation were defined in (26). FAZ horizontal diameter and FAZ vertical diameter were defined in (27). Besides, we quantified three OCTA metrics, including vessel skeleton density in the 1 mm circle, vessel skeleton density in the 2 mm circle, and FAZ eccentricity. The above OCTA metrics were defined in the supplementary material (**Supplementary Material**).

Other OCTA metrics automatically calculated by the instrument included center vessel density in the superficial layer (CVDS), superior vessel density in the superficial layer, inferior vessel density in the superficial layer, nasal vessel density in the superficial layer, temporal vessel density in the superficial layer, center vessel density in the deep layer, superior vessel density in the deep layer (SVDD), inferior vessel density in the deep layer (IVDD), nasal vessel density in the deep layer, and temporal vessel density in the deep layer, where they were within the circle with a 2.5-mm diameter centered on the center of foveal avascular zone, optic disc area, optic disc volume, optic cup area, optic cup volume, optic rim area, optic rim volume, the area ratio of cup and disc, the square root of the area ratio of cup and disc, the vertical diameter ratio of cup and disc, disc vertical diameter, disc horizontal diameter, mean thickness of macular in the retina layer, mean thickness of macular in the retinal nerve fiber layer (MTMRNFL), mean thickness of macular in the ganglion cell layer with inner plexiform layer (MTMGCL+), mean thickness of macular in the ganglion cell layer with inner plexiform layer and nerve fiber layer, mean thickness of macular in the choroidal layer, mean thickness of optic disc in the retina layer, mean thickness of optic disc in the retinal nerve fiber layer (MTODRNFL), mean thickness of optic disc in the ganglion cell layer with inner plexiform layer, mean thickness of optic disc in the ganglion cell layer with inner plexiform layer and nerve fiber layer (MTODGCL++), and mean thickness of optic disc in the choroidal layer (MTODCL).

Evaluation and Statistical Analysis

We used each eye with the inclusion criteria as a subject for analysis and calculated its OCTA metrics. All statistical analyses were performed in R. To ensure the quality of the data, histograms, box plots and scatter plots were made for quality control and checked by visual inspection. The data were preprocessed to replace the null values with the mean of other samples in the same group. To assess the reliability and reproducibility of the OCTA metrics based on MRVSM, Within-subject standard deviation (Sw) (28), coefficient of variation (CoV, 100 × Sw/overall mean) (28), and intraclass correlation coefficient (ICC) (29) were calculated. To assess the interocular correlation of the OCTA metrics, t-test was used to analyze whether the OCTA metrics were statistically different between eyes. Then, Pearson correlation coefficient, ICC, and Bland–Altman plots (30) were used to evaluate the interocular correlation. Univariate linear regression models with OCTA metrics as the dependent variables and systemic risk factors as independent variables were used to assess the effect of systemic risk factors on OCTA metrics. The significances (P-values) of univariate linear regression were corrected for FDR (31). For each diabetes case, we choose one matched control from the healthy group with replacement to make the identified systemic risk factors of the two cases as closely as possible. The distance of the two individuals was measured by Euclidean distance. Then, Wilcox test and FDR correction were used to test for differences in OCTA metrics between diabetic and healthy controls. Stepwise regression with Akaike information criterion (AIC) as objective function which repeatedly added most important independent variable or removed variables that are not important and highly correlated with other variables was used to build the model. Here, we used stepwise regression to determine the important OCTA metrics which were the independent variables of the logistic regression model to classify the DR participants and diabetes participants free of DR. Similarly, the OCTA metrics related to severe DR were found by stepwise regression which were the independent variables of the logistic regression model to classify severe DR participants and non-severe DR participants. The prediction powers of the models were measured by averaging the receiver operating characteristic curves (ROC), the area under the curve (32) (AUC) values, and the F_1 scores (33) under the 5-fold cross validation.

RESULTS

Assessment of MRVSM

Manually annotated were 11 OCTA images as ground truth and the performance of MRVSM was evaluated. The evaluation

metrics included sensitivity, specificity, accuracy, the absolute vessel volume difference (34), the dice similarity coefficient (35), and the Matthews correlation coefficient (36). Results show that MRVSM is close to the manual segmentation and outperforms state-of-the-art methods such as Reif et al. (16), Chu et al. (17), and Tang et al. (14) (**Figure 2A**). Apart from that, MRVSM can extract the blood vessels better in the low contrast area (**Figure 2B**). Compared with the state-of-the-art ones, MRVSM has the highest accuracy rate and the lowest error rate, which illustrates the advanced nature of MRVSM (**Figure 2C**). The performance evaluation data of the segmentation methods are shown in the supplementary material (**Supplementary Table 1**).

Reliability and Repeatability Analysis of OCTA Metrics Based on MRVSM

In the reliability analysis, the OCTA metrics based on MRVSM included big vessel density, small vessel density, big vessel skeleton density, small vessel skeleton density, vessel density, vessel skeleton density, vessel skeleton fractal dimension, vessel perimeter index, vessel diameter index, non-perfusion area, vessel complexity index, vessel skeleton density in the 1 mm circle, vessel density in 1 mm circle, vessel skeleton density in the 2 mm circle, and vessel density in 2 mm circle. Since Chu et al. (17), Reif et al. (16), and Tang et al. (14) cannot separate large and small blood vessels, we did not analyze the reliability and repeatability of big vessel skeleton density, small vessel skeleton density, big vessel skeleton density, and small vessel skeleton density. The eyes of 37 healthy participants from the Eye and ENT Hospital of Fudan University underwent four repeated OCTA imaging within 10 min. Sw and CoV were used to evaluate the repeatability of OCTA metrics, where low Sw and low CoV mean high repeatability. ICC was used to evaluate the reliability of OCTA metrics, where high ICC means high reliability. For the repeatability analysis of OCTA metrics, MRVSM had the lowest scores of Sw in vessel skeleton density

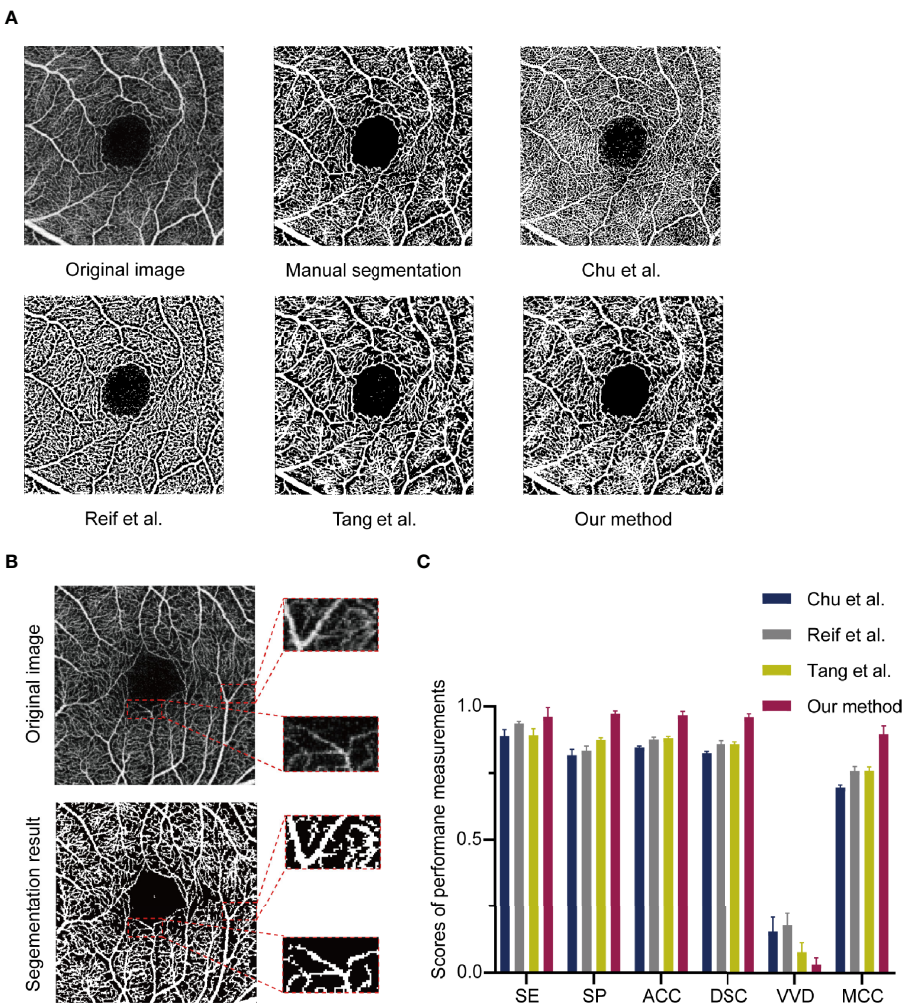


FIGURE 2 | The assessment of the multi-branch retinal vessel segmentation method (MRVSM). **(A)** Comparison of results of segmentation methods. **(B)** Segmentation results of MRVSM in the low contrast region. **(C)** The scores of sensitivity (SE), specificity (SP), accuracy (ACC), absolute vessel volume difference (VVD), dice similarity coefficient (DSC), and Matthews correlation coefficient (MCC) of the segmentation methods. Data were presented as the mean ± SD.

and vessel skeleton density in the 1 mm circle of the right eyes (**Supplementary Table 2**). **Figure 3** shows the scores of ICC of the OCTA metrics. For the vessel perimeter index and the vessel skeleton density in the 1 mm circle of the left eyes, and vessel perimeter index, vessel skeleton density, non-perfusion area, vessel skeleton density in the 1 mm circle, vessel skeleton density in the 2 mm circle, and vessel density in 2 mm circle of the right eyes, MRVSM has the highest scores of ICC (**Figure 3**). In addition, the ICC ranking of other OCTA quantitative metrics based on MRVSM is not at the bottom (**Figure 3**).

The Relationship Between Systemic Risk Factors and OCTA Metrics

Approximately 1,077 eyes of 550 healthy participants who underwent the measurement of systemic risk factors were used for the analysis of the relationship between systemic risk factors and all OCTA metrics. These participants have the following characteristics: females account for 55%, and age ranges from 20 to 79 years (mean [SD] age, 36 [14] years). Systemic risk factors associated with OCTA metrics were identified by the univariate linear regression model (F test). The results of the above correlation analysis were subjected to the Kolmogorov–Smirnov test to confirm the symmetry of the residuals of the univariate linear regression model and the FDR correction to adjust the p-values of the univariate linear regression model. Found to be associated with OCTA metrics were 28 systemic risk factors (**Figure 4**). The number of systemic risk factors which were strongly associated with OCTA metrics was calculated as 8 by the method described in (37). The eight systemic risk factors are uncorrected visual acuity, age, red cell distribution width coefficient of variation, hemoglobin, sex, creatinine, glucose, and uric acid. The results of the univariate linear regression model, Kolmogorov–Smirnov test, and FDR correction used to analyze the relationship between OCTA metrics and systemic risk factors are shown in **Supplementary Table 3**.

Analysis of OCTA Metrics Related to Diabetes

For each diabetes case, we choose one matched control from the healthy group with replacement to make the eight identified systemic risk factors of the two cases as closely as possible. Approximately 115 eyes of 68 diabetes participants (56% women; mean [SD] age, 70 [4] years) and 115 eyes of 81 healthy people (53% women; mean [SD] age, 36 [13] years) were used to analyze the relationship between OCTA metrics and diabetes. Approximately 34 OCTA metrics were found to be significantly different between the diabetes case and healthy case by Wilcox test and FDR correction (**Supplementary Table 4**). **Figure 5** shows the Wilcox results of non-perfusion area, FAZ circularity index, and vessel perimeter index which had the close relationships with diabetes (FDR-corrected P-values <0.001).

Analysis of OCTA Metrics Related to the Severity of DR

We further analyzed whether the above 34 OCTA metrics were related to DR and severe DR. This analysis was based on the 362 eyes of the 221 diabetes participants. Of the 362 eyes with diabetes, 172 (48%) eyes were mild non-proliferative diabetic retinopathy (NPDR), 41 (11%) eyes were moderate NPDR, and 27 (7%) eyes were severe NPDR.

SVDD, FAZ area, FAZ major axis diameter, FAZ minor axis diameter, vessel diameter index, MTMRNFL, vessel perimeter index, IVDD, vessel complexity index, and vessel skeleton density were found to be related to DR by stepwise regression (**Figure 6A**). The logistic regression model with the above metrics as independent variables further proved that they were

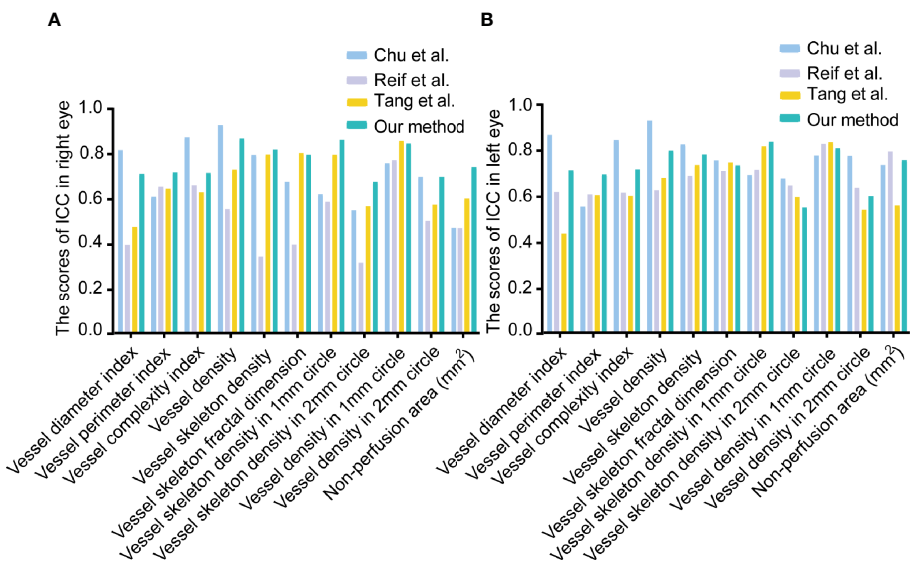


FIGURE 3 | The scores of intraclass correlation coefficient (ICC) of OCTA quantitative metrics. **(A)** The scores of ICC in the right eyes. **(B)** The scores of ICC in the left eyes.

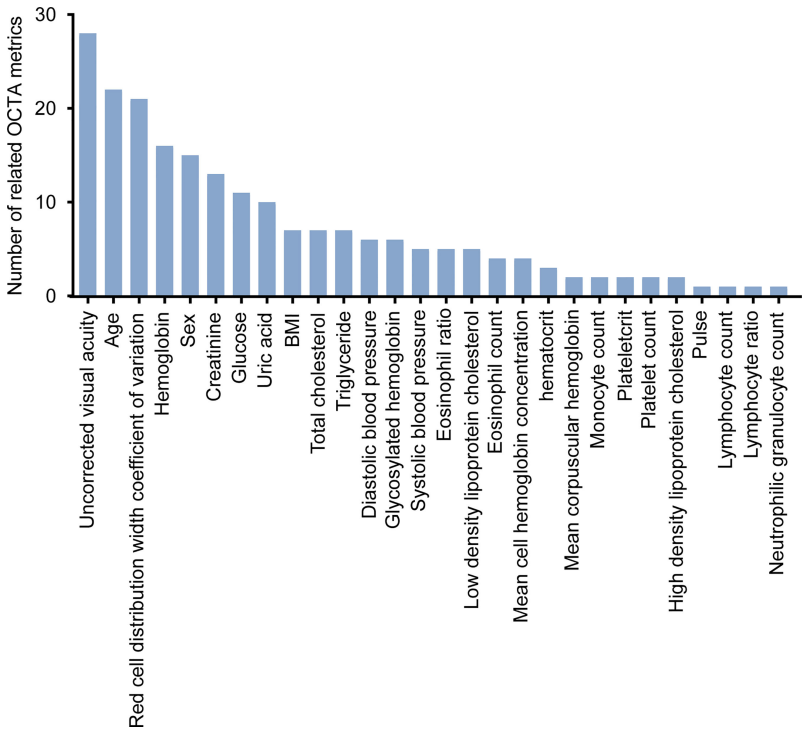


FIGURE 4 | The systemic risk factors associated with OCTA metrics. FDR-corrected P-values were calculated by Wilcox test.

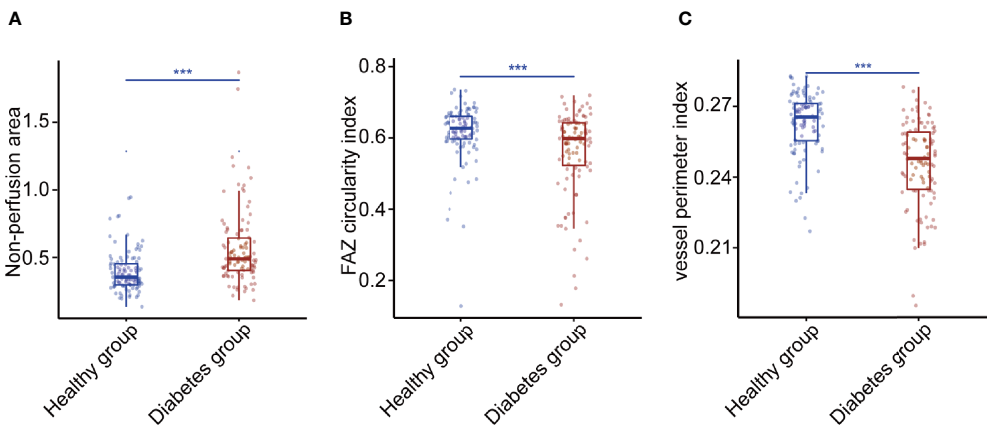


FIGURE 5 | Three OCTA metrics closely related to diabetes. **(A)** The Wilcox result of non-perfusion area. **(B)** The Wilcox result of foveal avascular zone circularity index. **(C)** The Wilcox result of vessel perimeter index. P-values were corrected by FDR. *p < 0.05, **p < 0.01, and ***p < 0.001. FAZ, foveal avascular zone.

closely related to DR (**Figure 6B**). To prove that the above OCTA metrics did correlate with DR, we first made a logistic regression using the systemic risk factors as independent variables to distinguish DR participants with diabetes participants free of DR as a baseline. A stepwise regression scheme was used to determine which systemic risk factors were included in the model. The model got an average AUC of 0.77 and an F_1 score of 0.73 under 5-fold cross validation (**Figure 7A**). Then we added the above OCTA metrics to the baseline model. Another round of stepwise regression was applied to rebuild the model. The performances of the models were improved (with AUC improved into 0.81 and F_1 score improved into 0.87, **Figure 7B**). These results that OCTA metrics selected in this section did correlate to the DR, even after the information of them encoded in systemic metrics deduced.

Figure 8 shows the ROC curve of the OCTA metrics associated with severe DR. There was a strong correlation between severe DR and vessel skeleton density, rim area, disc

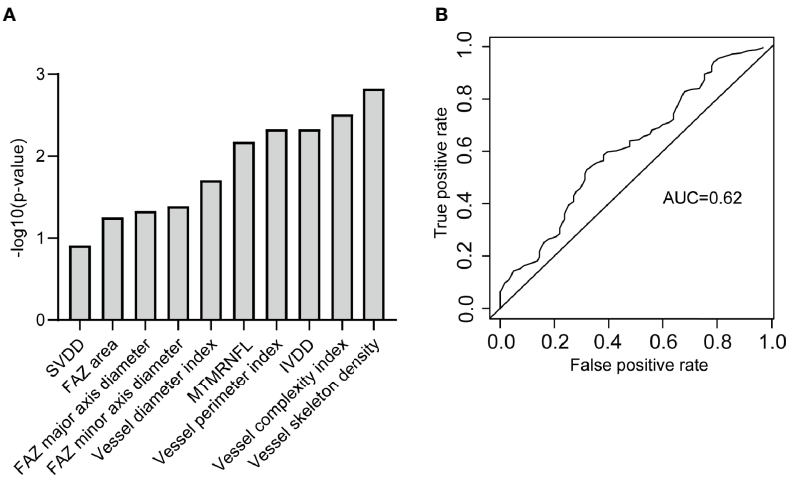


FIGURE 6 | The OCTA metrics related to DR. **(A)** The OCTA metrics related to DR and identified by stepwise regression. **(B)** The ROC curve of the logistic regression model to classify DR participants and diabetic free of DR participants. SVDD, superior vessel density in the deep layer; FAZ, foveal avascular zone; MTMRNFL, mean thickness of macular in the retinal nerve fiber layer; IVDD, inferior vessel density in the deep layer.

area, cup area, MTODRNFL, MTMGCL+, MTODCL, vertical diameter ratio of cup and disc, MTODGCL++, vessel skeleton fractal dimension, vessel complexity index, vessel diameter index, small vessel density, small vessel skeleton density, MTMRNFL, CVDS, and vessel density in 1 mm circle (**Figure 8A**). The logistic regression model constructed from these OCTA metrics can effectively distinguish severe DR participants from the non-severe DR participants with an accuracy of 0.80 and an AUC of 0.84 (**Figure 8B**).

DISCUSSION

In this study, we developed a multi-stage vessel segmentation method (MRVSM) which achieved effective segmentation on low-quality capillary and was comparable to manual segmentation. The reliability of OCTA metrics based on MRVSM is not inferior to that of OCTA metrics based on Reif et al. (16), Chu et al. (17), and Tang et al. (14). This provides a basis for further analysis of the relationship between OCTA metrics and diabetes and DR. We found eight systemic risk factors that are strongly correlated with OCTA metrics. After controlling for the eight systemic risk factors, we found 34 OCTA metrics related to diabetes, such as non-perfusion area, FAZ circularity index, and vessel perimeter index. Among the 34 OCTA metrics, 10 OCTA metrics (such as FAZ area, vessel diameter index, mean thickness of retinal nerve fiber layer centered in the macular) were found to be related to DR which can distinguish DR participants from diabetes participants free of DR even after the prediction power of systemic risk factors discounted. It is noteworthy that 17 OCTA metrics (such as vessel complexity index, vessel diameter index, and mean

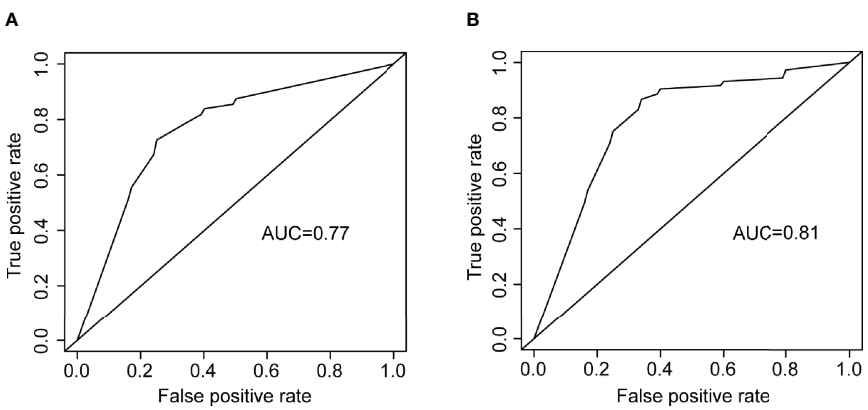


FIGURE 7 | The ROC curves of two logistic regression models for distinguishing DR participants from diabetes participants free of DR. **(A)** the logistic regression using systemic risk factors as independent variables. **(B)** the logistic regression using OCTA metrics as independent variables.

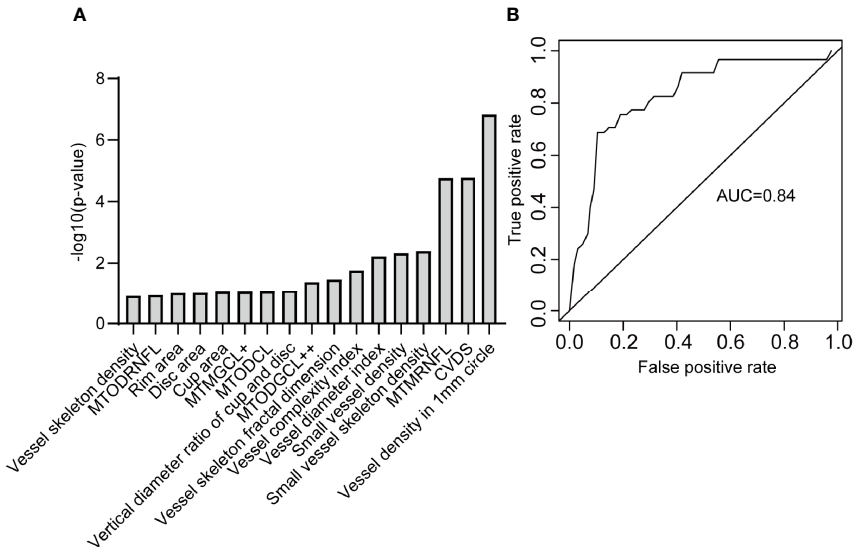


FIGURE 8 | The OCTA metrics related to severe DR. **(A)** The OCTA metrics related to severe DR and identified by stepwise regression. **(B)** The ROC curve of the logistic regression model to distinguish severe DR participants from non-severe DR participants. MTODRNFL, mean thickness of optic disc in the retinal nerve fiber layer; MTMGCL+, mean thickness of macular in the ganglion cell layer with inner plexiform layer; MTODCL, mean thickness of optic disc in the choroidal layer; MTODGCL++, mean thickness of optic disc in the ganglion cell layer with inner plexiform layer and nerve fiber layer; MTMRNFL, mean thickness of macular in the retinal nerve fiber layer; CVDS, center vessel density in the superficial layer.

thickness of retinal nerve fiber layer centered in the macular) were found to be related to severe DR. The model constructed from the 17 OCTA metrics can effectively distinguish the severe DR from non-severe DR. This study provides evidence to support the potential value of OCTA metrics in assessing and monitoring DR progression.

Although previous works had studied the OCTA metrics and the severity of DR, our study has precise automated blood vessel segmentation software, comprehensive quantification of OCTA metrics, a large number of samples, comprehensive physical examination data, and strict image quality control standards. Since the proposed segmentation method is not inferior to existing segmentation algorithms, it has better application prospects and provides a new idea in vessel segmentation fields. The OCTA metrics found related to DR can help clinicians to screen DR patients, thus facilitating early detection of DR patients and delaying the development of DR. It is noteworthy that the OCTA metrics found related to severe DR facilitate timely detection and therapeutic intervention for patients with severe NPDR, which can prevent blindness. The above conclusions are beneficial for people to better understand the changes in the retinal fundus caused by DR. The significant difference of OCTA metrics in different severity of DR is conducive to monitoring and management of DR, and has the potential to become a biomarker of DR.

There are still some shortcomings that need improvement. The OCTA image size of the macula in this paper is relatively small (3 mm x 3 mm), leading to the inability to obtain some vessel quantitative metrics. Large size means low resolution, so a balance between size and resolution is needed. In the follow-up study, we will increase the subject's population diversity and extend MRVSM to perform on large-scale images. Second, we only included high-resolution images, which may introduce bias. Third, the ratio of patients with moderate DR, severe DR is relatively small. We will recruit more participants with larger diversity in the follow-up studies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Eye and ENT Hospital of Fudan University and the Ethics Committee of School of Life Sciences of Fudan University. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

BingX, ZY, and ZC conceived the project and designed the study. JC and SZ recruited volunteers from the Eye and ENT Hospital of Fudan University to collect their data. JC graded the severity of DR on OCTA images. SS and XL collected volunteers' OCT and OCTA data at the Human Phenome Institute of Fudan

University. BinX proposed the automatic segmentation method and analyzed all the data, and wrote the manuscript. BingX, ZY, ZC, BinX, JC, SZ, SS, and XL revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by funds from the Shanghai Municipal Science and Technology Major Project (Grant No. 2017SHZDZX01 and No.2018SHZDZX01) and ZJLab, the National Key R&D Program of China Project (2017YFA0103900), the National Natural Science Foundation of China (31571083, 31970931), the Program for Professor of Special Appointment (Eastern Scholar of Shanghai, TP2014008), and the Shanghai Rising-Star Program (14QA1400800). This study received funding from State Key Laboratory of Translational Medicine and Innovative Drug Development, Jiangsu Simcere Diagnostics Co., Ltd. and Nanjing Simcere Medical Laboratory Science Co., Ltd. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication. All authors declare no other competing interests.

ACKNOWLEDGMENTS

We thank Jingchun Luo for supervising the data collection at the Human Phenome Institute of Fudan University, Yan Zheng for recruiting volunteers at the Human Phenome Institute of Fudan University, and all the volunteers for their help. We thank Tiemin Liu (Fudan University) for his comments and supervision.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2021.777552/full#supplementary-material

REFERENCES

1. Mohamed Q, Gillies MC, Wong TY. Management of Diabetic Retinopathy: A Systematic Review. *JAMA* (2007) 298:902–16. doi: 10.1001/jama.298.8.902
2. Aiello LP. Angiogenic Pathways in Diabetic Retinopathy. *N Engl J Med* (2005) 353:839–41. doi: 10.1056/NEJMe058142
3. Grover D, Li TJ, Chong CCW. Intravitreal Steroids for Macular Edema in Diabetes. *Cochrane Database Systematic Rev* (2008) CD005656. doi: 10.1002/14651858.CD005656.pub2
4. Cheung N, Mitchell P, Wong TY. Diabetic Retinopathy. *Lancet* (2010) 376:124–36. doi: 10.1016/S0140-6736(09)62124-3
5. Nishikawa T, Edelstein D, Du XL, Yamagishi S, Matsumura T, Kaneda Y, et al. Normalizing Mitochondrial Superoxide Production Blocks Three Pathways of Hyperglycaemic Damage. *Nature* (2000) 404:787–90. doi: 10.1038/35008121
6. Aiello LP, Davis MD, Girach A, Kles KA, Milton RC, Sheetz MJ, et al. Effect of Ruboxistaurin on Visual Loss in Patients With Diabetic Retinopathy. *Ophthalmology* (2006) 113:2221–30. doi: 10.1016/j.ophtha.2006.07.032
7. Abhary S, Hewitt AW, Burdon KP, Craig JE. A Systematic Meta-Analysis of Genetic Association Studies for Diabetic Retinopathy. *Diabetes* (2009) 58:2137–47. doi: 10.2337/db09-0059
8. Bolton WK, Cattran DC, Williams ME, Adler SG, Appel GB, Cartwright K, et al. Randomized Trial of an Inhibitor of Formation of Advanced Glycation End Products in Diabetic Nephropathy. *Am J Nephrol* (2004) 24:32–40. doi: 10.1159/000075627
9. Simó R, Hernández C. Neurodegeneration in the Diabetic Eye: New Insights and Therapeutic Perspectives. *Trends Endocrinol Metab* (2014) 25:23–33. doi: 10.1016/j.tem.2013.09.005
10. Cabrera DeBuc D, Somfai GM. Early Detection of Retinal Thickness Changes in Diabetes Using Optical Coherence Tomography. *Med Sci monitor: Int Med J Exp Clin Res* (2010) 16:MT15–21.
11. Cheung N, Wong TY. Diabetic Retinopathy and Systemic Vascular Complications. *Prog Retin Eye Res* (2008) 27:161–76. doi: 10.1016/j.preteyeres.2007.12.001
12. Ishibazawa A, Nagaoka T, Takahashi A, Omae T, Tani T, Sogawa K, et al. Optical Coherence Tomography Angiography in Diabetic Retinopathy: A Prospective Pilot Study. *Am J Ophthalmol* (2015) 160:35–44.e1. doi: 10.1016/j.ajo.2015.04.021
13. Sun Z, Tang F, Wong R, Lok J, Szeto SKH, Chan JCK, et al. OCT Angiography Metrics Predict Progression of Diabetic Retinopathy and Development of Diabetic Macular Edema: A Prospective Study. *Ophthalmology* (2019) 126:1675–84. doi: 10.1016/j.ophtha.2019.06.016
14. Tang FY, Ng DS, Lam A, Luk F, Wong R, Chan C, et al. Determinants of Quantitative Optical Coherence Tomography Angiography Metrics in Patients With Diabetes. *Sci Rep* (2017) 7:2575. doi: 10.1038/s41598-017-02767-0
15. Durbin MK, An L, Shemonski ND, Soares M, Santos T, Lopes M, et al. Quantification of Retinal Microvascular Density in Optical Coherence Tomographic Angiography Images in Diabetic Retinopathy. *JAMA Ophthalmol* (2017) 135:370–6. doi: 10.1001/jamaophthalmol.2017.0080
16. Reif R, Qin J, An L, Zhi Z, Dziennis S, Wang R. Quantifying Optical Microangiography Images Obtained From a Spectral Domain Optical Coherence Tomography System. *Int J BioMed Imaging* (2012) 2012:509783. doi: 10.1155/2012/509783
17. Chu Z, Lin J, Gao C, Xin C, Zhang Q, Chen CL, et al. Quantitative Assessment of the Retinal Microvasculature Using Optical Coherence Tomography Angiography. *J BioMed Opt* (2016) 21:66008. doi: 10.1117/1.JBO.21.6.066008
18. American Diabetes Association. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2019. *Diabetes Care* (2019) 42:S13–28. doi: 10.2337/dc19-S002
19. Early Treatment Diabetic Retinopathy Study Research Group. Grading Diabetic Retinopathy From Stereoscopic Color Fundus Photographs-An Extension of the Modified Airlie House Classification: ETDRS Report Number 10. *Ophthalmology* (1991) 98:786–806. doi: 10.1016/S0161-6420(13)38012-9
20. Ricci E, Perfetti R. Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification. *IEEE Trans Med Imaging* (2007) 26:1357–65. doi: 10.1109/TMI.2007.898551
21. Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY. A Local Search Approximation Algorithm for k-Means Clustering. *Comput Geom* (2004) 28:89–112. doi: 10.1016/j.comgeo.2004.03.003
22. Li Y, Fu Q, Ye F, Shouno H. Dark Channel Prior Based Blurred Image Restoration Method Using Total Variation and Morphology. *J Sys Eng Electron* (2015) 26:359–366. doi: 10.1109/JSEE.2015.00042
23. Díaz M, Novo J, Cutrín P, Gómez-Ulla F, Penedo MG, Ortega M. Automatic Segmentation of the Foveal Avascular Zone in Ophthalmological OCT-A Images. *PloS One* (2019) 14:e0212364.
24. Xu X, Chen C, Ding W, Yang P, Lu H, Xu F, et al. Automated Quantification of Superficial Retinal Capillaries and Large Vessels for Diabetic Retinopathy on Optical Coherence Tomographic Angiography. *J Biophotonics* (2019) 12:e201900103. doi: 10.1002/jbio.201900103
25. Alam M, Thapa D, Lim JI, Cao D, Yao X. Quantitative Characteristics of Sickle Cell Retinopathy in Optical Coherence Tomography Angiography. *BioMed Opt Express* (2017) 8:1741–53. doi: 10.1364/BOE.8.001741
26. Krawitz BD, Mo S, Geyman LS, Agemy SA, Scripsema NK, Garcia PM, et al. Acircularity Index and Axis Ratio of the Foveal Avascular Zone in Diabetic Eyes and Healthy Controls Measured by Optical Coherence Tomography Angiography. *Vision Res* (2017) 139:177–86. doi: 10.1016/j.visres.2016.09.019

27. Freiberg FJ, Pfau M, Wons J, Wirth MA, Becker MD, Michels S. Optical Coherence Tomography Angiography of the Foveal Avascular Zone in Diabetic Retinopathy. *Graefes Arch Clin Exp Ophthalmol* (2016) 254:1051–8. doi: 10.1007/s00417-015-3148-2
28. Fang D, Tang FY, Huang H, Cheung CY, Chen H. Repeatability, Interocular Correlation and Agreement of Quantitative Swept-Source Optical Coherence Tomography Angiography Macular Metrics in Healthy Subjects. *Br J Ophthalmol* (2019) 103:415–20. doi: 10.1136/bjophthalmol-2018-311874
29. Bartko JJ. The Intraclass Correlation Coefficient as a Measure of Reliability. *psychol Rep* (1966) 19:3–11. doi: 10.2466/pr0.1966.19.1.3
30. Bland JM, Altman DG. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *Int J Nurs Stud* (2010) 47:931–6. doi: 10.1016/j.ijnurstu.2009.10.001
31. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc (Methodological)* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
32. Fawcett T. An Introduction to ROC Analysis. *Pattern Recognition Lett* (2006) 27:861–74. doi: 10.1016/j.patrec.2005.10.010
33. Powers DMW. *Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation* (2020). Available at: http://arxiv.org/abs/2010.16061.
34. Eladawi N, Elmogy M, Helmy O, Aboelfetouh A, Riad A, Sandhu H, et al. Automatic Blood Vessels Segmentation Based on Different Retinal Maps From OCTA Scans. *Comput Biol Med* (2017) 89:150–61. doi: 10.1016/j.compbiomed.2017.08.008
35. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology* (1945) 26:297–302. doi: 10.2307/1932409
36. Powers DMW. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *J Mach Learn Technol* (2011) 2:37–63.
37. Pott S, Lieb JD. What are Super-Enhancers? *Nat Genet* (2015) 47:8–12. doi: 10.1038/ng.3167

Conflict of Interest: Author BingX was employed by company Jiangsu Simcere Diagnostics Co., Ltd., and company Nanjing Simcere Medical Laboratory Science Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 03 January 2022
doi: 10.3389/fendo.2021.781890

Association Between Serum Furin and Fasting Glucose: A Cross-Sectional Study in Chinese Adults

Yan He[1†], Hanyun Zhu[2†], Min Zhang[3†], Jing Li[1], Shengqi Ma[2], Yin Lu[1], Linan Chen[1], Mingzhi Zhang[2] and Hao Peng[1,4*]*

[1] Department of Epidemiology, School of Public Health, Medical College of Soochow University, Suzhou, China, [2] Department of Biostatistics, School of Public Health, Medical College of Soochow University, Suzhou, China, [3] Central Office, Suzhou National New and Hi-Tech Industrial Development Zone Center for Disease Control and Prevention, Suzhou, China, [4] Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Suzhou, China

OPEN ACCESS

Edited by:
Tao Huang,
Peking University, China

Reviewed by:
Eusebio Chiefari,
University Magna Graecia of Catanzaro, Italy
Marco Arieli Herrera-Valdez,
Universidad Nacional Autónoma de México, Mexico

***Correspondence:**
Hao Peng
penghao@suda.edu.cn
Mingzhi Zhang
zhangmingzhi@suda.edu.cn

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: 23 September 2021
Accepted: 08 December 2021
Published: 03 January 2022

Citation:
He Y, Zhu H, Zhang M, Li J, Ma S, Lu Y, Chen L, Zhang M and Peng H (2022) Association Between Serum Furin and Fasting Glucose: A Cross-Sectional Study in Chinese Adults. Front. Endocrinol. 12:781890. doi: 10.3389/fendo.2021.781890

Background: Furin has been associated with glucose metabolic phenotypes in small sampled clinical studies. However, this association has not yet been studied in Chinese. Here, we aimed to examine the association between serum furin and fasting glucose in Chinese adults.

Methods: Serum furin and fasting plasma glucose were assayed for 2,172 participants (mean aged 53 years, 38% men) in the Gusu cohort. A median regression model was applied to examine the association between serum furin and fasting glucose, adjusting for age, sex, education level, cigarette smoking, alcohol drinking, obesity, blood pressure, and lipids. To facilitate data interpretation, the association between serum furin and prevalent diabetes was also examined.

Results: Serum furin was negatively associated with fasting glucose (β=-0.18, P<0.001 for log-furin). In participants with diabetes, serum furin was significantly lower than those with normal glucose (median: 0.90 ng/mL *vs.* 1.05 ng/mL, P=0.001). Compared with participants in the highest quartile of serum furin, those in the lowest quartile had 42% and 80% increased risk of prevalent prediabetes (OR=1.42, 95%CI: 1.05-1.92, P=0.023) and diabetes (OR=1.80, 95%CI: 1.13-2.91, P=0.015), respectively.

Conclusions: Serum furin was negatively associated with prediabetes and diabetes in Chinese adults. Our findings suggest that serum furin may be a risk factor or a biomarker of diabetes.

Keywords: furin, diabetes, fasting plasma glucose, Chinese, cross-sectional study, risk factor

INTRODUCTION

Furin, ubiquitously expressed in all mammalian tissues and cells, is a member of the proprotein convertase subtilisin/Kexin (PCSK) family (1). Accumulating evidence has demonstrated that furin may be implicated in the process of glucose metabolism by converting numerous protein and peptide precursors into their bioactive forms. For example, cell-based studies found that furin could activate the precursors of insulin

and its receptor (2, 3) and regulate the proliferation and differentiation of pancreatic β-cells which determines the secretion of insulin (4, 5). Animal experiments found that knockout of the *furin* gene in β-cells resulted in glucose intolerance in mice (6). In humans, polymorphisms in the *FURIN* gene have been associated with metabolic syndrome (7), hypertension (8), and coronary artery disease (9). These findings suggest a potential role of furin in glucose metabolism. In fact, the levels of furin in the circulation have been associated with diabetes (10) and some relative phenotypes, such as obesity (11), metabolic syndrome (12), and diabetic cardiovascular disease (13). However, these results were largely derived from populations with European ancestry. It's not clear whether circulating furin could be associated with glucose metabolism in Chinese who have a different genetic background and risk profiles of diabetes in comparison to European populations. In Chinese adults, we previously found that serum furin was associated with obesity (14) and hypertension (15). Whether serum furin is associated with diabetes has not been studied in Chinese. Therefore, we aimed to examine the association between serum furin and fasting plasma glucose in Chinese adults.

METHODS

Participants

The Gusu cohort is a prospective study aiming to identify new risk factors for cardiovascular disease (CVD) in Chinese adults. The study design, survey methods, and laboratory measurements have been detailed previously (16). In brief, 2,498 community members aged over 30 years and free of CVD and chronic kidney disease were enrolled in the baseline examination in 2010. After excluding participants with missing data on serum furin (n=326), 2172 participants were finally included in the current analysis. The protocols of the current study were approved by the Soochow University Ethics Committee (Approval No. SUDA20200601H02). Written informed consent was obtained from all study participants.

Measurement of Serum Furin

Using the -80°C stored serums obtained in the baseline examination, furin concentrations were measured using commercial ELISA kits (Catalog: DL-FUR-Hu; DLDEVELOP, Wuxi, China) according to standard protocols as previously described (15). All the samples were processed in a duplicate assay. A standard curve was constructed and from which furin concentrations of unknown samples were determined.

Measurement of Fasting Plasma Glucose

Blood samples were obtained by venipuncture in the morning after a requested overnight fast (at least 8 h). Within 4 hours after venipuncture, fasting plasma glucose (FPG) was analyzed enzymatically on a Hitachi 7020 automatic biochemical analyzer using commercial reagents (Kangxiang Medical Appliances, Shanghai, PR of China). Intra- and inter-assay coefficients of variation were less than 2% and 4%, respectively. Diabetes was defined as the presence of one of the following: (a) a self-reported previous diagnosis by health care professionals and current use of either insulin or oral hypoglycemic medication, and (b) an FPG level of 7.0 mmol/L or higher (17). Prediabetes was defined as an FPG level between 5.6 and 6.9 mmol/L without self-reported physician-diagnosed diabetes.

Assessment of Conventional Risk Factors

Data on demographic information, lifestyle risk factors, and personal medical history were collected with standard questionnaires in the Chinese language administered by trained staff. Current smoking was defined as having smoked at least 100 cigarettes in the entire lifetime, smoke cigarettes regularly, and smoke currently. Current drinking was defined as having consumed alcohol ≥ 12 times in the past year and drinking currently. Education level was estimated as years a participant stays in the education system. Body mass index (BMI) was calculated as weight in kilograms divided by the square of the subject's height in squared meters. Three consecutive sitting blood pressure measurements (3 min between each) were taken by trained staff, using a standard mercury sphygmomanometer, according to standard protocol (18), after the subjects had been resting for at least 5 min. The first and fifth Korotkoff sounds were recorded as systolic blood pressure (SBP) and diastolic blood pressure (DBP), respectively. The mean of the three records was used in the analysis. Total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) were examined for all participants.

Statistical Analysis

Baseline characteristics of study participants were presented according to quartiles of serum furin. Base-10 logarithmic transformation was applied to maximal normal distribution of serum furin and the generated values (log-furin) were used in downstream analyses. To examine the association between serum furin and FPG, we constructed a median regression model in which FPG was the dependent variable and serum furin (continuous log-furin or categorical furin in quartiles) was the independent variable, adjusting for age, sex, education level, cigarette smoking, alcohol drinking, BMI, SBP, LDL-C, and HDL-C. Median regression was used here to account for the skewed distribution data of FPG. To facilitate data interpretation, we further examined the association between serum furin and prevalent diabetes by constructing a logistic regression model in which prevalent diabetes (yes/no) was the dependent variable and serum furin (log-furin or furin quartiles) was the independent variable, adjusting for the confounding factors listed above. The association between serum furin and prediabetes was similarly examined. All statistical analyses were conducted using SAS statistical software (version 9.1, Cary, NC). A two-tailed *P* value less than 0.05 was considered statistically significant.

RESULTS

Baseline Characteristics

There were 2,172 participants (823 men and 1349 women) with a mean age of 53.2 years included in the present study. Of them,

198 (9.12%) participants including 85 patients under hypoglycemic treatments were diagnosed with prevalent diabetes. Their baseline characteristics according to quartiles of serum furin are shown in **Table 1**. Compared with participants with a higher level of serum furin, those with a lower level of serum furin were more likely to be older, current smokers, current drinkers and had higher levels of BMI, SBP, DBP, and FPG (all $P<0.05$).

Levels of Serum Furin in Participants With Different Statuses of Glucose Metabolism

Intra- and inter-assay coefficients of variation of the measurements of serum furin were less than 10% and 12%, respectively. Compared to participants with a normal FPG, the median levels of serum furin was significantly lower in those with prediabetes (median: 0.96 ng/mL *vs.* 1.05 ng/mL, $P=0.003$) and diabetes (median: 0.90 ng/mL *vs.* 1.05 ng/mL, $P=0.001$) (**Figure 1**). We did not find any significant difference in serum furin levels between participants with prediabetes and diabetes ($P=0.353$).

Association Between Serum Furin and Glucose Metabolic Status

After adjusting for age, sex, education level, cigarette smoking, alcohol drinking, BMI, SBP, LDL-C, and HDL-C, serum furin was negatively associated with FPG (β=-0.18, $P<0.001$ for log-furin, **Table 2**). Compared with participants at the highest quartile of serum furin, those at the lowest quartile had a median of 0.19 mmol/L higher FPG (P<0.001). The regression with prevalent prediabetes and diabetes revealed that a higher level of serum furin seemed to be nominally associated with a lower risk of having prediabetes (OR=0.72, $P=0.051$) and diabetes (OR=0.62, $P=0.052$) (**Table 3**). Compared with participants at the highest quartile of serum furin, those at the lowest quartile had 42% and 80% increased risk of having prediabetes ($P=0.023$) and diabetes ($P=0.015$), respectively.

DISCUSSION

We are the first to report an association between serum furin and fasting glucose in Chinese adults in a cross-sectional study including a relatively large sample of participants in the Gusu cohort. Participants with a lower level of serum furin were more likely to have prevalent prediabetes and diabetes. This association was independent of behavioral and metabolic factors. Our findings indicated that furin may participate in glucose metabolism through mechanisms beyond metabolic factors. Furin deficiency may be a marker or even a potential risk factor for diabetes.

The identified association between serum furin and glucose metabolism found in our study has also been suggested by other studies. For example, β cell-specific *furin* gene knockout mice developed glucose intolerance and had smaller islets with lower insulin content than wild controls (6). In humans, polymorphisms of the *FURIN* gene encoding furin protein have been associated with some diabetic complications, such as metabolic syndrome (7), hypertension (8), and coronary artery disease (9). Further, furin protein in circulation has also been studied by some clinical studies. For example, a case-control study including 25 diabetic patients with complications, 25 diabetic patients without complications, and 25 healthy controls found that serum furin was significantly associated with diabetes complicated with cardiovascular disease (13). A cross-sectional study including 138 participants found that furin was significantly associated with metabolic syndrome (12). A prospective study including 4,678 participants reported an association of serum furin with incident diabetes among Swedish in the Malmö Diet and Cancer Study (10). However, these results were mainly generated from White populations whose genetic background and diabetes risk profiles differ from Chinese. Our study is the first to examine the association between serum furin and diabetes in Chinese adults. Nevertheless, we found an inconsistent result that a lower level

TABLE 1 | Baseline characteristics of study participants according to serum furin levels (N = 2172).

Characteristics	Serum furin, ng/mL				*P*-value for trend
	Quartile 1 (~0.59)	Quartile 2 (0.60~1.01)	Quartile 3 (1.02~1.75)	Quartile 4 (1.76~)	
No. of participants	544	543	543	542	
Age, years, mean ± SD	54.37 ± 9.51	53.49 ± 9.51	51.90 ± 9.52	51.58 ± 9.57	<0.001
Sex, men (%)	254 (46.69)	208 (38.31)	180 (33.15)	181 (33.39)	<0.001
Current smoking, n (%)	146 (26.84)	122 (22.47)	96 (17.68)	121 (22.32)	0.022
Current drinking, n (%)	128 (23.53)	103 (18.97)	88 (16.21)	76 (14.02)	<0.001
Education level, years, mean ± SD	7.18 ± 3.31	7.32 ± 3.14	7.04 ± 3.26	6.83 ± 3.21	0.033
BMI, kg/m^2, mean ± SD	24.91 ± 3.70	25.10 ± 3.46	24.78 ± 3.46	24.51 ± 3.83	0.031
SBP, mmHg, mean ± SD	132.9 ± 15.5	131.1 ± 14.3	129.1 ± 15.8	128.3 ± 20.1	<0.001
DBP, mmHg, mean ± SD	86.3 ± 7.6	85.6 ± 8.1	84.4 ± 8.9	84.0 ± 11.3	<0.001
Fasting glucose, mmol/L	5.3 (4.8-5.8)	5.2 (4.7-5.7)	5.1 (4.7-5.6)	5.0 (4.6-5.6)	0.002
Total cholesterol, mmol/L	5.05 (4.50-5.63)	5.10 (4.56-5.74)	5.09 (4.53-5.77)	5.11 (4.55-5.68)	0.154
Triglycerides, mmol/L	1.09 (0.79-1.50)	1.19 (0.82-1.70)	1.08 (0.77-1.68)	1.11 (0.75-1.65)	0.428
LDL cholesterol, mmol/L	2.96 (2.56-3.50)	2.97 (2.51-3.46)	2.93 (2.50-3.37)	2.92 (2.48-3.39)	0.164
HDL cholesterol, mmol/L	1.44 (1.24-1.71)	1.43 (1.19-1.70)	1.50 (1.26-1.75)	1.45 (1.22-1.72)	0.612

Results are expressed as median (interquartile range) unless otherwise noted. P-values for trend were tested by linear regression model for continuous variables and chi-square trend test for categorical variables.

BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; LDL, low-density lipoprotein; HDL, high-density lipoprotein.

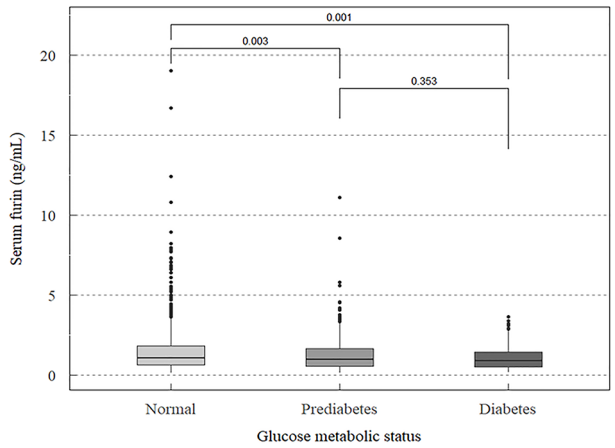


FIGURE 1 | Median levels of serum furin in participants with different statuses of glucose metabolism.

of serum furin was associated with a higher risk of prevalent diabetes. Different genetic backgrounds of study participants may explain the discrepancy. Notably, basic studies demonstrated that furin deficiency caused by gene knockout could result in glucose intolerance (6). This finding may indicate a feedback adjustment mechanism that increased glucose in diabetic patients stimulates secretion of furin in feedback to facilitate activation of insulin and its receptors. Therefore, researchers may observe a positive association between serum furin and diabetes in populations as listed above. These results appeared to indicate that the relation between serum furin and diabetes in populations is more complex than that in cell and animal experiments. As a wide proprotein convertase, furin acts as an upstream regulator of the glucose metabolism system, such as the natriuretic peptides system. This system plays a critical role in maintaining blood glucose balance and blood pressure through BNP which was activated by furin. As a result, furin expression and excretion may be upregulated in compensatory for high-risk individuals of diabetes, e.g., hypertensive and obesity patients. Also, our group previously found that serum furin was also significantly associated with obesity and hypertension that share many risk factors and molecular mechanisms with diabetes (14, 15). All these findings, even conflicting, together with ours, suggest that furin may participate in glucose metabolism and therefore could be a therapeutic target for diabetes.

In addition to population evidence, the possible mechanisms underlying the association between serum furin and diabetes could deepen our understanding of the role of furin in diabetes. Furin, which belongs to the PCSK family and converts numerous proteins and peptide precursors into their biologically active forms, has been demonstrated to play an important role in glucose metabolism and several relative processes associated with diabetes, such as insulin resistance. One possible mechanism could be that furin deficiency influences the maturation of insulin receptors (2, 19) which is critical for the maintenance of glucose homeostasis. The second possible mechanism may be the involvement of furin in the proliferation and differentiation of pancreatic β-cells (5) where glucose metabolism mainly occurs. Another possible mechanism might be related to the role of furin in the activation of B-type natriuretic peptides (20, 21) which is one key component of the natriuretic peptides system. This system has been demonstrated to play an integral role in glucose metabolism and participate in the development of diabetes (20, 22, 23).

TABLE 2 | Association between serum furin and fasting plasma glucose.

serum furin(ng/mL)	Un-adjusted		Adjusted*	
	β (SE)	*P*-value	β (SE)	*P*-value
log-furin	-0.27 (0.06)	<0.001	-0.18 (0.05)	<0.001
Categorical				
Quartile 4	ref		ref	
Quarntile 3	0.10 (0.05)	0.053	0.08 (0.04)	0.062
Quarntile 2	0.20 (0.05)	<0.001	0.04 (0.05)	0.393
Quarntile 1	0.30 (0.05)	<0.001	0.19 (0.04)	<0.001

**Adjusting for age, sex, education level, cigarette smoking, alcohol consumption, SBP, BMI, LDL, and HDL.*

TABLE 3 | Associations between serum furin and prevalent prediabetes and diabetes.

serum furin(ng/mL)	Prediabetes			Diabetes		
	Cases	OR (95%CI)*	*P*-value	Cases	OR (95%CI)*	*P*-value
log-furin	474	0.72 (0.52-1.00)	0.051	198	0.62 (0.38-1.00)	0.052
Categorical						
Quartile 4	108	ref		33	ref	
Quartile 3	107	1.05 (0.76-1.43)	0.779	55	1.74 (1.08-2.82)	0.024
Quartile 2	114	0.98 (0.72-1.34)	0.910	48	1.28 (0.79-2.10)	0.315
Quartile 1	145	1.42 (1.05-1.92)	0.023	62	1.80 (1.13-2.91)	0.015

**Adjusting for age, sex, education level, cigarette smoking, alcohol consumption, SBP, BMI, LDL, and HDL.*

To the best of our knowledge, our study is the first to investigate the association between serum furin and diabetes in Chinese adults. The strengths of this study include careful and systemic analyses of the association between serum furin and glucose metabolic states including prediabetes and diabetes and comprehensive adjustments of many conventional risk factors including behavioral and metabolic factors. However, our study also has several limitations that deserve clarification. First, the cross-sectional study designs prevented a causal inference. However, our results showed that serum furin was not only associated with diabetes but also correlated with prediabetes. This finding increased the probability that furin deficiency may be a risk factor for diabetes. Nevertheless, the causality between serum furin and diabetes is not established and needs further evidence from clinical trials. Second, our study population was comprised of ethnic Han individuals only, whether the results can be generalizable to other racial/ethnic groups is not clear.

CONCLUSION

In conclusion, our study demonstrated that serum furin was negatively associated with prediabetes and diabetes in Chinese adults. These results suggested that furin might play a potential role in glucose metabolism and the deficiency of furin might serve as a risk factor or biomarker for diabetes. However, the causal association between furin and diabetes still needs more evidence and the underlying molecular mechanisms still need further investigations.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Soochow University Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YH, MZZ, and HP conceived and designed the study. HZ and MZ analyzed and interpreted the data. YH and HZ drafted the paper. MZ, YH, JL, YL, LC, and SM collected the data. MZZ and HP revised and gave the final approval of the version to be published, and all authors agreed to be accountable for all aspects of the work.

FUNDING

This study was supported by the National Natural Science Foundation of China (NO. 82173596, 81903384, and 81872690), the Natural Science Foundation of Jiangsu Province (NO. BK20180841), the Suzhou Municipal Science and Technology Bureau (NO. SS201853, SKJY2021040, and SYS2020091), and a Project of the Priority Academic Program Development of Jiangsu Higher Education Institutions.

ACKNOWLEDGMENTS

We are deeply appreciative of the cooperation and participation of the members in the Gusu cohort and thank all staff for their support and assistance. Especially, we thank the Gusu Center for Disease Prevention and Control for their support in the recruitment of participants.

REFERENCES

1. Seidah NG, Day R, Marcinkiewicz M, Chrétien M. Precursor Convertases: An Evolutionary Ancient, Cell-Specific, Combinatorial Mechanism Yielding Diverse Bioactive Peptides and Proteins. *Ann N Y Acad Sci* (1998) 839:9–24. doi: 10.1111/j.1749-6632.1998.tb10727.x
2. Bravo DA, Gleason JB, Sanchez RI, Roth RA, Fuller RS. Accurate and Efficient Cleavage of the Human Insulin Proreceptor by the Human Proprotein-Processing Protease Furin. Characterization and Kinetic Parameters Using the Purified, Secreted Soluble Protease Expressed by a Recombinant Baculovirus. *J Biol Chem* (1994) 269:25830–7. doi: 10.1016/S0021-9258(18)47322-X
3. Vollenweider F, Kaufmann J, Irminger JC, Halban PA. Processing of Proinsulin by Furin, PC2, and PC3 in (Co) Transfected COS (Monkey Kidney) Cells. *Diabetes* (1995) 44:1075–80. doi: 10.2337/diab.44.9.1075
4. Louagie E, Taylor NA, Flamez D, Roebroek AJ, Bright NA, Meulemans S, et al. Role of Furin in Granular Acidification in the Endocrine Pancreas: Identification of the V-ATPase Subunit Ac45 as a Candidate Substrate. *Proc Natl Acad Sci USA* (2008) 105:12319–24. doi: 10.1073/pnas.0800340105
5. Kayo T, Sawada Y, Suda M, Konda Y, Izumi T, Tanaka S, et al. Proprotein-Processing Endoprotease Furin Controls Growth of Pancreatic Beta-Cells. *Diabetes* (1997) 46:1296–304. doi: 10.2337/diabetes.46.8.1296
6. Brouwers B, Coppola I, Vints K, Dislich B, Jouvet N, Van Lommel L, et al. Loss of Furin in β-Cells Induces an Mtorc1-ATF4 Anabolic Pathway That Leads to β-Cell Dysfunction. *Diabetes* (2021) 70:492–503. doi: 10.2337/db20-0474
7. Ueyama C, Horibe H, Yamase Y, Fujimaki T, Oguri M, Kato K, et al. Association of FURIN and ZPR1 Polymorphisms With Metabolic Syndrome. *Biomed Rep* (2015) 3:641–7. doi: 10.3892/br.2015.484
8. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic Variants in Novel Pathways Influence Blood Pressure and Cardiovascular Disease Risk. *Nature* (2011) 478:103–9. doi: 10.1038/nature10405
9. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al. Large-Scale Association Analysis Identifies New Risk Loci for Coronary Artery Disease. *Nat Genet* (2013) 45:25–33. doi: 10.1038/ng.2480
10. Fernandez C, Rysä J, Almgren P, Nilsson J, Engström G, Orho-Melander M, et al. Plasma Levels of the Proprotein Convertase Furin and Incidence of Diabetes and Mortality. *J Internal Med* (2018) 284:377–87. doi: 10.1111/joim.12783
11. Swärd P, Rosengren BE, Jehpsson L, Karlsson MK. Association Between Circulating Furin Levels, Obesity and Pro-Inflammatory Markers in Children. *Acta Paediatr (Oslo Norway 1992)* (2021) 110:1863–8. doi: 10.1111/apa.15774

12. Harlid S, Myte R, Van Guelpen B. The Metabolic Syndrome, Inflammation, and Colorectal Cancer Risk: An Evaluation of Large Panels of Plasma Protein Markers Using Repeated, Prediagnostic Samples. *Mediators Inflamm* (2017) 2017:4803156. doi: 10.1155/2017/4803156
13. Fathy SA, Abdel Hamid FF, Zabut BM, Jamee AF, Ali MA, Abu Mustafa AM. Diagnostic Utility of BNP, Corin and Furin as Biomarkers for Cardiovascular Complications in Type 2 Diabetes Mellitus Patients. *Biomarkers* (2015) 20:460–9. doi: 10.3109/1354750X.2015.1093032
14. He Y, Ren L, Zhang Q, Zhang M, Shi J, Hu W, et al. Deficient Serum Furin Predicts Risk of Abdominal Obesity: Findings From a Prospective Cohort of Chinese Adults. *Postgrad Med J* (2021) 97:234–8. doi: 10.1136/postgradmedj-2019-137422
15. He Y, Ren L, Zhang Q, Zhang M, Shi J, Hu W, et al. Serum Furin as a Biomarker of High Blood Pressure: Findings From a Longitudinal Study in Chinese Adults. *Hypertens Res* (2019) 42:1808–15. doi: 10.1038/s41440-019-0295-6
16. Peng H, Zhang Q, Cai X, Liu Y, Ding J, Tian H, et al. Association Between High Serum Soluble Corin and Hypertension: A Cross-Sectional Study in a General Population of China. *Am J Hypertens* (2015) 28:1141–9. doi: 10.1093/ajh/hpv002
17. American Dibetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. *Diabetes Care* (2020) 43:S14–s31. doi: 10.2337/dc20-S002
18. Weber MA, Schiffrin EL, White WB, Mann S, Lindholm LH, Kenerson JG, et al. Clinical Practice Guidelines for the Management of Hypertension in the Community: A Statement by the American Society of Hypertension and the International Society of Hypertension. *J Clin Hypertens (Greenwich Conn)* (2014) 16:14–26. doi: 10.1111/jch.12237
19. Bernot D, Stalin J, Stocker P, Bonardo B, Scroyen I, Alessi MC, et al. Plasminogen Activator Inhibitor 1 is an Intracellular Inhibitor of Furin Proprotein Convertase. *J Cell Sci* (2011) 124:1224–30. doi: 10.1242/jcs.079889
20. Verboven K, Hansen D, Jocken JWE, Blaak EE. Natriuretic Peptides in the Control of Lipid Metabolism and Insulin Sensitivity. *Obes Rev* (2017) 18:1243–59. doi: 10.1111/obr.12598
21. Semenov AG, Tamm NN, Seferian KR, Postnikov AB, Karpova NS, Serebryanaya DV, et al. Processing of Pro-B-Type Natriuretic Peptide: Furin and Corin as Candidate Convertases. *Clin Chem* (2010) 56:1166–76. doi: 10.1373/clinchem.2010.143883
22. Pfister R, Sharp S, Luben R, Welsh P, Barroso I, Salomaa V, et al. Mendelian Randomization Study of B-Type Natriuretic Peptide and Type 2 Diabetes: Evidence of Causal Association From Population Studies. *PloS Med* (2011) 8: e1001112. doi: 10.1371/journal.pmed.1001112
23. Heinisch BB, Vila G, Resl M, Riedl M, Dieplinger B, Mueller T, et al. B-Type Natriuretic Peptide (BNP) Affects the Initial Response to Intravenous Glucose: A Randomized Placebo-Controlled Cross-Over Study in Healthy Men. *Diabetologia* (2012) 55:1400–5. doi: 10.1007/s00125-011-2392-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 25 January 2022
doi: 10.3389/fendo.2021.724278

Identification of Shared Genes and Pathways in Periodontitis and Type 2 Diabetes by Bioinformatics Analysis

Junho Kang[1†], Eun Jung Kwon[2†], Mihyang Ha[2], Hansong Lee[2], Yeuni Yu[2], Ji Wan Kang[2], Yeongjoo Kim[2], Eun Young Lee[3], Ji-Young Joo[4], Hye Jin Heo[5], Eun Kyoung Kim[5], Tae Woo Kim[6], Yun Hak Kim[5,7] and Hae Ryoun Park[3*]*

[1] Medical Research Institute, Pusan National University, Busan, South Korea, [2] Interdisciplinary Program of Genomic Data Science, Pusan National University, Busan, South Korea, [3] Department of Oral Pathology, School of Dentistry, Pusan National University, Yangsan, South Korea, [4] Department of Periodontology, School of Dentistry, Pusan National University, Yangsan, South Korea, [5] Departmment of Anatomy, School of Medicine, Pusan National University, Yangsan, South Korea, [6] Department of Orthopaedic Surgery, Pusan National University Yangsan Hospital, School of Medicine, Pusan National University, Yangsan, South Korea, [7] Department of Biomedical Informatics, School of Medicine, Pusan National University, Yangsan, South Korea

OPEN ACCESS

Edited by:
Jie Li,
Guangdong Provincial People's Hospital, China

Reviewed by:
Jinhui Liu,
Nanjing Medical University, China
Simin Li,
Southern Medical University, China
Wen-Li Lu,
China Medical University Children's Hospital, Taiwan

***Correspondence:**
Yun Hak Kim
yunhak10510@pusan.ac.kr
Hae Ryoun Park
parkhr@pusan.ac.kr

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *20 July 2021*
Accepted: *20 December 2021*
Published: *25 January 2022*

Citation:
Kang J, Kwon EJ, Ha M, Lee H, Yu Y, Kang JW, Kim Y, Lee EY, Joo J-Y, Heo HJ, Kim EK, Kim TW, Kim YH and Park HR (2022) Identification of Shared Genes and Pathways in Periodontitis and Type 2 Diabetes by Bioinformatics Analysis. Front. Endocrinol. 12:724278. doi: 10.3389/fendo.2021.724278

Introduction: It is well known that the presence of diabetes significantly affects the progression of periodontitis and that periodontitis has negative effects on diabetes and diabetes-related complications. Although this two-way relationship between type 2 diabetes and periodontitis could be understood through experimental and clinical studies, information on common genetic factors would be more useful for the understanding of both diseases and the development of treatment strategies.

Materials and Methods: Gene expression data for periodontitis and type 2 diabetes were obtained from the Gene Expression Omnibus database. After preprocessing of data to reduce heterogeneity, differentially expressed genes (DEGs) between disease and normal tissue were identified using a linear regression model package. Gene ontology and Kyoto encyclopedia of genes and genome pathway enrichment analyses were conducted using R package '*vsn*'. A protein-protein interaction network was constructed using the search tool for the retrieval of the interacting genes database. We used molecular complex detection for optimal module selection. CytoHubba was used to identify the highest linkage hub gene in the network.

Results: We identified 152 commonly DEGs, including 125 upregulated and 27 downregulated genes. Through common DEGs, we constructed a protein-protein interaction and identified highly connected hub genes. The hub genes were up-regulated in both diseases and were most significantly enriched in the Fc gamma R-mediated phagocytosis pathway.

Discussion: We have identified three up-regulated genes involved in Fc gamma receptor-mediated phagocytosis, and these genes could be potential therapeutic targets in patients with periodontitis and type 2 diabetes.

Keywords: differentially expressed genes, hub genes, gene expression omnibus, biomarker, Fc gamma receptor

INTRODUCTION

Periodontitis (PD) and diabetes are complex chronic diseases, and the link between the two diseases has been continuously reported (1, 2). Previous epidemiologic studies have reported that the risk of periodontitis increases by about three times in diabetics compared to non-diabetic patients. In addition, adults with HbA1c levels above 9% had a higher prevalence of severe periodontitis than adults without diabetes (3). These epidemiological studies indicate that depending on the state of diabetes, it is associated with an increase in the prevalence and severity of periodontitis. The study that suggested the possibility of diabetes incidence according to periodontal disease status was a study by the Gila River Indian Community, a Native American population with a high prevalence of diabetes (4). In this study, patients with untreated severe periodontitis were associated with poor glycemic control (HbA1c >9.0%, 75 mmol/mol), suggesting that diabetes control may be inhibited. Thus, the identification of factors about common pathogenic changes of them might provide an important clue for the treatment.

PD is one of the most representative chronic inflammatory diseases, which is caused by periodontopathogenic bacteria harbored in a gingival crevice, whereas type 2 diabetes (T2DM) is a metabolic disease with complicated pathogenic factors. Thus, the diseases are independent of each other with respect to disease classification, but several studies have shown that the two diseases share risk factors such as socioeconomic status, smoking, and age (5). Although understanding environmental causes, such as demographic factors may play an important role in the prevention and resolution of diseases, clarification of specific pathogenic determinants beyond simple demographic factors is needed to develop a strategy for curing the two diseases effectively. Studies have suggested the importance of inflammation as a linking factor, based on the clinical findings of a decrease in HbA1c levels after surgical and/or non-surgical periodontal therapy and the resultant alleviation of PD (6). The problem is that inflammation is not a particular etiological agent but a generic term for a pathological process comprising complicated processes. Although numerous cytokines that are involved in the inflammatory process could be suggested as major factors in both diseases, these may be simply modifying factors that could modulate the pathogenesis of the diseases. Thus, clinical and/or experimental studies have limitations in clearly revealing the common pathogenic mechanisms.

Microarray technology provides systematic biological solutions from hardware to software systems. Simultaneously scans the hybridization signal of tens of thousands of gene probes on the chip and enables quantitative analysis of the transcript profile of the sample (7–9). Microarray analysis can identify biomarkers for disease and provide insight into novel therapeutic targets. The wide application of microarray analysis has generated large amounts of data, most of which are already stored in public databases such as GEO. Integrating and analyzing this data can lead to novel and important clues about the disease (10). In recent years, several independent studies have performed microarray analyzes of periodontitis and diabetes (11, 12). In independent studies, due to the heterogeneity of tissues or samples, most microarray results were limited or inconsistent, and were generated in a single cohort study. To minimize these issues, we tried to identify potential biomarkers of periodontitis and diabetes by integrating expression data from several independent cohorts through bioinformatics methods.

MATERIALS AND METHODS

Data Acquisition

Expression profiling data by array of the PD and T2DM were obtained from the Gene Expression Omnibus (GEO) database (13). The accession numbers were GSE10334 (12), GSE16134 (14), GSE23586 (15), GSE20966 (11), and GSE25724 (16). The keywords used to select the datasets are shown in **Table 1**. All included datasets were related to gene expression data from disease (PD, T2DM) groups and/or normal groups and were downloaded for differentially expressed genes (DEG) analysis. The flow chart of this study is shown in **Figure 1** and the characteristics of the included datasets are listed in **Table 2**.

Data Preprocessing

There is inevitable heterogeneity between studies published in GEO. To minimize heterogeneity, we pre-processed raw data using the same method and only included studies conducted on the same platform. Briefly, data preprocessing (background correction and normalization) for raw probe intensities was performed using the "*vsn*" (variance stabilizing normalization, v 3.24.0) R/Bioconductor package (17). Based on the annotation file downloaded from GEO, the probe id was converted to a

TABLE 1 | Electric search strategy in GEO.

Database	GEO	
Disease	**Periodontitis**	**Type 2 diabetes mellitus**
Search Term	[("Periodontitis" (MeSH Terms) AND "Homosapiens"(porgn: txid9606)]	[("diabetes mellitus, type 2"(MeSH Terms) OR type 2 diabetes (All Fields)] AND ["pancreas" (MeSH Terms) OR pancreatic (All Fields)] AND ["tissues"(MeSH Terms) OR tissue (All Fields)] AND "Homo sapiens"(porgn) AND ["gds"(Filter) AND "Expression profiling by array"(Filter)].
Filter	Expression profiling by array Human	Expression profiling by array Human
Results	19	4

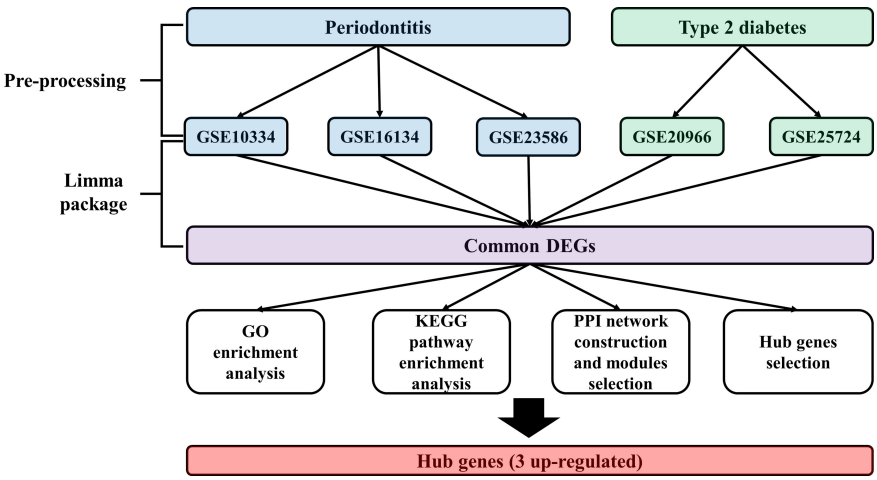


FIGURE 1 | Study flow chart.

genetic symbol. When multiple probes corresponded to a single gene, the average level of the probes was calculated as the expression value of the gene. In addition, probes that did not correspond to the genetic symbols were removed. To screen the intersectional genes significantly expressed in each cohort, R package "*venn*" was used to plot Venn diagrams.

Differential Expression Analysis

To identify DEGs between disease and normal tissue, we performed the linear regression model package, "limma", in R language (v 3.42.2) (18). We considered genes with adjusted p-value < 0.05 to be significantly differentially expressed. The method used for adjusted p-value is false discovery rate. The $\log_2$ fold change > 0.5 was set to upregulated genes, and $\log_2$ fold change < 0.5 was set to downregulated genes.

Functional and Pathway Enrichment Analysis

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genome (KEGG) pathway enrichment analyses were performed using R package "clusterProfiler" (v 3.14.3) (19). The common DEGs identified in each disease were included in GO term enrichment analysis and KEGG pathway enrichment analysis, which were performed separately for up and downregulated

TABLE 2 | Characteristics of included datasets.

	GSE10334		GSE16134		GSE23586	
	Normal	**Periodontitis**	**Normal**	**Periodontitis**	**Normal**	**Periodontitis**
Type of tissue/cell	Gingival tissue		Gingival tissue		Gingival tissue	
Platform	Affymetrix Human Genome U133 Plus 2.0 Array		Affymetrix Human Genome U133 Plus 2.0 Array		Affymetrix Human Genome U133 Plus 2.0 Array	
Sample size	64	183	69	241	3	3
Age (mean)	42.0		39.3		35.0	58.0
Gender (Male: Female)	5:5		5:5		8:6	7:7
Pocket depth (mm)	1 ~ 4	5 ~ 11	3.9 ± 0.7		1.6 ± 0.2	6.5 ± 2.0
References	Demmer et al. (12)		Kebschull et al. (14)		Abe et al. (15)	

	GSE20966		GSE25724	
	Normal	**T2DM**	**Normal**	**T2DM**
Type of tissue/cell	Beta-cells		Pancreatic islet	
Platform	Affymetrix Human X3P Array		Affymetrix Human Genome U133A Array	
Sample size	10	10	7	6
Gender (Male: Female)	6: 4	7: 3	4: 33: 3	
Age (year)	60.3 ± 3.0	67.3 ± 4.3	58.1 ± 12.8	70.5 ± 7.3
BMI	30.6 ± 3.2	30.9 ± 4.0	24.8 ± 1.9	26.0 ± 1.8
References	Marselli et al. (11)		Dominguez et al. (16)	

genes. Terms with ≥ 5 associated genes and ≥ 2 DEGs in each experiment were included; FDR adjustment of p-values was performed using the Benjamini-Hochberg method.

PPI Network Construction and Module Selection

Protein–protein interaction (PPI) network was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database (20) to identify the relevant pathways and functions of common DEGs. The minimum required interaction score used to construct the PPI network was 0.4. We used molecular complex detection (MCODE) (21) for optimal module selection. The parameters used in MCODE were as follows: degree cutoff, 2; cluster finding, node score cutoff, 0.2; k-core, 2; and max.depth, 100. CytoHubba (22) was used to identify the highest linkage hub gene in the network. The parameters used in CytoHubba were as follows: the top 10 nodes ranked: degree, display options: display the expanded subnetwork.

Receiver Operating Characteristic (ROC) Curves Analysis

To confirm the diagnostic performance of hub genes, ROC curves were analyzed. The package used for the analysis was the r package, pROC (v1.18.0).

Experimental Validation of Hub Genes

For experimental validation, we recruited samples from 5 normal, 9 patients with periodontitis, and 4 patients with periodontitis with diabetes. Inclusion criteria included individuals aged 20 years or older. Exclusion criteria included subjects with current or previous history of significant systemic diseases, such as cardiovascular accidents, cancer, or renal failure. Patients with a prior history of using non-steroidal anti-inflammatory drugs, antibiotics or blood thinner were also excluded. Periodontal condition of subjects was assessed, then moderate or severe periodontitis patients were selected for periodontitis group. Periodontal conditions are according to the classification established in 1999 International Workshop of the American Academy of Periodontology (APP) for a Classification of Periodontal Diseases and Conditions. The presence of diabetes was determined by the value of Hb1Ac, and periodontitis patients showing more than 6.5% Hb1Ac were grouped into periodontitis with diabetes mellitus. Controls exhibited healthy periodontal tissue of less than 2 mm probing depth and no inflammatory signs such as gingival swelling, redness, and bleeding on probing. Briefly, tissue samples were obtained from periodontal pockets which were affected by periodontitis. Tissues of healthy controls without periodontitis were harvested from gingiva during gingivectomy or crown lengthening. Characteristics of clinical samples used for experimental validation is shown in **Supplementary Table 1**.

Quantitative Real-Time Polymerase Chain Reaction

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Pusan National University Dental Hospital (IRB: PNUDH 2020-032). Total RNA was extracted using the gingival tissues from healthy control, periodontitis patients and periodontitis patients with diabetes. Complementary DNA (cDNA) was synthesized using the Smart Gene Compact cDNA Synthesis kit (Smart Gene, South Korea). Quantitative real-time PCR was performed using the QuantStudio™ 3 Real-Time PCR System (Applied Biosystems, USA). Target mRNA expression relative to housekeeping gene expression (GAPDH) was calculated using the delta-delta Ct (ΔΔCT) method. A list of primers used in the quantitative real-time polymerase chain reaction is shown in **Supplementary Table 2**.

Statistical Analysis

We performed one-way ANOVA to check the differences between each group, and Tukey's multiple comparison test was used for the *post-hoc* test. P values less than 0.05 were considered significant. All statistical analyses were performed using R (v4.0.5).

RESULTS

Identification of Common DEGs for Each Disease

First, background correction and normalization of all expressed microarray data sets were performed to proceed with the DEG analysis. We identified DEGs in individual cohorts. We then identified overlapping genes among DEGs from individual cohorts. As a result, we identified 125 up-regulated and 27 down-regulated genes. Venn diagrams and volcano plots of DEGs identified in individual cohorts are shown in **Figure 2**.

Functional Enrichment Analysis of Common DEGs

GO and KEGG pathway analyses were performed to confirm the function of 152 common DEGs (**Supplementary Table 3**). The top 10 significantly enriched GO terms of the up-regulated and down-regulated genes are shown in **Supplementary Figure 1**, and the results of KEGG pathway analysis are shown in **Supplementary Figure 2**. The results of GO, KEGG analyses in each disease were described in **Supplementary Table 3**.

PPI Network Construction, Module Selection, and Identification of Hub Genes

To identify the protein–protein interactions of common DEGs, we used the STRING network-based protein interaction assay to create a PPI network. The most significant module consisted of 41 nodes and 93 edges. We identified highly connected genes in the most significantly constructed modules, and *PTPRC*, *HGF*, *RAC2*, *INPP5D*, *ENG*, *NES*, *CYBB*, *NCAM1*, *PDGFRA*, and *GATA2* were the top 10 highly connected genes (**Figure 3**). We then defined genes with a degree score of 10 or higher as hub genes. The hub genes identified in this study were *PTPRC*, *HGF*, *RAC2*, and *INPP5D*. The expression values of the hub genes in individual cohorts are shown in **Figure 3**, and the characteristics

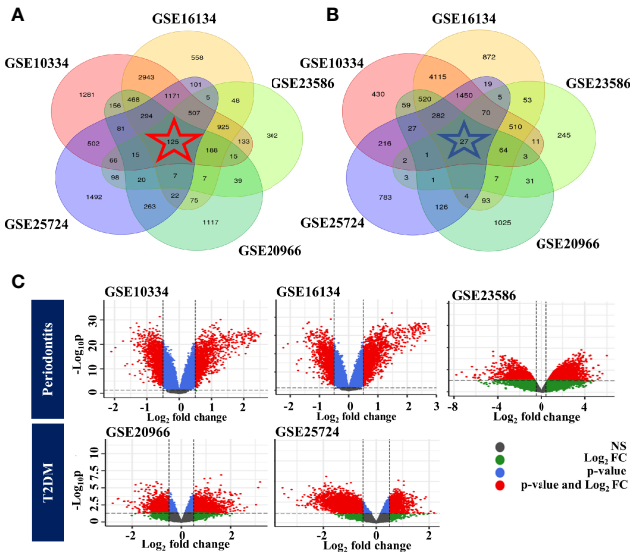


FIGURE 2 | Identification of common DEGs for each disease. **(A)** Venn diagram of upregulated genes. **(B)** Venn diagram of downregulated genes. **(C)** Volcano plot of DEGs identified in individual cohorts. The color of the points in the volcano plot indicates gray: not significant, green: passed only the log2 fold change, blue: passed only the p-value, red: passed the p-value and log2 fold change. PD, Below: T2DM.

are shown in **Supplementary Table 4**. We performed functional enrichment analysis of the hub genes. In GO-terms, the negative regulation of interleukin-6 production and regulation of B cell proliferation were significantly enriched (adjusted $p < 0.05$, **Figure 3**). In the KEGG pathway, the Fc gamma R-mediated phagocytosis and the Fc epsilon RI signaling pathway were significantly enriched (adjusted $p < 0.05$, **Figure 3**).

Identification of Diagnostic Performance of Hub Genes in Periodontitis and Type 2 Diabetes

ROC analysis was performed using the included GEO datasets and clinical samples to evaluate the diagnostic performance of each hub gene in T2DM and PD. GSE10334 dataset, and their AUC values (RAC2:0.863, INPP5D:0.818, PTPRC:0.749, HGF:0.697) were showed in **Supplementary Figure 3A**. GSE16134 dataset, and their AUC values (RAC2:0.890, INPP5D:0.836, PTPRC: 0.785, HGF:0.722) were showed in **Supplementary Figure 3B**. GSE23586 dataset, and their AUC values (RAC2:1.000, INPP5D:0.889, PTPRC:1.000, HGF:1.000) were showed in **Supplementary Figure 3C**. GSE20966 dataset, and their AUC values(RAC2:0.620, INPP5D:0.880, PTPRC:0.710, HGF:0.630) were showed in **Supplementary Figure 3D**. GSE25724 dataset, and their AUC values (RAC2:0.952, INPP5D:0.952, PTPRC:0.952, HGF:0.786) were showed in **Supplementary Figure 3E**. In periodontitis samples, and their AUC values (RAC2:0.956, INPP5D:1.000, PTPRC:1.000, HGF:0.733) were shown in **Supplementary Figure 4**. In periodontitis patients with diabetes samples, and their AUC values (RAC2:0.900, INPP5D:0.850, PTPRC:0.850, HGF:0.450) were showed in **Supplementary Figure 5**. These

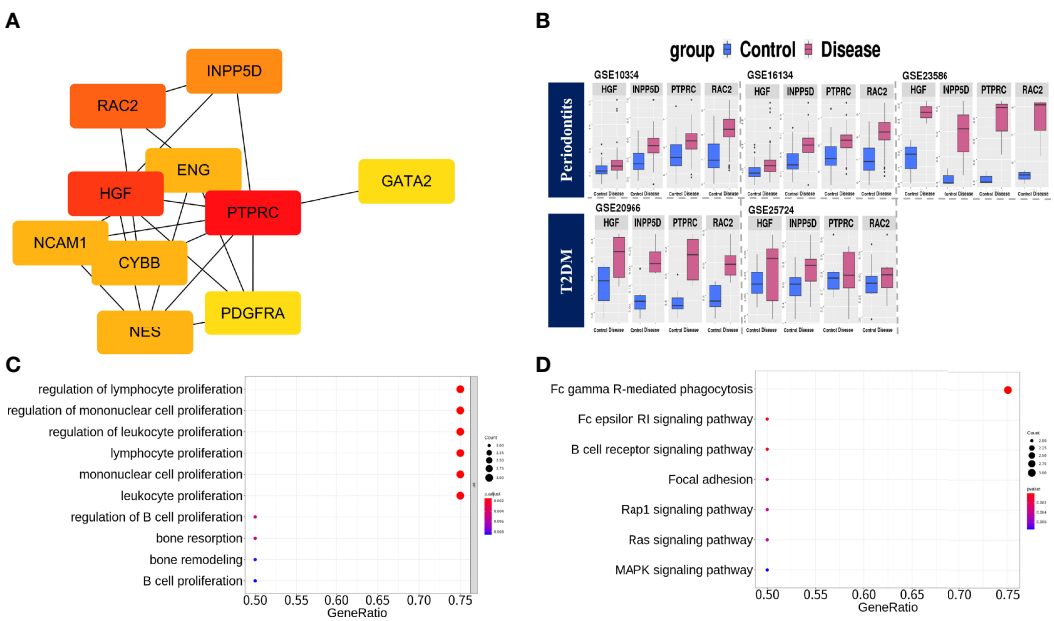


FIGURE 3 | PPI network construction, module selection, and identification of hub genes. **(A)** The top 10 highly connected genes based on degree score. **(B)** Boxplots of expression values in the individual cohorts of hub genes. Blue: Control group, Red: Disease group. **(C)** The significantly enriched BP terms of GO analysis with hub genes. The meaning of the dot color is as follows: red: adjusted p-value < 0.002, purple: < 0.004, blue < 0.008. **(D)** KEGG pathway analysis using hub genes. The meaning of the dot color is as follows: red: adjusted p-value < 0.002, purple: < 0.004, blue < 0.006.

results suggest that all three hub genes except HGF have good diagnostic performance.

Validation of the Hub Genes in Clinical Samples

We performed a Real-Time qPCR assay using gingival tissues from healthy controls, periodontitis patients, and periodontitis patients with diabetes to experimentally validate the hub gene. As a result, there was no significant difference in HGF between the groups, and INPP5D was significantly increased only in the periodontitis patient group ($p < 0.05$). RAC2 and PTPRC were significantly increased in the periodontitis ($p<0.005$) and periodontitis patients with T2DM group ($p < 0.005$) compared with the healthy control group. Real-Time qPCR results are shown in **Figure 4**.

DISCUSSION

Recently, studies on the molecular pathogenesis of diseases using bioinformatics tools have emerged in the field of biomedicine. Such studies have revealed numerous findings on crucial factors that are important for the development and progression of the disease as well as for linking genes between various diseases. PD and diabetes are chronic diseases with a complex mechanism and a bidirectional relationship (23). The risk of PD is known to increase by two to three times in people with diabetes (24) and the level of glycemic control is key in determining the risk (25). To date, research on PD and diabetes has mainly focused on T2DM, but studies have shown that type 1 diabetes is also associated with periodontal destruction (26). These previous findings indicate that there may be a potential common key gene that causes both PD and diabetes.

In this study, we identified 152 common DEGs, including 125 up-regulated genes and 27 down-regulated genes. 152 common DEGs were significantly enriched in focal adhesion, cAMP signaling pathway, and Fc gamma R-mediated phagocytosis signaling pathway, which are known to be associated with immune response. Focal adhesion is known to up-regulate several adhesion molecules due to the activation of endothelial cells in the inflammatory response and trigger the interaction of these cells with white blood cells (27). cAMP regulates a number of key pathways that affect the immune system and, depending on the cell type, can trigger both pro-inflammatory and anti-inflammatory effects. It is known that poor regulation of this signaling can affect the pathogenesis of inflammatory skin diseases such as atopic dermatitis and psoriasis (28). Fc gamma R mediated phagocytosis controls humoral and innate immunity essential for response to infection and chronic inflammation (29). These results indirectly indicate that the signaling pathways that control various inflammations are involved in both chronic inflammatory diseases. However, we tried to identify potential target genes or signaling pathways that could be more important of these signaling pathways. Therefore, we identified highly connected hub genes among 152 common DEGs and tried to find signaling pathways containing these genes. As a result, the highly linked hub genes were *PTPRC*, *RAC2*, *HGF*, and *INPP5D*. In functional enrichment analysis of these genes, it was the most significantly enriched Fc gamma R mediated phagocytosis, and

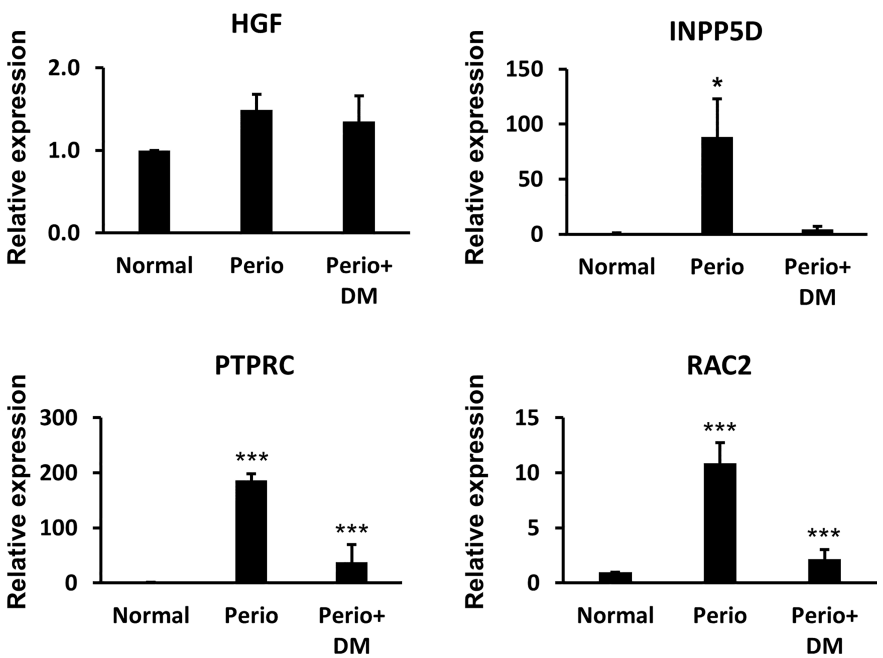


FIGURE 4 | The results of Quantitative reverse transcription-PCR between healthy controls, periodontitis patients, and periodontitis patients with diabetes. *$p < 0.05$, ***$p < 0.005$.

PTPRC, *RAC2*, *INPP5D* hub genes included in this signaling pathway were thought to be potential therapeutic targets.

Protein tyrosine phosphatase, receptor type, C (PTPRC) plays a key role in the signaling regulation of integrin and cytokine receptors, lymphocyte development, and antigen receptor signaling (30). Moreover, it differentially regulates Fc-gamma R-mediated tyrosine phosphorylation in polymorphonuclear neutrophil activation (31). In diabetes, protein tyrosine phosphatase is known to act as a negative regulator of insulin signaling, and expression of PTPRC is associated with residual β-cell function in type 1 diabetes (32). In addition, there is sufficient evidence that PTPRC acts as an important regulator of immune cell function, protection against viral infection, and is associated with immunodeficiency and viral susceptibility (1, 33). These previous findings and our findings support the hypothesis that upregulation of PTPRC regulates viral susceptibility and immune cell function and contributes to the pathogenesis of PD and diabetes. *RAC2* encodes a small GTPase of guanosine triphosphate (GTP) metabolic protein. Genetic mutations in *RAC2* cause severe phagocytic immunodeficiency characterized by severe infection in infancy (34). There are three isoforms of RAC, and *RAC1* and *RAC2* are differentially activated. Previous studies have shown decreased NADPH oxidase activation and IgG-mediated phagocytosis in neutrophils in *Rac2*-/- mice (35). In addition, NADPH oxidase preferentially interacts with *RAC2* (36). *RAC2* is a T1D candidate gene located at the 22q12.3 chromosomal risk locus (37). Functional knockdown experiments of *Rac2* showed that *Rac2* knockdown increases cytokine-induced apoptosis (38). Moreover, *Rac2*-deficient mice are much more susceptible to PD than WT controls, but the IL-17 response has not been determined. Nonetheless, mice deficient in *Rac2* display abundant mononuclear cell infiltration in the junctional epithelium and gingival connective tissue (39). *INPP5D*, also known as *SHIP1*, is an enzyme with phosphatase activity. Loss of *Inpp5d* function in mice causes both macrophages and B lymphocytes to become hypersensitive to stimuli (40). *INPP5D* partially attenuates BCR signaling through its association with Fc gamma receptor IIB and acts as an effector for other inhibitory receptors in numerous immune cell types (41). Small molecule inhibitors of *INPP5D* are known to improve obesity and the metabolic syndrome related to aging and diet (42). The *RAC2* and *INPP5D* genes are still insufficiently studied in both chronic diseases, but given the role revealed in previous studies, the upregulation of both genes could be a novel biomarker for chronic hyperglycemia.

Fc gamma R-mediated phagocytosis occurs when the Fc gamma receptor of monocyte-macrophages or neutrophils binds to IgG (43). The Fc gamma receptor is an essential participant in many immune system effector functions, such as the release of inflammatory mediators, antibody-dependent cytotoxicity, and phagocytosis (44, 45). Recent studies have confirmed that Fc gamma receptor glycan modifications are also important for interactions with antibodies and downstream immune responses (46, 47). Additionally, these previous studies suggest a role in inflammatory diseases and potential new therapeutic targets. Our findings support previous findings. Therefore, we believe that three hub genes that are commonly upregulated in both chronic diseases may be potential therapeutic targets.

There are several limitations to our study, and they are as follows: 1. In the included studies, information on various factors related to the disease was not provided and thus could not be evaluated. 2. Although we tried to minimize the heterogeneity of expression values according to the platform, only the microarray studies were included, so genes with low expression values may have been excluded. 3. Although the hub genes were experimentally validated through clinical samples, heterogeneity exists in the results because they were obtained from gingival tissue from PD patients with T2DM, not from pancreatic tissue.

CONCLUSIONS

The three hub genes identified in our study are involved in immune response regulation and insulin resistance and are all involved in the Fc gamma R mediated phagocytosis signaling pathway. In addition, previous studies have suggested an important role of the Fc gamma receptor in chronic inflammatory diseases and as a new therapeutic target. Therefore, we believe that up-regulation of these three genes affects two chronic diseases and can be potential therapeutic targets. Finally, we believe that further validation through experiments on these genes is also necessary.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

KJ: conceptualization, data analysis, data curation, and writing original draft. KEJ: conceptualization, data analysis, and data curation. HM: data curation. LH: investigation. YY: investigation. KJW: investigation. KY: investigation. LEY: methodology and investigation. JJY: methodology and investigation. HHJ: experimental validation. KEK: experimental validation. KTW: experimental validation, writing review. KYH: supervision, writing review, and editing. PHR: supervision, writing review, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Research Foundation of Korea (2020R1A2C1005203, 2020R1C1C1003741, and NRF-2018R1A5A2023879).

ACKNOWLEDGMENTS

The authors thank the contributors to the GEO (http://www.ncbi.nlm.nih.gov/geo/) database for sharing data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2021.724278/full#supplementary-material

Supplementary Figure 1 | The results of GO analysis using common DEGs. The color of the dots indicates the p-value and size is the number of included genes. The meaning of the dot color is as follows: red: adjusted p-value < 0.002, purple: < 0.004, blue < 0.008. **(A)** upregulated genes **(B)** downregulated genes

Supplementary Figure 2 | The results of KEGG pathway analysis using common DEGs. The color of the dots indicates p-value and size is the numbers of included genes. The meaning of the dot color is as follows: red: adjusted p-value < 0.002, purple: < 0.004, blue < 0.006. A: upregulated genes B: downregulated gene.

Supplementary Figure 3 | The results of ROC analysis for hub genes in included gene expression omnibus datasets. **(A)** GSE10334, **(B)** GSE16134, **(C)** GSE23586, **(D)** GSE20966, **(E)** GSE25724.

Supplementary Figure 4 | The results of ROC analysis for hub genes in periodontitis patient samples. **(A)** RAC2, **(B)** INPP5D, **(C)** PTPRC, **(D)** HGF.

Supplementary Figure 5 | The results of ROC analysis for hub genes in periodontitis with diabetes patient samples. **(A)** RAC2, **(B)** INPP5D, **(C)** PTPRC, **(D)** HGF.

Supplementary Table 3 | The results of GO and KEGG pathway analyses.

REFERENCES

1. Hendriks WJ, Pulido R. Protein Tyrosine Phosphatase Variants in Human Hereditary Disorders and Disease Susceptibilities. *Biochim Biophys Acta* (2013) 1832(10):1673–96. doi: 10.1016/j.bbadis.2013.05.022
2. Khumaedi AI, Purnamasari D, Wijaya IP, Soeroso Y. The Relationship of Diabetes, Periodontitis and Cardiovascular Disease. *Diabetes Metab Syndr* (2019) 13(2):1675–8. doi: 10.1016/j.dsx.2019.03.023
3. Tsai C, Hayes C, Taylor GW. Glycemic Control of Type 2 Diabetes and Severe Periodontal Disease in the US Adult Population. *Community Dent Oral Epidemiol* (2002) 30(3):182–92. doi: 10.1034/j.1600-0528.2002.300304.x
4. Emrich LJ, Shlossman M, Genco RJ. Periodontal Disease in non-Insulin-Dependent Diabetes Mellitus. *J Periodontol* (1991) 62(2):123–31. doi: 10.1902/jop.1991.62.2.123
5. Buchwald H, Oien DM. Metabolic/bariatric Surgery Worldwide 2011. *Obes Surg* (2013) 23(4):427–36.
6. Sundar C, Ramalingam S, Mohan V, Pradeepa R, Ramakrishnan M. Periodontal Therapy as an Adjunctive Modality for HbA1c Reduction in Type-2 Diabetic Patients. *J Educ Health promotion* 7(2018).
7. Peng Y. A Novel Ensemble Machine Learning for Robust Microarray Data Classification. *Comput Biol Med* (2006) 36(6):553–73.
8. Gan Z, Zou F, Zeng N, Xiong B, Liao L, Li H, et al. Wavelet Denoising Algorithm Based on NDOA Compressed Sensing for Fluorescence Image of Microarray. *IEEE Access* (2019) 7:13338–46.
9. Gan Z, Zeng N, Zou F, Chen J, Du M, Liao L, et al. Multilevel Segmentation Optimized by Physical Information for Gridding of Microarray Images. *IEEE Access* (2019) 7:32146–53.
10. Slonim DK, Yanai I. Getting Started in Gene Expression Microarray Analysis. *PloS Comput Biol* (2009) 5(10):e1000543. doi: 10.1371/journal.pcbi.1000543
11. Marselli L, Thorne J, Dahiya S, Sgroi DC, Sharma A, Bonner-Weir S, et al. Gene Expression Profiles of Beta-Cell Enriched Tissue Obtained by Laser Capture Microdissection From Subjects With Type 2 Diabetes. *PloS One* (2010) 5(7):e11499.
12. Demmer RT, Behle JH, Wolf DL, Handfield M, Kebschull M, Celenti R, et al. Transcriptomes in Healthy and Diseased Gingival Tissues. *J Periodontol* (2008) 79(11):2112–24. doi: 10.1902/jop.2008.080139
13. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for Functional Genomics Data Sets–Update. *Nucleic Acids Res* (2013) 41(Database issue):D991–5. doi: 10.1093/nar/gks1193
14. Kebschull M, Demmer RT, Grun B, Guarnieri P, Pavlidis P, Papapanou PN. Gingival Tissue Transcriptomes Identify Distinct Periodontitis Phenotypes. *J Dent Res* (2014) 93(5):459–68. doi: 10.1177/0022034514527288
15. Abe D, Kubota T, Morozumi T, Shimizu T, Nakasone N, Itagaki M, et al. Altered Gene Expression in Leukocyte Transendothelial Migration and Cell Communication Pathways in Periodontitis-Affected Gingival Tissues. *J Periodontal Res* (2011) 46(3):345–53. doi: 10.1111/j.1600-0765.2011.01349.x
16. Dominguez V, Raimondi C, Somanath S, Bugliani M, Loder MK, Edling CE, et al. Class II Phosphoinositide 3-Kinase Regulates Exocytosis of Insulin Granules in Pancreatic Beta Cells. *J Biol Chem* (2011) 286(6):4216–25. doi: 10.1074/jbc.M110.200295
17. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance Stabilization Applied to Microarray Data Calibration and to the Quantification of Differential Expression. *Bioinformatics* (2002) 18 Suppl:1, S96–104. doi: 10.1093/bioinformatics/18.suppl_1.s96
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007
19. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
20. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING V11: Protein-Protein Association Networks With Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res* (2019) 47(D1):D607–13. doi: 10.1093/nar/gky1131
21. Bader GD, Hogue CW. An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks. *BMC Bioinf* (2003) 4:2. doi: 10.1186/1471-2105-4-2
22. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. Cytohubba: Identifying Hub Objects and Sub-Networks From Complex Interactome. *BMC Syst Biol* (2014) 8 Suppl:4, S11. doi: 10.1186/1752-0509-8-S4-S11
23. Preshaw PM, Bissett SM. Periodontitis and Diabetes. *Br Dent J* (2019) 227 (7):577–84.
24. Preshaw PM, Alba AL, Herrera D, Jepsen S, Konstantinidis A, Makrilakis K, et al. Periodontitis and Diabetes: A Two-Way Relationship. *Diabetologia* (2012) 55(1):21–31. doi: 10.1007/s00125-011-2342-y
25. Tsai C, Hayes C, Taylor GW. Glycemic Control of Type 2 Diabetes and Severe Periodontal Disease in the US Adult Population. *Community Dent Oral Epidemiol* (2002) 30(3):182–92.
26. Lalla E, Cheng B, Lal S, Kaplan S, Softness B, Greenberg E, et al. Diabetes Mellitus Promotes Periodontal Destruction in Children. *J Clin Periodontol* (2007) 34(4):294–8. doi: 10.1111/j.1600-051X.2007.01054.x
27. Gonzalez-Amaro R, Diaz-Gonzalez F, Sanchez-Madrid F. Adhesion Molecules in Inflammatory Diseases. *Drugs* (1998) 56(6):977–88. doi: 10.2165/00003495-199856060-00003
28. Raker VK, Becker C, Steinbrink K. The cAMP Pathway as Therapeutic Target in Autoimmune and Inflammatory Diseases. *Front Immunol* (2016) 7:123.
29. Ben Mkaddem S, Benhamou M, Monteiro RC. Understanding Fc Receptor Involvement in Inflammatory Diseases: From Mechanisms to New Therapeutic Tools. *Front Immunol* (2019) 10:811. doi: 10.3389/fimmu.2019.00811
30. Hermiston ML, Xu Z, Weiss A. CD45: A Critical Regulator of Signaling Thresholds in Immune Cells. *Annu Rev Immunol* (2003) 21:107–37. doi: 10.1146/annurev.immunol.21.120601.140946
31. Hoffmeyer F, Witte K, Gebhardt U, Schmidt RE. The Low Affinity Fc Gamma RIIa and Fc Gamma RIIIb on Polymorphonuclear Neutrophils are Differentially Regulated by CD45 Phosphatase. *J Immunol* (1995) 155 (8):4016–23.

32. Akerman L, Ludvigsson J, Casas R. Low C-Peptide Levels and Decreased Expression of TNF and CD45 in Children With High Risk of Type 1 Diabetes. *Clin Immunol* (2013) 148(1):4–15. doi: 10.1016/j.clim.2013.03.011
33. Thiel N, Zischke J, Elbasani E, Kay-Fedorov P, Messerle M. Viral Interference With Functions of the Cellular Receptor Tyrosine Phosphatase CD45. *Viruses* (2015) 7(3):1540–57. doi: 10.3390/v7031540
34. Kurkchubasche AG, Panepinto JA, Tracy TFJr., Thurman GW, Ambruso DR. Clinical Features of a Human Rac2 Mutation: A Complex Neutrophil Dysfunction Disease. *J Pediatr* (2001) 139(1):141–7. doi: 10.1067/mpd.2001.114718
35. Kim C, Dinauer MC. Rac2 is an Essential Regulator of Neutrophil Nicotinamide Adenine Dinucleotide Phosphate Oxidase Activation in Response to Specific Signaling Pathways. *J Immunol* (2001) 166(2):1223–32.
36. Dorseuil O, Reibel L, Bokoch GM, Camonis J, Gacon G. The Rac Target NADPH Oxidase P67phox Interacts Preferentially With Rac2 Rather Than Rac1 (∗). *J Biol Chem* (1996) 271(1):83–8.
37. Courjal F, Chuchana P, Theillet C, Fort P. Structure and Chromosomal Assignment to 22q12 and 17qter of the Ras-Related Rac2 and Rac3 Human Genes. *Genomics* (1997) 44(2):242–6. doi: 10.1006/geno.1997.4871
38. Floyel T, Mirza AH, Kaur S, Frorup C, Yarani R, Storling J, et al. The Rac2 GTPase Contributes to Cathepsin H-Mediated Protection Against Cytokine-Induced Apoptosis in Insulin-Secreting Cells. *Mol Cell Endocrinol* (2020) 110993. doi: 10.1016/j.mce.2020.110993
39. Sima C, Gastfreund S, Sun C, Glogauer M. Rac-Null Leukocytes are Associated With Increased Inflammation-Mediated Alveolar Bone Loss. *Am J Pathol* (2014) 184(2):472–82.
40. Helgason CD, Damen JE, Rosten P, Grewal R, Sorensen P, Chappel SM, et al. Targeted Disruption of SHIP Leads to Hemopoietic Perturbations, Lung Pathology, and a Shortened Life Span. *Genes Dev* (1998) 12(11):1610–20. doi: 10.1101/gad.12.11.1610
41. Pauls SD, Marshall AJ. Regulation of Immune Cell Signaling by SHIP1: A Phosphatase, Scaffold Protein, and Potential Therapeutic Target. *Eur J Immunol* (2017) 47(6):932–45. doi: 10.1002/eji.201646795
42. Srivastava N, Iyer S, Sudan R, Youngs C, Engelman RW, Howard KT, et al. A Small-Molecule Inhibitor of SHIP1 Reverses Age- and Diet-Associated Obesity and Metabolic Syndrome. *JCI Insight* 1(11). doi: 10.1172/jci.insight.88544
43. Anderson CL, Shen L, Eicher DM, Wewers MD, Gill JK. Phagocytosis Mediated by Three Distinct Fc Gamma Receptor Classes on Human Leukocytes. *J Exp Med* (1990) 171(4):1333–45. doi: 10.1084/jem.171.4.1333
44. Salmon JE, Edberg JC, Brogle NL, Kimberly RP. Allelic Polymorphisms of Human Fc Gamma Receptor IIA and Fc Gamma Receptor IIIB. Independent Mechanisms for Differences in Human Phagocyte Function. *J Clin Invest* (1992) 89(4):1274–81. doi: 10.1172/JCI115712
45. Lobell RB, Austen KF, Katz HR. Fc Gamma R-Mediated Endocytosis and Expression of Cell Surface Fc Gamma RIIb1 and Fc Gamma RIIb2 by Mouse Bone Marrow Culture-Derived Progenitor Mast Cells. *J Immunol* (1994) 152 (2):811–8.
46. Hayes JM, Wormald MR, Rudd PM, Davey GP. Fc Gamma Receptors: Glycobiology and Therapeutic Prospects. *J Inflamm Res* (2016) 9:209–19. doi: 10.2147/JIR.S121233
47. Hayes JM, Frostell A, Cosgrave EF, Struwe WB, Potter O, Davey GP, et al. Fc Gamma Receptor Glycosylation Modulates the Binding of IgG Glycoforms: A Requirement for Stable Antibody Interactions. *J Proteome Res* (2014) 13 (12):5471–85. doi: 10.1021/pr500414q

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fendo.2022.818831

Coffee Types and Type 2 Diabetes Mellitus: Large-Scale Cross-Phenotype Association Study and Mendelian Randomization Analysis

Xinpei Wang[1], *Jinzhu Jia*[1,2*] *and Tao Huang*[3,4,5*]

[1] Department of Biostatistics, School of Public Health, Peking University, Beijing, China, [2] Center for Statistical Science, Peking University, Beijing, China, [3] Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, China, [4] Department of Global Health, School of Public Health, Peking University, Beijing, China, [5] Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing, China

OPEN ACCESS

Edited by:
Changwei Li,
Tulane University School of Public Health and Tropical Medicine, United States

Reviewed by:
Ping Zeng,
Xuzhou Medical University, China
Hao Peng,
Soochow University, China

***Correspondence:**
Jinzhu Jia
jzjia@math.pku.edu.cn
Tao Huang
huangtaotao@pku.edu.cn

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *20 November 2021*
Accepted: *10 January 2022*
Published: *11 February 2022*

Citation:
Wang X, Jia J and Huang T (2022) Coffee Types and Type 2 Diabetes Mellitus: Large-Scale Cross-Phenotype Association Study and Mendelian Randomization Analysis. Front. Endocrinol. 13:818831. doi: 10.3389/fendo.2022.818831

Purpose: To explore whether coffee intake is associated with the risk of type 2 diabetes mellitus (T2DM) from a genetic perspective, and whether this association remains the same among different types of coffee consumers.

Methods: We utilized the summary-level results of 12 genome-wide association studies. First, we used linkage disequilibrium score regression and cross-phenotype association analysis to estimate the genetic correlation and identify shared genes between coffee intake and T2DM in addition to some other T2DM-related phenotypes. Second, we used Mendelian randomization (MR) analysis to test whether there is a significant genetically predicted causal association between coffee intake and the risk of T2DM or other T2DM-related phenotypes. For all the analyses above, we also conducted a separate analysis for different types of coffee consumers, in addition to total coffee intake.

Results: Genetically, choice for ground coffee was significantly negatively associated with the risk of T2DM and some other related risks. While coffee intake and choice for decaffeinated/instant coffee had significant positive correlation with these risks. Between these genetically related phenotypes, there were 1571 genomic shared regions, of which 134 loci were novel. Enrichment analysis showed that these shared genes were significantly enriched in antigen processing related biological processes. MR analysis indicated that higher genetically proxied choice for ground coffee can reduce the risk of T2DM (T2DM: b: -0.2, p-value: 4.70×10^{-10}; T2DM adjusted for body mass index (BMI): b: -0.11, p-value: 4.60×10^{-5}), and BMI (b: -0.08, p-value: 6.50×10^{-5}).

Conclusions: Compared with other types of coffee, ground coffee has a significant negative genetic and genetically predicated causal relationship with the risk of T2DM. And this association is likely to be mediated by immunity. The effect of different coffee types on T2DM is not equal, researchers on coffee should pay more attention to distinguishing between coffee types.

Keywords: type 2 diabetes mellitus, coffee types, coffee intake, cardiac metabolic risks, Mendelian randomization

1 INTRODUCTION

Coffee is one of the most widely consumed beverages in the world, with about 500 billion cups consumed yearly (1). Coffee contains a variety of chemical substances, although some of which have clear health effects, such as chlorogenic acids (CGAs), which can protect the body from oxidative damage (2), and ochratoxin A, which would aggravate obesity (3), there are still some coffee ingredients whose functions have not been clarified yet. And the content of components in coffee is also affected by various factors such as processing (4). Therefore, research on the effects of coffee types on human health despite being complicated is important.

In recent years, a plethora of studies have been published focusing on the association between coffee intake and type 2 diabetes mellitus (T2DM). Cumulative evidence from observational studies suggests a significant relationship between coffee intake and T2DM (5–7), but observational research is susceptible to confounding factors and reverse causality (8). This problem is somewhat overcome by randomized clinical trials (RCTs), but current RCTs on the impact of coffee intake on T2DM are faced with problems of short intervention period and inconsistent results (9–11). Some researchers used genetic variations as instrumental variables (IVs) to study the effect of genetic proxied coffee intake on T2DM by Mendelian randomization (MR), but these studies hardly distinguished coffee types and the results are not consistent with observational studies, or even completely opposite (12–14). Because of the inconsistent nature of findings regarding the impact of habitual coffee intake on T2DM, more research is necessary before health care workers can make evidence-based recommendations.

It has been indicated that people's choices for coffee types vary (15) and there are differences in the chemical content of different coffee types (16–18). Previous research typically only investigated the association between total coffee consumption and T2DM risk without distinguishing among coffee types. Therefore, the results obtained from previous studies may not be applicable to every type of coffee, which may limit the clinical application thereof. Recently, the summary statistics of genome-wide association study (GWAS) on the total coffee intake and choices for different coffee types from the UK Biobank (UKB) large cohort have been released (19). Some of the loci found in the GWAS also have been reported to be associated with T2DM (19, 20), suggesting that there may be some genetic or causal links between them.

Therefore, the main objective of the present study is to explore whether coffee intake is associated with T2DM and some other T2DM-related phenotypes from a genetic perspective, and whether this association remains the same among different types of coffee consumers. Our study will shed light on the mechanism of the association between coffee consumption and T2DM and provide more precise guidance for coffee consumption.

2 MATERIALS AND METHODS

2.1 Study Design

This study aims to identify the association of coffee intake/coffee types and T2DM from a genetic perspective. In order to explore the association more comprehensively, we also included some other T2DM-related phenotypes. The flowchart of the study is shown in **Supplementary Figure 1**. First, we calculated the genetic correlation between coffee intake/coffee types and T2DM with the use of linkage disequilibrium (LD) score regression. For the phenotypes with significant genetic correlation, we further conducted cross-phenotype association analysis to identify the shared genes and enrichment analysis to find out the enriched biological processes and pathways of these genes. Second, we used MR to estimate the effect of genetically proxied coffee intake/coffee types on T2DM and T2DM-related phenotypes. It is hoped that this study will lead to new insights of the association between coffee intake and T2DM.

In addition to total coffee intake, our research also included an analysis of the choice for different coffee types, including decaffeinated coffee (any type), instant coffee, ground coffee (include espresso, filter etc), and other types of coffee. In addition to T2DM, we also analyzed some other T2DM-related phenotypes that consists of body mass index (BMI), fasting glucose (FG), fasting insulin (FI), insulin resistance (HOMA-IR), and beta-cell function (HOMA-β).

2.2 Data Sources

The data used in this study were obtained from 12 large-scale GWASs (**Supplementary Table 1**). The data of coffee intake and different coffee type choices came from GWAS of more than 320,000 UKB participants, who were asked how many cups of coffee they drank per day and what type of coffee they usually consumed (decaffeinated coffee (any type), instant coffee, ground coffee (include espresso, filter etc), and other types of coffee) (21), released by Neale lab (19).

The T2DM (adjusted and unadjusted for BMI) data came from the largest T2DM GWAS meta-analysis, with a total of 898,130 participants (22). BMI data was obtained from a meta-analysis of ∼700,000 individuals (23). And the data of FG, FI, HOMA-IR, HOMA - β were obtained from the meta-analysis of 21 GWAS by Jos é e Dupuis et al. (24). For details of these studies, please refer to **Supplementary Table 1** and corresponding literature.

2.3 Statistical Analysis

2.3.1 Linkage Disequilibrium Score Regression Analysis

As a convenient and frequently used tool in the study of genetic correlation between different phenotypes, LDSC relies on that the product of z-scores from two studies of phenotypes with non-zero genetic correlation is related to the LD score of the SNP under a polygenic model (25). This method requires only GWAS summary statistics and therefore is faster than other methods. We used LDSC to calculate the genetic correlation between coffee intake, choice for different coffee types and T2DM. We have adopted a series of quality control processes to ensure the accuracy of the results. To ensure the quality of SNP genotyping and imputation, only SNPs present in Hapmap3 were included in the analysis, and a variant was removed from the analysis if it had one of the following conditions: had missing values, INFO score <= 0.9, minor allele frequency (MAF) <= 0.01, p-value<0 or p-value>1, low sample size, not SNPs

(e.g., indels), strand ambiguous SNPs, had duplicated rs numbers. Besides, we also checked that the median value of the signed summary statistic column was close to the null median in order to make sure that this column is not mislabeled. Bonferroni correction was performed on the obtained p-value and tests were judged statistically significant at p-value < 1.43×10^{-3} (0.05/5/7).

2.3.2 Cross-Phenotype Association Analysis

Having identified the significantly genetically related phenotype-pairs, we further conducted cross-phenotype association (26) analysis to identify the shared genes of these phenotype pairs. This was implemented using the R software package 'CPASSOC' (26). This package uses the square root of the sample size of each phenotype as the weight and estimates the correlation matrix through summary statistics of all independent SNPs in the two phenotypes (26). This method not only considers the heterogeneity within the same phenotype or between different phenotypes, but also accounts for the potential kinship or population stratification between the participants (26).

SNPs with cross-phenotype association analysis p-value < 5×10^{-8} and single trait GWAS p-value < 0.05 were thought to have significant influence on both phenotypes. Then we used PLINK 1.9 software (27) to divide these SNPs into different independent clumping regions. SNPs with a distance less than 10,000kb and LD score R^2 > 0.001 were divided into the same clumping region. And the SNP with the lowest p-value in each region was taken as the index SNP. A genomic region was judged as a novel shared genomic region if it meets the following four conditions at the same time: (1) was identified in the cross-phenotype association study (p-value < 5×10^{-8}); (2) was not identified in the single phenotype GWAS (p-value > 5×10^{-8}); (3) was not reported to be associated with either of the two phenotypes in the GWAS catalog (28); (4) was not reported to be associated with either of the two phenotypes in the Phenoscanner (29, 30).

2.3.3 KEGG and GO Enrichment Analysis

In order to have a deeper understanding of the shared mechanism between coffee intake and T2DM, we performed Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analysis on the shared genes identified in the previous step. All analysis is done through the R package 'clusterProfiler' (31).

2.3.4 Mendelian Randomization Analysis

MR is currently a widely accepted method for assessing potential causal relationship for its unique advantages, compared with observational studies and RCTs (8). The core of MR is IV. In genetic epidemiology, IVs refer to genetic variations related to exposure but not directly related to outcomes and confounders. The selected IVs must satisfy three basic assumptions (8): first, it must be related to the exposure of interest; second, it must be independent of confounding factors; third, given exposure and confounders, the IVs are independent of the outcome.

We used two packages which are widely available to conduct MR: "TwoSampleMR" (32, 33) and "CAUSE" (34). "TwoSampleMR" is a convenient and frequently used tool in the study of two-sample MR. Using this package, we selected independent genetic variations with the threshold of p-value 5×10^{-8} and clump-kb 10000, clump-r2 0.01 as IVs, calculated causal estimates based on the association effect sizes of IVs with outcome and exposure, and then used a series of methods including inverse variance weighted method (IVW), MR-Egger regression, simple median, weighted median, penalized weighted median, simple mode, and weighted mode to obtain the overall causal estimate. We also tested whether the MR-Egger regression (33) intercept item was zero to check whether the horizontal pleiotropy was balanced.

Different from "TwoSampleMR", as a recently published method, "CAUSE" uses all SNPs in estimating causal effects, not just variants that are strongly correlated with exposure (34). This method relies on that if exposure has a causal effect on the outcome, then for any SNP that has a non-zero effect on the exposure, its association effect size with exposure and outcome should be related. Based on this, this method is thought to be able to distinguish between causality and horizontal pleiotropy (related and unrelated). If the result shows that the causal model is better to fit the data than the shared model (p-value <0.05), then we think that exposure has a causal effect on the outcome.

3 RESULTS

3.1 Genetic Correlation

There were extensive significant genetic correlations between the amount of coffee intake or choices for different coffee types and T2DM in addition to other T2DM-related phenotypes (**Figure 1**, **Supplementary Tables 2–6**). Without distinguishing between coffee types, we found that the amount of coffee intake was positively genetically associated with BMI (Rg = 0.2617, p-value = 1.12×10^{-30}).

When considering different coffee types, there was a considerable difference between various coffee type choices. Choosing decaffeinated coffee had a significant positive genetic correlation with T2DM (Rg = 0.1496, p-value = 1.00×10^{-4}) and BMI (Rg = 0.1541, p-value = 2.96×10^{-6}), compared to other types of coffee. Choice for instant coffee showed more significant associations. It was positively associated with T2DM (Rg = 0.1090, p-value = 1.00×10^{-3}), BMI (Rg = 0.1522, p-value = 1.73×10^{-7}), FG (Rg = 0.2084, p-value = 1.00×10^{-3}), FI (Rg = 0.3643, p-value = 6.40×10^{-6}), and HOMA-IR (Rg = 0.3989, p-value = 4.86×10^{-6}).

Conversely, choice for ground coffee showed a favorable relationship with all six phenotypes except HOMA-β (T2DM: Rg = -0.1805, p-value = 2.81×10^{-13}; T2DM adjusted for BMI: Rg = -0.0979, p-value = 3.00×10^{-4}; BMI: Rg = -0.2104, p-value = 6.40×10^{-26}; FG: Rg = -0.1903, p-value = 9.39×10^{-5}; FI: Rg = -0.3584, p-value = 6.88×10^{-8}; HOMA-IR: Rg = -0.3809, p-value = 2.72×10^{-8}). There was no significant genetic association between choice for coffee other than the above-mentioned types and T2DM or any related phenotype.

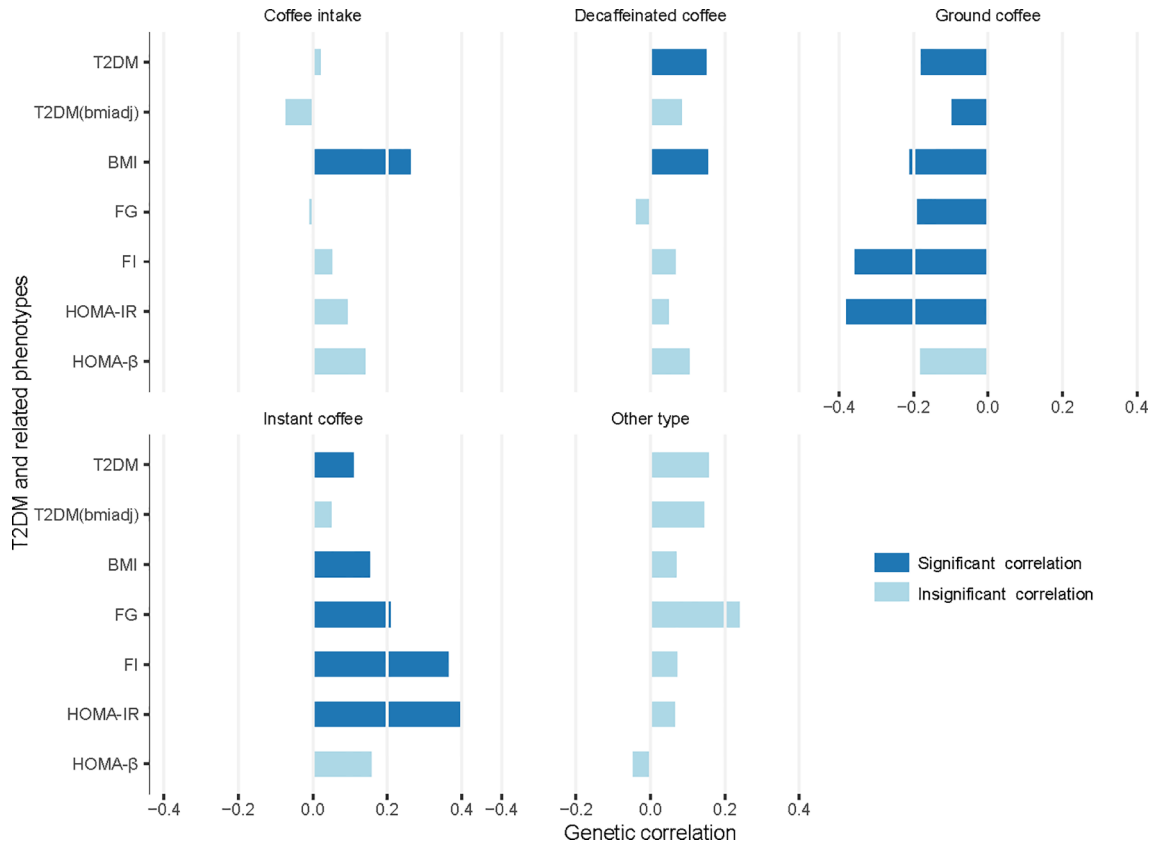


FIGURE 1 | Genetic correlation between coffee intake or choices for different coffee types and T2DM/related phenotypes. T2DM, type 2 diabetes mellitus; T2DM (bmiadj), T2DM adjusted for body mass index (BMI); FG, fasting glucose; FI, fasting insulin; HOMA-IR, insulin resistance; HOMA-B, beta-cell function; Significant correlation, p-value < 1.43×10^{-3} (0.05/5/7).

3.2 Cross-Phenotype Association Analysis

We identified multiple common genomic regions shared between the amount of coffee intake or choice for coffee types and T2DM or other T2DM-related phenotypes through cross-phenotype association analysis of trait pairs with significant genetic correlations (**Supplementary Tables 7–20**), including many novel regions which have not been reported in previous studies (**Tables 1A–D**).

In summary, there were 304 shared genomic regions between total amount of coffee consumed and BMI. Choice for decaffeinated coffee had 75 shared genetic regions with T2DM, 238 shared genetic regions with BMI. Choice for instant coffee and T2DM, BMI, FG, FI, HOMA-IR, had 111, 260, 4, 1, 1 shared regions, respectively. Choice for ground coffee shared the most genomic regions with T2DM and related phenotypes, including 136 shared with T2DM, 88 shared with T2DM (adjusted for BMI), 321 shared with BMI, 15 shared with FG, 7 shared with FI, 10 shared with HOMA-IR.

One of the most important findings of this analysis was the novel genomic regions which had not been reported by previous studies. We identified 134 novel regions totally. For example, we found that rs10492872 located on chromosome 16 was the shared site of choice for instant coffee and choice for ground coffee and BMI, and the effects of this locus on choice for instant coffee and ground coffee were opposite, similar to rs11660753 on chromosome 18. Another SNP, rs4988483, also located on chromosome 16, was the shared site of choice for ground coffee and T2DM (adjusted and unadjusted for BMI), suggesting that this SNP may affect both choice of ground coffee and risk of T2DM through a non-BMI-mediated pathway. Further analysis of these sites will help to understand the association between coffee and T2DM.

The discovery of these shared sites suggested that the genetic correlations between coffee intake or choice for different coffee types and T2DM were likely to be driven by these shared regions. Further research on these regions will help to gain a deeper understanding of the pathogenesis of T2DM and its association with coffee.

3.3 KEGG and GO Enrichment Analysis

The KEGG pathways and GO terms enriched by shared genes are shown in **Supplementary Figures 2–8**. We performed

TABLE 1A | Novel shared genomic regions shared between coffee intake and T2DM or T2DM-related phenotypes.

T2DM/related phenotypes	CHR	N	POS	KB	Index SNP	A1	A2	Coffee type		T2DM/related phenotypes		CPASSOC
								beta	p	beta	p	p
BMI	1	7	chr1:189946070-189961726	15.657	rs12046184	A	T	0.014	4.49E-04	-0.020	1.20E-06	2.93E-10
	1	19	chr1:99184056-99242884	58.829	rs1982703	G	A	0.005	6.02E-03	-0.009	2.90E-06	1.21E-08
	2	1	chr2:32920358-32920358	0.001	rs7564044	G	T	-0.008	2.21E-04	0.009	1.10E-04	3.75E-08
	2	1	chr2:49894154-49894154	0.001	rs10190188	C	T	-0.005	2.45E-03	0.009	8.70E-06	1.41E-08
	3	1	chr3:29088999-29088999	0.001	rs6796246	G	A	0.007	1.89E-05	0.008	1.30E-06	5.51E-09
	4	1	chr4:159851039-159851039	0.001	rs10031800	T	C	0.005	2.69E-03	-0.008	1.10E-05	3.71E-08
	4	1	chr4:173167398-173167398	0.001	rs836322	A	G	-0.005	8.46E-03	0.008	3.50E-06	4.95E-08
	4	1	chr4:28647811-28647811	0.001	rs10027492	T	A	0.005	6.90E-03	-0.008	1.40E-06	6.34E-09
	5	1	chr5:158488326-158488326	0.001	rs17718288	C	G	0.004	1.06E-02	-0.008	5.90E-06	3.97E-08
	8	1	chr8:118946541-118946541	0.001	rs10955841	A	G	0.005	3.61E-03	0.010	1.70E-07	1.64E-08
	8	1	chr8:78956658-78956658	0.001	rs2219968	A	G	-0.006	2.97E-04	-0.009	1.10E-06	3.87E-08
	9	1	chr9:73798371-73798371	0.001	rs1329767	A	C	-0.004	4.49E-02	0.010	1.40E-07	6.50E-09
	10	2	chr10:93644552-93790523	145.972	rs7917710	A	C	-0.006	1.56E-03	-0.010	3.50E-07	3.82E-08
	12	1	chr12:56895503-56895503	0.001	rs2657909	C	T	-0.006	7.81E-04	-0.010	3.90E-07	3.08E-08
	12	1	chr12:61854732-61854732	0.001	rs2009164	C	T	0.004	2.51E-02	-0.011	5.90E-08	1.36E-09
	14	1	chr14:30575613-30575613	0.001	rs6571334	G	C	0.008	6.72E-05	-0.009	2.50E-05	1.55E-09
	16	6	chr16:5850793-5854110	3.318	rs17792339	C	A	-0.007	7.59E-05	0.007	1.80E-05	1.17E-09
	17	2	chr17:41495423-41498187	2.765	rs11079338	G	A	0.006	1.15E-03	0.010	9.10E-08	9.23E-09
	18	1	chr18:22632927-22632927	0.001	rs9963409	T	C	-0.004	1.84E-02	0.009	9.80E-07	1.16E-08
	18	14	chr18:73124672-73159173	34.502	rs12165099	A	G	0.004	2.30E-02	-0.009	3.20E-07	4.50E-09
	22	6	chr22:38156183-38268922	112.74	rs2413485	T	C	0.005	3.23E-03	-0.010	7.00E-08	1.69E-10

TABLE 1B | Novel shared genomic regions shared between choice for decaffeinated coffee and T2DM or T2DM-related phenotypes.

T2DM/related phenotypes	CHR	N	POS	KB	Index SNP	A1	A2	Coffee type		T2DM/related phenotypes		CPASSOC
								beta	p	beta	p	p
T2DM	2	1	chr2:60588713-60588713	0.001	rs243016	T	A	-0.015	2.61E-02	-0.036	8.00E-08	1.22E-08
	3	1	chr3:136005792-136005792	0.001	rs7653249	G	C	0.016	3.26E-02	0.039	1.80E-07	2.68E-08
	3	24	chr3:94029191-94051397	22.207	rs4857339	T	C	0.015	1.96E-02	-0.034	1.80E-07	8.31E-09
	6	9	chr6:127798402-127817401	19	rs11154414	T	C	0.016	1.73E-02	0.034	2.20E-07	2.06E-08
	6	1	chr6:20725694-20725694	0.001	rs62397653	C	A	0.027	4.76E-02	-0.072	2.70E-07	4.95E-08
	7	2	chr7:102086552-102086605	0.054	rs77655131	T	C	-0.024	1.50E-02	0.053	5.40E-08	3.71E-09
	7	1	chr7:28222877-28222877	0.001	rs139048357	A	C	-0.063	2.65E-02	-0.160	4.60E-07	4.26E-08
	7	1	chr7:77044764-77044764	0.001	rs9641219	G	T	0.015	2.00E-02	-0.033	3.90E-07	3.37E-08
	8	3	chr8:110059269-110073120	13.852	rs6469227	G	T	-0.013	4.18E-02	-0.034	2.60E-07	4.13E-08
	8	23	chr8:116464988-116549176	84.189	rs12114740	C	T	0.014	3.39E-02	0.034	1.40E-07	2.17E-08
	10	1	chr10:12252217-12252217	0.001	rs2271804	A	G	-0.013	3.69E-02	-0.033	1.40E-07	3.57E-08
	10	1	chr10:99161831-99161831	0.001	rs7079477	G	C	-0.015	3.86E-02	0.038	1.30E-07	1.21E-08
	13	1	chr13:80591311-80591311	0.001	rs7328113	A	C	0.014	3.91E-02	-0.035	2.30E-07	2.59E-08
	14	1	chr14:74954059-74954059	0.001	rs7149930	T	C	-0.013	3.74E-02	-0.034	2.00E-07	2.35E-08
	20	3	chr20:20066701-20069826	3.126	rs73125628	T	C	-0.015	3.55E-02	-0.038	1.40E-07	1.79E-08
BMI	1	1	chr1:177394187-177394187	0.001	rs10913339	G	A	0.018	4.15E-02	0.013	6.30E-08	4.09E-08
	1	2	chr1:221715102-221766906	51.805	rs12079987	A	G	0.057	5.80E-03	-0.028	6.40E-07	1.47E-08
	2	3	chr2:241344240-241367465	23.226	rs10199929	T	A	-0.014	3.45E-02	-0.010	5.20E-08	3.39E-08
	8	4	chr8:50957454-51029371	71.918	rs1903311	G	A	-0.023	2.50E-03	-0.010	8.10E-08	1.46E-09
	9	1	chr9:28544375-28544375	0.001	rs10968649	G	T	-0.021	2.21E-02	0.013	2.20E-07	3.05E-08
	10	6	chr10:970426-1001215	30.79	rs2282419	C	G	-0.025	4.44E-03	0.013	9.20E-08	1.69E-09
	11	1	chr11:371265-371265	0.001	rs11246136	A	C	-0.022	3.48E-02	-0.016	6.90E-08	2.37E-08
	14	4	chr14:97236360-97255679	19.32	rs10149171	A	G	-0.039	4.25E-04	0.014	2.60E-06	7.04E-09
	14	2	chr14:99757151-99758235	1.085	rs2748805	T	C	0.025	2.29E-04	-0.009	2.50E-06	4.76E-09
	17	1	chr17:5005311-5005311	0.001	rs6502843	G	A	-0.022	7.62E-04	-0.008	6.90E-07	2.50E-08

TABLE 1C | Novel shared genomic regions shared between choice for ground coffee and T2DM or T2DM-related phenotypes.

T2DM/related phenotypes	CHR	N	POS	KB	Index SNP	A1	A2	Coffee type		T2DM/related phenotypes		CPASSOC
								beta	p	beta	p	p
T2DM	1	1	chr1:229594310-229594310	0.001	rs10916480	G	A	-0.023	1.45E-02	0.051	6.10E-08	8.17E-09
	1	2	chr1:26726127-26726967	0.841	rs113466616	A	G	-0.015	3.13E-02	0.035	1.80E-07	4.62E-08
	2	5	chr2:163639107-163649690	10.584	rs12614955	T	C	0.017	1.69E-02	0.035	9.70E-07	2.60E-08

(Continued)

TABLE 1C | Continued

T2DM/related phenotypes	CHR	N	POS	KB	Index SNP	A1	A2	Coffee type		T2DM/related phenotypes		CPASSOC
								beta	p	beta	p	p
	2	4	chr2:208910738-208921257	10.52	rs10932228	G	A	-0.017	7.71E-03	-0.033	3.30E-07	8.38E-09
	2	1	chr2:27541053-27541053	0.001	rs4665966	G	C	0.015	3.57E-02	0.042	6.00E-08	3.45E-09
	2	2	chr2:30878707-30879117	0.411	rs2602778	A	G	-0.018	8.21E-03	0.034	3.60E-07	2.84E-08
	3	3	chr3:170562564-170572896	10.333	rs12488260	T	G	-0.015	2.29E-02	-0.034	1.90E-07	8.69E-09
	3	1	chr3:186698711-186698711	0.001	rs6775869	C	T	-0.016	1.64E-02	-0.034	5.80E-07	1.63E-08
	5	1	chr5:44800917-44800917	0.001	rs10462082	C	A	0.015	2.02E-02	0.032	1.00E-06	3.05E-08
	5	1	chr5:86520965-86520965	0.001	rs13185500	A	T	-0.015	4.65E-02	0.038	1.30E-07	4.57E-08
	6	7	chr6:126072853-126097457	24.605	rs1977141	G	A	0.015	2.21E-02	0.033	3.60E-07	2.11E-08
	8	3	chr8:129579579-129592794	13.216	rs2395822	T	C	-0.017	1.46E-02	-0.034	6.30E-07	2.28E-08
	8	5	chr8:19780310-19813180	32.871	rs113023641	A	G	0.030	9.09E-04	-0.048	1.50E-07	5.32E-09
	8	1	chr8:41440849-41440849	0.001	rs183323983	C	A	-0.107	6.03E-04	-0.170	1.20E-06	1.06E-08
	8	1	chr8:41586822-41586822	0.001	rs7016707	T	G	-0.023	2.94E-02	0.053	2.80E-07	2.85E-08
	9	10	chr9:96970677-96994063	23.387	rs10821317	T	C	-0.014	3.77E-02	0.036	1.70E-07	3.54E-08
	10	6	chr10:13480317-13482073	1.757	rs4750356	G	A	0.016	1.81E-02	0.034	5.00E-07	1.78E-08
	11	3	chr11:2163932-2164990	1.059	rs7111447	T	A	-0.031	8.72E-03	0.067	1.60E-07	2.91E-08
	11	18	chr11:43635499-43696917	61.419	rs2862961	A	G	0.017	1.37E-02	-0.036	2.00E-07	1.72E-08
	12	4	chr12:26346146-26347517	1.372	rs2343869	G	A	-0.014	2.84E-02	-0.032	4.50E-07	2.63E-08
	14	3	chr14:104008159-104011429	3.271	rs12891360	T	C	0.030	1.78E-05	0.032	8.00E-06	4.53E-09
	16	1	chr16:1129010-1129010	0.001	rs4988483	A	C	0.036	1.00E-02	0.079	1.40E-06	2.39E-08
	16	1	chr16:69653696-69653696	0.001	rs39999	G	C	0.036	3.97E-05	-0.041	4.60E-06	4.38E-08
	17	1	chr17:76792179-76792179	0.001	rs1044486	A	G	-0.015	1.92E-02	-0.031	9.40E-07	4.61E-08
	20	1	chr20:31903533-31903533	0.001	rs721970	G	A	0.026	3.39E-02	0.069	2.00E-07	8.02E-09
T2DM (adjusted for BMI)	2	2	chr2:112878404-112880587	2.184	rs62157841	A	G	0.015	2.44E-02	0.039	5.00E-07	3.66E-08
	3	1	chr3:23252089-23252089	0.001	rs79433447	C	A	0.038	4.77E-02	0.110	1.40E-07	2.47E-08
	3	1	chr3:54876573-54876573	0.001	rs111494834	T	C	-0.036	3.33E-02	0.110	1.80E-07	6.70E-09
	9	1	chr9:139256766-139256766	0.001	rs3829109	A	G	-0.016	2.68E-02	-0.044	1.60E-07	2.07E-08
	10	2	chr10:64598621-64620042	21.422	rs911610	T	C	0.018	5.84E-03	-0.038	6.60E-07	3.55E-08
	11	1	chr11:65405600-65405600	0.001	rs2306363	T	G	0.017	3.57E-02	-0.049	1.80E-07	3.85E-08
	15	1	chr15:77847802-77847802	0.001	rs34075648	C	G	-0.014	4.07E-02	0.043	8.20E-08	1.67E-08
	16	1	chr16:1129010-1129010	0.001	rs4988483	A	C	0.036	1.00E-02	0.091	4.30E-07	1.86E-08
	18	5	chr18:46158145-46169673	11.529	rs299727	A	G	-0.023	1.91E-02	0.057	5.00E-07	2.83E-08
	22	1	chr22:50771629-50771629	0.001	rs149127232	C	A	0.025	2.87E-03	0.049	1.70E-06	1.84E-08
BMI	1	2	chr1:110686949-110687493	0.545	rs6682201	G	T	-0.017	6.82E-03	-0.008	3.60E-06	1.94E-08
	1	2	chr1:195316969-195322984	6.016	rs12731553	G	T	-0.017	1.45E-02	-0.009	1.40E-06	2.80E-08
	1	1	chr1:221057646-221057646	0.001	rs2738755	T	C	-0.015	2.68E-02	-0.009	2.10E-06	4.55E-08
	2	1	chr2:16617417-16617417	0.001	rs4281911	T	C	-0.014	3.27E-02	0.010	1.40E-07	4.47E-08
	3	7	chr3:141701527-141715543	14.017	rs3817176	T	C	0.035	7.59E-05	-0.012	3.50E-07	5.79E-09
	3	4	chr3:38625709-38647780	22.072	rs7374289	T	C	0.018	4.88E-03	0.009	1.80E-07	1.17E-09
	4	1	chr4:104978031-104978031	0.001	rs950882	C	T	0.018	1.13E-02	0.009	3.00E-06	4.60E-08
	4	1	chr4:180156210-180156210	0.001	rs17747559	A	G	0.017	9.83E-03	-0.009	3.80E-07	4.00E-08
	5	1	chr5:142892428-142892428	0.001	rs6876238	A	C	-0.014	2.79E-02	-0.009	6.80E-07	4.77E-08
	5	1	chr5:153490778-153490778	0.001	rs11948898	C	G	0.017	7.33E-03	0.008	2.00E-06	1.06E-08
	5	1	chr5:164585515-164585515	0.001	rs4400143	T	C	-0.017	1.26E-02	-0.009	1.00E-06	2.37E-08
	6	49	chr6:45017228-45328265	311.038	rs12205860	A	G	-0.015	2.63E-02	-0.010	1.90E-07	2.44E-09
	7	1	chr7:100774705-100774705	0.001	rs2227666	A	G	0.053	9.76E-05	-0.020	5.90E-07	7.26E-09
	9	1	chr9:104415321-104415321	0.001	rs1323425	C	T	-0.013	3.42E-02	-0.009	6.00E-07	1.74E-08
	10	1	chr10:10269019-10269019	0.001	rs7907578	T	G	0.042	3.08E-04	-0.015	1.30E-06	4.51E-08
	11	1	chr11:66568943-66568943	0.001	rs682842	T	C	-0.027	4.60E-05	-0.007	1.20E-04	4.54E-08
	13	1	chr13:111603363-111603363	0.001	rs7998604	C	A	-0.021	3.66E-03	-0.009	1.20E-05	3.78E-08
	13	1	chr13:98091008-98091008	0.001	rs1304392	C	T	-0.014	3.79E-02	-0.010	1.60E-07	8.30E-09
	14	1	chr14:60171047-60171047	0.001	rs9323353	G	A	0.027	4.80E-05	-0.009	7.80E-07	1.09E-08
	16	2	chr16:54120330-54123512	3.183	rs10492872	G	T	0.015	2.76E-02	0.009	1.30E-06	1.84E-08
	18	6	chr18:73122598-73147492	24.895	rs11660753	A	G	-0.016	2.04E-02	-0.009	4.90E-07	1.74E-08

TABLE 1D | Novel shared genomic regions shared between choice for instant coffee and T2DM or T2DM-related phenotypes.

T2DM/related phenotypes	CHR	N	POS	KB	Index SNP	A1	A2	Coffee type		T2DM/related phenotype		CPASSOC
								beta	p	beta	p	p
T2DM	1	1	chr1:112277693-112277693	0.001	rs127204	C	A	0.0152	4.32E-03	0.0320	1.00E-06	3.45E-08

(Continued)

TABLE 1D | Continued

T2DM/related phenotypes	CHR	N	POS	KB	Index SNP	A1	A2	Coffee type		T2DM/related phenotype		CPASSOC
								beta	p	beta	p	p
	1	6	chr1:26726127-26743903	17.777	rs113466616	A	G	0.0170	2.12E-03	0.0350	1.80E-07	5.92E-09
	1	1	chr1:229594310-229594310	0.001	rs10916480	G	A	0.0161	3.49E-02	0.0510	6.10E-08	1.27E-08
	2	1	chr2:165770457-165770457	0.001	rs16849922	C	A	-0.0122	2.40E-02	0.0330	3.30E-07	3.45E-08
	2	4	chr2:161354368-161368808	14.441	rs79523138	G	A	-0.0187	1.96E-02	0.0540	5.20E-08	3.05E-09
	4	17	chr4:157676228-157745957	69.73	rs4691382	G	A	-0.0150	1.33E-02	-0.0400	9.80E-08	6.78E-09
	5	1	chr5:59398807-59398807	0.001	rs73759013	C	A	-0.0188	4.67E-03	-0.0410	3.10E-07	2.53E-08
	6	3	chr6:153367613-153391618	24.006	rs7767938	C	T	0.0192	1.55E-03	0.0360	9.60E-07	2.77E-08
	7	2	chr7:104415434-104420060	4.627	rs4460308	C	T	0.0176	1.23E-03	0.0330	1.00E-06	1.63E-08
	7	1	chr7:28222877-28222877	0.001	rs139048357	A	C	0.0566	1.49E-02	-0.1600	4.60E-07	1.41E-08
	7	1	chr7:74134911-74134911	0.001	rs35005436	C	T	0.0174	1.52E-02	0.0470	2.50E-07	3.08E-08
	8	4	chr8:19780310-19808131	27.822	rs113023641	A	G	-0.0212	4.70E-03	-0.0480	1.50E-07	1.02E-08
	9	1	chr9:19093017-19093017	0.001	rs10963961	G	C	-0.0116	4.68E-02	0.0360	4.40E-07	4.34E-08
	9	8	chr9:97772285-97795176	22.892	rs10993460	C	T	0.0238	9.79E-03	-0.0580	3.00E-07	5.10E-09
	11	3	chr11:45875392-45877688	2.297	rs7951225	T	A	-0.0155	1.70E-02	-0.0410	2.50E-07	2.89E-08
	12	19	chr12:4237568-4257553	19.986	rs113966250	G	C	-0.0385	3.17E-02	-0.1200	7.00E-08	2.04E-09
	14	3	chr14:69446960-69447386	0.427	rs8016537	A	G	0.0164	1.81E-03	0.0310	1.00E-06	3.22E-08
	14	2	chr14:101309692-101309759	0.068	rs2295388	A	G	0.0154	1.56E-02	-0.0390	3.70E-07	1.75E-08
	17	1	chr17:37402280-37402280	0.001	rs2100651	A	C	-0.0198	9.53E-04	-0.0350	2.00E-06	4.62E-08
	17	2	chr17:17612187-17739163	126.977	rs77745346	T	C	-0.0234	4.17E-02	0.0700	1.60E-07	9.42E-09
	19	9	chr19:13139616-13212025	72.41	rs12151248	T	C	0.0274	7.43E-04	-0.0520	4.10E-07	1.18E-09
BMI	3	1	chr3:66668365-66668365	0.001	rs11128158	G	A	-0.012	3.40E-02	0.010	1.10E-07	7.94E-09
	4	2	chr4:81194090-81199966	5.877	rs3796606	G	A	0.016	3.52E-03	-0.008	2.70E-06	4.73E-08
	7	2	chr7:104435213-104441909	6.697	rs2470961	A	G	0.020	2.02E-04	0.008	4.60E-06	3.28E-08
	10	4	chr10:124929763-124947286	17.524	rs12218544	C	G	0.019	2.05E-02	-0.014	1.40E-07	4.29E-09
	12	1	chr12:117630765-117630765	0.001	rs4492907	C	T	0.011	3.13E-02	0.009	1.00E-07	2.34E-08
	12	6	chr12:84242417-84306965	64.549	rs922160	G	A	0.016	2.76E-03	0.009	3.90E-07	1.93E-08
	16	1	chr16:54120330-54120330	0.001	rs10492872	G	T	-0.013	1.75E-02	0.009	1.30E-06	2.43E-08
	17	1	chr17:35061859-35061859	0.001	rs6607342	C	T	0.017	2.53E-02	0.013	1.20E-07	4.55E-08
	18	1	chr18:73122598-73122598	0.001	rs11660753	A	G	0.013	2.18E-02	-0.009	4.90E-07	4.22E-08
	20	5	chr20:50403227-50505843	102.617	rs6021437	C	T	-0.016	4.41E-03	-0.010	9.90E-08	5.73E-09
	20	1	chr20:8352693-8352693	0.001	rs6086425	G	A	0.016	2.16E-02	-0.011	3.30E-07	1.10E-08

POS, position of shared genomic regions; A1, effect allele of index SNP; A2, non-effect allele of index SNP; CPASSOC, cross-phenotype association analysis using R package "CPASSOC"; T2DM, type 2 diabetes mellitus; BMI, body mass index.

enrichment analysis on the four shared gene sets of total coffee intake/choice for different coffee types (decaffeinated coffee, ground coffee, instant coffee) and T2DM or related phenotypes. In the KEGG analysis, in addition to total coffee intake, the remaining three gene sets all showed significant enrichment in allograft rejection (hsa05330), autoimmune thyroid disease (hsa05320), Human T-cell leukemia virus 1 infection (hsa05166), and viral myocarditis (hsa05416).

In GO analysis, all four gene sets were significantly enriched in the following five biological processes related to antigen processing: antigen processing and presentation (GO:0019882), antigen processing and presentation of endogenous peptide antigen (GO:0002483), antigen processing and presentation of endogenous peptide antigen *via* MHC class I (GO:0019885), antigen processing and presentation of exogenous peptide antigen (GO:0002478), antigen processing and presentation of peptide antigen (GO: 0048002).

3.4 MR

We used eight different methods to estimate the genetically predicated causal effects of coffee intake and coffee type choice on T2DM or related phenotypes. The results are given in **Supplementary Tables 21–25** and shown graphically in **Figure 2**.

For total coffee intake, its harmful associations with T2DM (b_{CAUSE}: 0.31, p-value $_{CAUSE}$: 3.30×10^{-2}) and BMI (b_{CAUSE}: 0.36, p-value $_{CAUSE}$: 1.8×10^{-5}) were found by seven methods (except MR-Egger regression). Even after adjusting BMI, its enhancement of T2DM risk was also reflected in the results of six models (except CAUSE and MR-Egger regression).

Analysis of different coffee types indicated various results. Compared with other types of coffee, choice for decaffeinated coffee was found to increase the risk of T2DM (CAUSE; b_{CAUSE}: 0.29, p-value $_{CAUSE}$: 2.30×10^{-2}) and BMI (IVW, simple median, weighted median, penalized weighted median) by some methods. In the research of instant coffee, the results obtained by different methods were not completely consistent. CAUSE's result supported the view that choice for instant coffee could increase the risk of T2DM (b_{CAUSE}: 0.17, p-value $_{CAUSE}$: 2.90×10^{-2}) and this causality was mediated by BMI, but weighted median, penalized weighted median, simple mode, and weighted mode showed the opposite result. Different from the above two types of coffee, choice for ground coffee showed a protective association with glucose homeostasis. It was indicated to

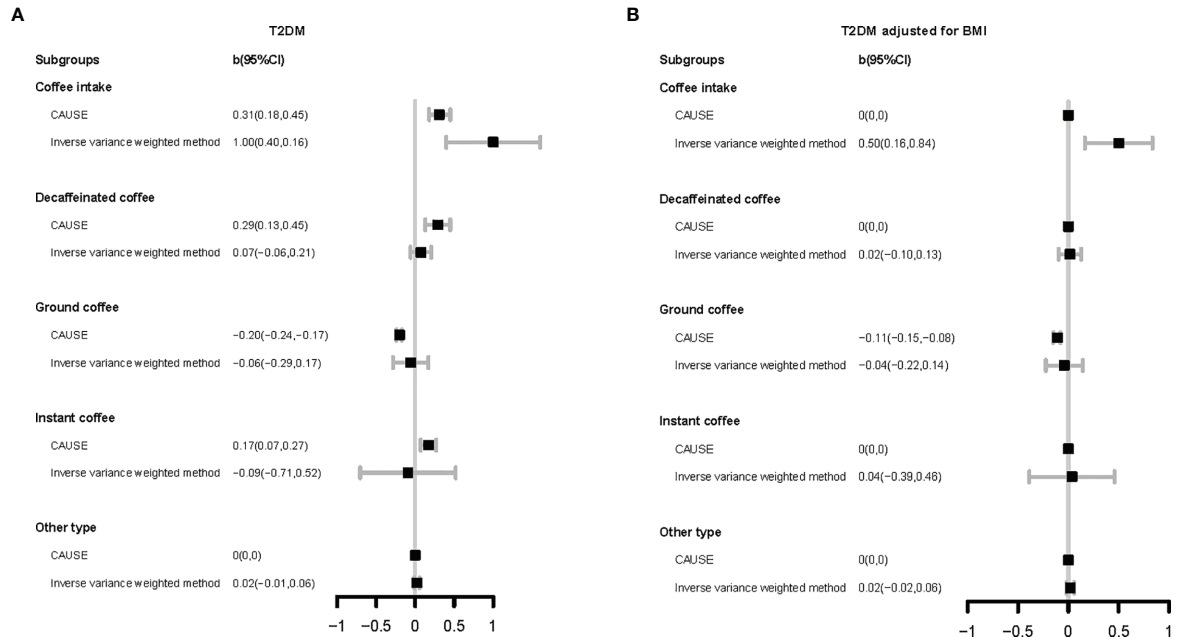


FIGURE 2 | Forest plot for Mendelian randomization results. The vertical axis represents different coffee types and different method. The horizontal axis represents the size of causal effect estimation. T2DM, type 2 diabetes mellitus; BMI, body mass index. CAUSE, causal analysis using summary effect estimates (a mendelian randomization method); CI, confidence interval.

decrease BMI (b $_{CAUSE}$: -0.08, p-value $_{CAUSE}$: 6.50×10^{-5}) by CAUSE. Regardless of whether BMI was adjusted or not, the results of most methods showed that choice for ground coffee can reduce the risk of T2DM (except IVW and MR-Egger regression; T2DM: b $_{CAUSE}$: -0.2, p-value $_{CAUSE}$: 4.70×10^{-10}; T2DM adjusted for BMI: b $_{CAUSE}$: -0.11, p-value $_{CAUSE}$: 4.60×10^{-5}).

4 DISCUSSION

In the present study, we found that genetically proxied total coffee intake, choice for decaffeinated or instant coffee were significantly associated with increased T2DM risk, whereas genetically proxied choice for ground coffee was associated with decreased risk. We have identified the genes shared by these trait pairs and further determined the biological processes and pathways that these genes were enriched in, and the results indicated the association between coffee intake and T2DM was likely mediated by immune system. This is of great significance for a better understanding of the impact of coffee on T2DM.

In the genetic association analysis, our results showed that those opting for ground coffee (include espresso, filter etc) were less genetically susceptible to T2DM, while those who tended to choose instant coffee or decaffeinated coffee (any type) were more genetically susceptible to these risks (**Figure 1**). Because of the heterogeneity among the consumers of different coffee types showed above, we suggested that combining the data of different coffee beverage consumers in statistical analyses should be done with caution.

We have identified many shared genomic regions between coffee intake or choice for different coffee types and T2DM or related phenotypes (**Supplementary Tables 7–20**). Some of these regions have been discovered in GWAS of single trait, but a considerable part of them has never been discovered till now (**Tables 1A–D**). These regions should be paid more attention because they suggest common unknown pathways which may affect both phenotypes and these regions may partly account for the "missing heritability" (35) in mono-phenotype GWAS.

The results of GO enrichment analysis showed that shared genes were significantly enriched in the biological processes related to antigen processing (**Supplementary Figures 2, 4, 6, 8**). This result suggests that the association between coffee intake and T2DM is likely to be related to immune regulation. The KEGG enrichment results also supported this inference (**Supplementary Figures 3, 5, 7**). Moreover, in the cross-phenotype association study, the notable identified SNP rs10492872 mapped to *FTO* is shared by two trait pairs of ground coffee/BMI and instant coffee/BMI, and its effects on ground coffee and instant coffee are in opposite directions, so as rs11660753 mapped to *SMIM21* (**Tables 1C, D**). *FTO* gene is the first candidate gene of obesity, and it is also related to a variety of other phenotypes, such as alcohol intake (36), C-reactive protein levels (37), etc. A recent study has found that *FTO* can help cancer cells escape immune surveillance (38). These findings suggest that *FTO* is likely to participate in immune activities. Compared with *FTO*, there is less research on *SMIM21*, but epidemiological studies have also found its significant association

with rheumatoid arthritis (39). The above research results suggest that the relationship between coffee intake and T2DM and the differences among various coffee type drinkers are likely to be related to the immune system.

In our MR analysis, we used eight different methods for MR research, with their own characteristics (**Figure 2**, **Supplementary Tables 21–25**). Heterogeneity test showed that heterogeneity was common (**Supplementary Tables 26–30**), which means that the IVs selected according to the conventional principles may be invalid or at least partially invalid. Under this circumstance, the results obtained by IVW may be problematic (40). The test of the intercept term of MR-Egger regression (33) showed that the horizontal pleiotropy did not exist or had reached equilibrium (**Supplementary Table 31**). However, it should be noted that MR-Egger regression can only test the unrelated pleiotropy, that is, the horizontal pleiotropy of IVs on the outcome is not related to confounding factors, but cannot be used to identify related pleiotropy. Our genetic correlation results showed that these phenotypes had extensive genetic correlations (**Figure 1**), which suggested that related pleiotropy was likely to exist. The median/mode method (41) has lower requirements for IVs, but considering that it uses little information, the results need to be carefully understood. As a newly proposed method, the CAUSE method (34) can deal with related and unrelated pleiotropy to identify causal models and shared models, but the validity of its results needs to be tested in more applications. In short, we recommend that MR results should be obtained from a combination of different methods such as we did here to address relationships between lifestyle factors and health optimally.

We have noticed that there have been MR studies on coffee intake and the risk of T2DM published (12, 14, 42). Among them, the results of the MR study by Shuai Yuan et al. (12) was similar to our results, they found that coffee intake could increase the risk of T2DM, but this effect disappeared after adjusting for BMI. Whereas the remaining two studies (14, 42) found no significant effect of coffee intake on the risk of T2DM. Compared with these researches, our study has several outstanding advantages. First, in addition to total coffee intake, we also analyzed the association between the choice for different coffee types and T2DM, and found significant differences between various coffee type drinkers. This is not considered by the previous MR studies on coffee intake and T2DM. Second, instead of using only a few variants related to caffeine metabolism, we used more variants as IVs to achieve a greater proportion of explanation for the variance of coffee intake. Finally, we used multiple MR methods with different advantages to overcome the limitations of a single method. Thus, we believe that our research is more reliable and meaningful.

Both the results of genetic correlation study and MR analysis showed that choice for ground coffee was associated with lower risk of T2DM, compared to instant coffee and decaffeinated coffee. This indicates that the content of coffee's beneficial cardiovascular components is different in various coffee types. This is consistent with the results of the laboratory investigation. A study investigating the content of CGAs and caffeine in 83 commercially available coffee species found that unblended ground coffee had the highest CGAs content and lowest mean caffeine/CGAs ratio (16). Study on total phenol content has reached similar conclusions (4). In instant coffee, the lower beneficial effects of chlorogenic acid and total phenols are likely to be offset by the harmful effects of added sugar, creamer, and other ingredients, as the results of observational studies on instant coffee showed (43).

When we directly analyzed total coffee intake without distinguishing between types, we found significant positive associations. This may be because in the UKB participants, compared with ground coffee, people who drink instant coffee or decaffeinated coffee account for a larger proportion. Actual data support this inference (decaffeinated coffee: N=64,717, instant coffee: N=185,482; ground coffee: N=73,906; other type of coffee: N=5,566). Although we are very clear about the cardiovascular beneficial effects of certain components in coffee, previous studies on the association between coffee intake and cardiovascular risks have not reached a consistent result. Some studies thought that moderate coffee drinking can benefit cardiovascular health (44), but there are also research suggesting the harmful effects of coffee (12, 13). Our results suggest a possible reason for this phenomenon. We suggest that research on the health effects of coffee should distinguish between coffee types.

There are several notable strengths of our work. First, this is the first article to simultaneously study the genetic and potential causal relationship between coffee intake and T2DM. We identified many shared genomic regions and found significant causal relationship, which provided new insights into the mechanism of their associations. In addition to the amount of total coffee intake, we also studied the relationship between choice for different types of coffee and T2DM. The differences in the results suggest the importance of distinguishing different types of coffee in coffee research. In addition to T2DM, we have also investigated other T2DM-related phenotypes (BMI, FG, FI, HOMA-IR, HOMA-β), which will facilitate a more thorough understanding of the association between coffee and T2DM. The GWAS summary statistics selected for our study are all from large-scale and high-quality GWAS. We identify many shared loci and found many significant causal relationships that have not been found in previous less-powerful MR studies.

Our research still has some areas that can be further improved. First, for different types of coffee, it should be noted that we are studying the relationship between choosing this type of coffee and T2DM, but not the amount of this type of coffee consumed. In the future, if relevant data is available, the association between the consumption amount of each type of coffee and T2DM can be further studied. Second, although it is observed that there is a great difference between the results of instant coffee and ground coffee, our research results still cannot answer whether this difference is due to the composition difference caused by the processing or the usual addition of saccharin and other additives in instant coffee. Later, we will

further analyze whether the preference of adding milk/cream/sugar to coffee will affect its association with T2DM. We believe that these sites will be much useful for studying the pathogenesis of T2DM and its relationship with coffee.

In summary, we have shown that choice for ground coffee (include espresso, filter etc) has extensive significant negative genetic correlation with T2DM and related phenotypes. MR analysis using all variants indicated genetically proxied choice for ground coffee can decrease BMI and the risk of T2DM, while other types of coffee may increase the risk of T2DM. This study provides new insights and evidence for the health effect of coffee. The results of different coffee types suggests that research on coffee's health effect should pay more attention to distinguishing between coffee types.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XW, JJ, and TH designed the study. XW performed the statistical analysis. XW wrote the manuscript. All authors helped interpret the data, reviewed, and edited the final paper, and approved the submission.

FUNDING

The study was supported by grants from the Peking University Start-up Grant (71013Y2114), High-performance Computing Platform of Peking University and Beijing Technology and Business University Grant(88442Y0033). The funding organization had no role in the preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.818831/full#supplementary-material

REFERENCES

1. Kilpatrick J. *Beanie Coffee. 2020 December 9* (2020). Available at: https://www.beaniecoffee.com/top-coffee-consuming-countries-world/.
2. Williamson G. Protection Against Developing Type 2 Diabetes by Coffee Consumption: Assessment of the Role of Chlorogenic Acid and Metabolites on Glycaemic Responses. *Food Funct* (2020) 11(6):4826–33. doi: 10.1039/d0fo01168a
3. Chen C, Wu F. The Need to Revisit Ochratoxin A Risk in Light of Diabetes, Obesity, and Chronic Kidney Disease Prevalence. *Food Chem Toxicol* (2017) 103:79–85. doi: 10.1016/j.fct.2017.03.001
4. Olechno E, Puścion-Jakubik A, Markiewicz-Żukowska R, Socha K. Impact of Brewing Methods on Total Phenolic Content (TPC) in Various Types of Coffee. *Molecules* (2020) 25(22):5274. doi: 10.3390/molecules25225274.
5. Kokubo Y, Iso H, Saito I, Yamagishi K, Yatsuya H, Ishihara J, et al. The Impact of Green Tea and Coffee Consumption on the Reduced Risk of Stroke Incidence in Japanese Population: The Japan Public Health Center-Based Study Cohort. *Stroke* (2013) 44(5):1369–74. doi: 10.1161/strokeaha.111.677500
6. Choi Y, Chang Y, Ryu S, Cho J, Rampal S, Zhang Y, et al. Coffee Consumption and Coronary Artery Calcium in Young and Middle-Aged Asymptomatic Adults. *Heart* (2015) 101(9):686–91. doi: 10.1136/heartjnl-2014-306663
7. O'Keefe JH, DiNicolantonio JJ, Lavie CJ. Coffee for Cardioprotection and Longevity. *Prog Cardiovasc Dis* (2018) 61(1):38–42. doi: 10.1016/j.pcad.2018.02.002
8. Benn M, Nordestgaard BG. From Genome-Wide Association Studies to Mendelian Randomization: Novel Opportunities for Understanding Cardiovascular Disease Causality, Pathogenesis, Prevention, and Treatment. *Cardiovasc Res* (2018) 114(9):1192–208. doi: 10.1093/cvr/cvy045
9. Karusheva Y, Kunstein L, Bierwagen A, Nowotny B, Kabisch S, Groener JB, et al. An 8-Week Diet High in Cereal Fiber and Coffee But Free of Red Meat Does Not Improve Beta-Cell Function in Patients With Type 2 Diabetes Mellitus: A Randomized Controlled Trial. *Nutr Metab (Lond)* (2018) 15:90. doi: 10.1186/s12986-018-0324-5
10. Alperet DJ, Rebello SA, Khoo EY, Tay Z, Seah SS, Tai BC, et al. The Effect of Coffee Consumption on Insulin Sensitivity and Other Biological Risk Factors for Type 2 Diabetes: A Randomized Placebo-Controlled Trial. *Am J Clin Nutr* (2020) 111(2):448–58. doi: 10.1093/ajcn/nqz306
11. van Dijk AE, Olthof MR, Meeuse JC, Seebus E, Heine RJ, van Dam RM. Acute Effects of Decaffeinated Coffee and the Major Coffee Components Chlorogenic Acid and Trigonelline on Glucose Tolerance. *Diabetes Care* (2009) 32(6):1023–5. doi: 10.2337/dc09-0207
12. Yuan S, Larsson SC. An Atlas on Risk Factors for Type 2 Diabetes: A Wide-Angled Mendelian Randomisation Study. *Diabetologia* (2020) 63(11):2359–71. doi: 10.1007/s00125-020-05253-x
13. Nicolopoulos K, Mulugeta A, Zhou A, Hyppönen E. Association Between Habitual Coffee Consumption and Multiple Disease Outcomes: A Mendelian Randomisation Phenome-Wide Association Study in the UK Biobank. *Clin Nutr* (2020) 39(11):3467–76. doi: 10.1016/j.clnu.2020.03.009
14. Nordestgaard AT, Thomsen M, Nordestgaard BG. Coffee Intake and Risk of Obesity, Metabolic Syndrome and Type 2 Diabetes: A Mendelian Randomization Study. *Int J Epidemiol* (2015) 44(2):551–65. doi: 10.1093/ije/dyv083
15. Wong THT, Sui Z, Rangan A, Louie JCY. Discrepancy in Socioeconomic Status Does Not Fully Explain the Variation in Diet Quality Between Consumers of Different Coffee Types. *Eur J Nutr* (2018) 57(6):2123–31. doi: 10.1007/s00394-017-1488-x
16. Jeon JS, Kim HT, Jeong IH, Hong SR, Oh MS, Yoon MH, et al. Contents of Chlorogenic Acids and Caffeine in Various Coffee-Related Products. *J Adv Res* (2019) 17:85–94. doi: 10.1016/j.jare.2019.01.002
17. Sirota R, Gorelik S, Harris R, Kohen R, Kanner J. Coffee Polyphenols Protect Human Plasma From Postprandial Carbonyl Modifications. *Mol Nutr Food Res* (2013) 57(5):916–9. doi: 10.1002/mnfr.201200557
18. Restuccia D, Spizzirri UG, Parisi OI, Cirillo G, Picci N. Brewing Effect on Levels of Biogenic Amines in Different Coffee Samples as Determined by LC-UV. *Food Chem* (2015) 175:143–50. doi: 10.1016/j.foodchem.2014.11.134
19. team P-U. *Pan-UKB* (2020). Available at: https://pan.ukbb.broadinstitute.org.
20. Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic Analysis of Quantitative Traits in the Japanese Population Links Cell Types to Complex Human Diseases. *Nat Genet* (2018) 50(3):390–400. doi: 10.1038/s41588-018-0047-6
21. Canela-Xandri O, Rawlik K, Tenesa A. An Atlas of Genetic Associations in UK Biobank. *Nat Genet* (2018) 50(11):1593–9. doi: 10.1038/s41588-018-0248-z
22. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps. *Nat Genet* (2018) 50(11):1505–13. doi: 10.1038/s41588-018-0241-6
23. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-Analysis of Genome-Wide Association Studies for Height and Body Mass Index in ~700000 Individuals of European Ancestry. *Hum Mol Genet* (2018) 27(20):3641–9. doi: 10.1093/hmg/ddy271
24. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New Genetic Loci Implicated in Fasting Glucose Homeostasis and Their

Impact on Type 2 Diabetes Risk. *Nat Genet* (2010) 42(2):105–16. doi: 10.1038/ng.520
25. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An Atlas of Genetic Correlations Across Human Diseases and Traits. *Nat Genet* (2015) 47(11):1236–41. doi: 10.1038/ng.3406
26. Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, et al. Meta-Analysis of Correlated Traits via Summary Statistics From GWASs With an Application in Hypertension. *Am J Hum Genet* (2015) 96(1):21–36. doi: 10.1016/j.ajhg.2014.11.011
27. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *Gigascience* (2015) 4:7. doi: 10.1186/s13742-015-0047-8
28. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of Published Genome-Wide Association Studies, Targeted Arrays and Summary Statistics 2019. *Nucleic Acids Res* (2019) 47(D1):D1005–d1012. doi: 10.1093/nar/gky1120
29. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: A Database of Human Genotype-Phenotype Associations. *Bioinformatics* (2016) 32(20):3207–9. doi: 10.1093/bioinformatics/btw373
30. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: An Expanded Tool for Searching Human Genotype-Phenotype Associations. *Bioinformatics* (2019) 35(22):4851–3. doi: 10.1093/bioinformatics/btz469
31. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *Omics* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
32. Pierce BL, Burgess S. Efficient Design for Mendelian Randomization Studies: Subsample and 2-Sample Instrumental Variable Estimators. *Am J Epidemiol* (2013) 178(7):1177–84. doi: 10.1093/aje/kwt084
33. Bowden J, Davey Smith G, Burgess S. Mendelian Randomization With Invalid Instruments: Effect Estimation and Bias Detection Through Egger Regression. *Int J Epidemiol* (2015) 44(2):512–25. doi: 10.1093/ije/dyv080
34. Morrison J, Knoblauch N, Marcus JH, Stephens M, He X. Mendelian Randomization Accounting for Correlated and Uncorrelated Pleiotropic Effects Using Genome-Wide Summary Statistics. *Nat Genet* (2020) 52 (7):740–7. doi: 10.1038/s41588-020-0631-4
35. Génin E. Missing Heritability of Complex Diseases: Case Solved? *Hum Genet* (2020) 139(1):103–13. doi: 10.1007/s00439-019-02034-
36. Hubacek JA, Pikhart H, Peasey A, Malyutina S, Pajak A, Tamosiunas A, et al. The Association Between the FTO Gene Variant and Alcohol Consumption and Binge and Problem Drinking in Different Gene-Environment Background: The HAPIEE Study. *Gene* (2019) 707:30–5. doi: 10.1016/j.gene.2019.05.002
37. Mitra SR, Tan PY, Amini F. Effect of FTO Rs9930506 on Obesity and Interaction of the Gene Variants With Dietary Protein and Vitamin E on C-Reactive Protein Levels in Multi-Ethnic Malaysian Adults. *J Hum Nutr Diet* (2018) 31(6):758–72. doi: 10.1111/jhn.12593
38. Liu Y, Liang G, Xu H, Dong W, Dong Z, Qiu Z, et al. Tumors Exploit FTO-Mediated Regulation of Glycolytic Metabolism to Evade Immune Surveillance. *Cell Metab* (2021) 33(6):1221–33.e11. doi: 10.1016/j.cmet.2021.04.001
39. Bossini-Castillo L, de Kovel C, Kallberg H, van 't Slot R, Italiaander A, Coenen M, et al. A Genome-Wide Association Study of Rheumatoid Arthritis Without Antibodies Against Citrullinated Peptides. *Ann Rheum Dis* (2015) 74(3):e15. doi: 10.1136/annrheumdis-2013-204591
40. Burgess S, Dudbridge F, Thompson SG. Combining Information on Multiple Instrumental Variables in Mendelian Randomization: Comparison of Allele Score and Summarized Data Methods. *Stat Med* (2016) 35(11):1880–906. doi: 10.1002/sim.6835
41. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization With Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* (2016) 40(4):304–14. doi: 10.1002/gepi.21965
42. Said MA, van de Vegte YJ, Verweij N, van der Harst P. Associations of Observational and Genetically Determined Caffeine Intake With Coronary Artery Disease and Diabetes Mellitus. *J Am Heart Assoc* (2020) 9(24):e016808. doi: 10.1161/JAHA.120.016808
43. Kim HJ, Cho S, Jacobs DRJr., Park K. Instant Coffee Consumption may be Associated With Higher Risk of Metabolic Syndrome in Korean Adults. *Diabetes Res Clin Pract* (2014) 106(1):145–53. doi: 10.1016/j.diabres.2014.07.007
44. Zhou A, Hyppönen E. Long-Term Coffee Consumption, Caffeine Metabolism Genetics, and Risk of Cardiovascular Disease: A Prospective Analysis of Up to 347,077 Individuals and 8368 Cases. *Am J Clin Nutr* (2019) 109(3):509–16. doi: 10.1093/ajcn/nqy297

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 18 February 2022
doi: 10.3389/fendo.2022.809445

Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic Etiology and Causal Relationship

Yanna Chi[1], *Xinpei Wang*[1], *Jinzhu Jia*[1,2*] *and Tao Huang*[3,4,5*]

[1] Department of Biostatistics, School of Public Health, Peking University, Beijing, China, [2] Center for Statistical Science, Peking University, Beijing, China, [3] Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, China, [4] Department of Global Health, School of Public Health, Peking University, Beijing, China, [5] Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing, China

OPEN ACCESS

Edited by:
Kezhong Zhang,
Wayne State University, United States

Reviewed by:
Qinghua Sun,
Zhejiang Chinese Medical University, China
Shuai Yuan,
Karolinska Institutet (KI), Sweden
Wuxiang Xie,
Peking University, China

***Correspondence:**
Jinzhu Jia
jzjia@pku.edu.cn
Tao Huang
huang.tao@pku.edu.cn

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: 05 November 2021
Accepted: 18 January 2022
Published: 18 February 2022

Citation:
Chi Y, Wang X, Jia J and Huang T (2022) Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic Etiology and Causal Relationship. Front. Endocrinol. 13:809445. doi: 10.3389/fendo.2022.809445

Objective: This study aimed to explore shared genetic etiology and the causality between smoking status and type 2 diabetes (T2D), cardiovascular diseases (CVDs), and related metabolic traits.

Methods: Using summary statistics from publicly available genome-wide association studies (GWASs), we estimated genetic correlations between smoking status and T2D, 6 major CVDs, and 8 related metabolic traits with linkage disequilibrium score regression (LDSC) analysis; identified shared genetic loci with large-scale genome-wide cross-trait meta-analysis; explored potential shared biological mechanisms with a series of post-GWAS analyses; and determined causality with Mendelian randomization (MR).

Results: We found significant positive genetic associations with smoking status for T2D (Rg = 0.170, $p = 9.39 \times 10^{-22}$), coronary artery disease (CAD) (Rg = 0.234, $p = 1.96 \times 10^{-27}$), myocardial infarction (MI) (Rg = 0.226, $p = 1.08 \times 10^{-17}$), and heart failure (HF) (Rg = 0.276, $p = 8.43 \times 10^{-20}$). Cross-trait meta-analysis and transcriptome-wide association analysis of smoking status identified 210 loci (32 novel loci) and 354 gene–tissue pairs jointly associated with T2D, 63 loci (12 novel loci) and 37 gene–tissue pairs with CAD, 38 loci (6 novel loci) and 17 gene–tissue pairs with MI, and 28 loci (3 novel loci) and one gene–tissue pair with HF. The shared loci were enriched in the exo-/endocrine, cardiovascular, nervous, digestive, and genital systems. Furthermore, we observed that smoking status was causally related to a higher risk of T2D ($\beta = 0.385$, $p = 3.31 \times 10^{-3}$), CAD ($\beta = 0.670$, $p = 7.86 \times 10^{-11}$), MI ($\beta = 0.725$, $p = 2.32 \times 10^{-9}$), and HF ($\beta = 0.520$, $p = 1.53 \times 10^{-6}$).

Conclusions: Our findings provide strong evidence on shared genetic etiology and causal associations between smoking status and T2D, CAD, MI, and HF, underscoring the potential shared biological mechanisms underlying the link between smoking and T2D and CVDs. This work opens up a new way of more effective and timely prevention of smoking-related T2D and CVDs.

Keywords: smoking status, type 2 diabetes, cardiovascular disease, shared genetic etiology, causality

INTRODUCTION

Despite concerted efforts to combat the global tobacco epidemic, tobacco smoking remains the leading preventable cause of morbidity and mortality (1). Smoking has multiple well-known adverse health effects (2, 3), and its association with type 2 diabetes (T2D) and cardiovascular diseases (CVDs) has been a major public health concern. Considerable studies, both prospective cohort studies among different population groups (4–6) and meta-analyses (7–10), have provided compelling evidence of the important role of smoking in increasing the risk of T2D and CVDs. Approximately 30%~40% of the increased risk of T2D (2) and 20%~30% of all CVD deaths (11, 12) compared to never smokers are attributed to smoking. In addition, previous twin or family studies have shown that smoking, T2D, and many CVDs, such as coronary artery disease (CAD), are heritable traits (13–15), and the heritability was estimated to range from 4% to 19% for smoking phenotypes (16, 17), 17%~23% for T2D (18), and 14% ~21% for CAD (19, 20) in recent large-scale genome-wide association studies (GWASs). Furthermore, genetic correlations between several smoking phenotypes and T2D or CVDs have been observed (16, 21). For example, two recent large-scale GWASs on tobacco use revealed that smoking initiation was genetically positively correlated with T2D, CAD, myocardial infarction (MI), and heart failure (HF) and that cigarettes per day and smoking cessation were genetically positively correlated with CAD. More interestingly, single-nucleotide polymorphisms (SNPs) in some genes have been reported to have effects on both smoking and T2D or CVDs (22–24).

These lines of evidence suggest two possibilities to account for such associations between smoking and T2D or CVDs. One is pleiotropy. Smoking and T2D or CVDs may share common genetic variants that simultaneously influence two or more of these traits or disorders by engaging in common pathways or controlling common risk factors. An alternative possibility is that causal associations may exist between smoking and T2D or CVDs. In recent years, large publicly available GWAS datasets and multiple state-of-the-art statistical analysis methods including linkage disequilibrium score regression (LDSC) (25), cross-trait meta-analysis (26), transcriptome-wide association studies (TWAS) (27), and Mendelian randomization (MR) analysis (28–31), can be utilized to facilitate investigations of whether the comorbidity and risk interrelationship of these traits or disorders can be explained by common genetic variants or causality. Given these possibilities and methodological advances, it is now important and feasible for us to elucidate the mechanisms underlying the comorbidity between smoking and T2D and CVDs. As is apparent from the literature, the associations between smoking and T2D or CVDs varied due to the differences in the measurement of smoking in different studies (3–5, 8). In our study, we chose smoking status, an ordinal categorical variable, which is divided into current smokers, former smokers, and never smokers according to smoking intensity and recency.

To our knowledge, no genetic study has systematically explored the common genetic etiology between smoking status and T2D and CVDs. Therefore, in the present study, we conducted a comprehensive analysis using summary statistics from publicly available GWASs to explore shared genetic etiology and the causality between smoking status and T2D, CVDs, and related metabolic traits.

MATERIALS AND METHODS

Study Design and Data Summary

The whole study design is shown in **Figure 1**. Summary statistics used in this study were extracted from publicly available GWASs. The dataset of smoking status was from Gene ATLAS, consisting of 452,264 participants (32, 33). We retrieved summary statistics from the Diabetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium for T2D (N = 898,130) (18). Generally, CVDs encompass a broad range of disorders of the heart and blood vessels including coronary heart disease, cerebrovascular disease, and other conditions. In this study, we chose six common or devastating CVDs including CAD (N = 148,715) (20) and MI (N = 163,665) (34) from the Coronary Artery Disease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus the Coronary Artery Disease (C4D) Genetics (CARDIoGRAMplusC4D) consortium, HF (N = 977,323) (35) from the Heart Failure Molecular Epidemiology for Therapeutic Targets (HERMES), ischemic stroke (IS; N = 521,612) from the METASTROKE collaboration (36), intracerebral hemorrhage (ICH; N = 3,026) from the International Stroke Genetics Consortium (37), and atrial fibrillation (AF; N = 133,073) from the Atrial Fibrillation Genetics Consortium (38). In addition, several important T2D/CVD-related metabolic traits were considered in this study, including glycemic traits [fasting glucose (FG; N = 46,186), fasting insulin (FI; N = 38,238), and the surrogate estimates of β-cell function (HOMA-β; N = 36,466) and insulin resistance (HOMA-IR; N = 37,037) derived from fasting variables by homeostasis model assessment from the Meta-Analyses of Glucose and Insulin-related traits Consortium (39) and blood lipids [high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC), and triglyceride (TG), N = 188,577] from the Global Lipids Genetics Consortium (40). The majority of the participants were of European ancestry in each GWAS (**Supplementary Table 1**). Detailed disease definition and baseline characteristics for each study were described in previous studies (18, 20, 32–40). For example, smoking status, an ordinal categorical variable based on several questions about smoking intensity and recency, includes the categories of current smokers (those who have smoked 100 cigarettes in their lifetime and currently smoke cigarettes), former smokers (those who have smoked at least 100 cigarettes in their lifetime but had quit smoking at the time of interview), and never smokers (those who have never smoked or who have smoked less than 100 cigarettes in their lifetime) (32, 33). T2D status was defined based on multiple sources of evidence, including a self-reported history of T2D, doctor-diagnosed T2D, antidiabetic treatment, fasting plasma glucose >7.0 mmol/L, or 2-h plasma glucose >11.1 mmol/L (18). In CARDIoGRAMplusC4D, CAD status was defined by an inclusive CAD diagnosis, including MI, percutaneous transluminal coronary angioplasty (PTCA), coronary artery bypass grafting (CABG), chronic ischemic heart

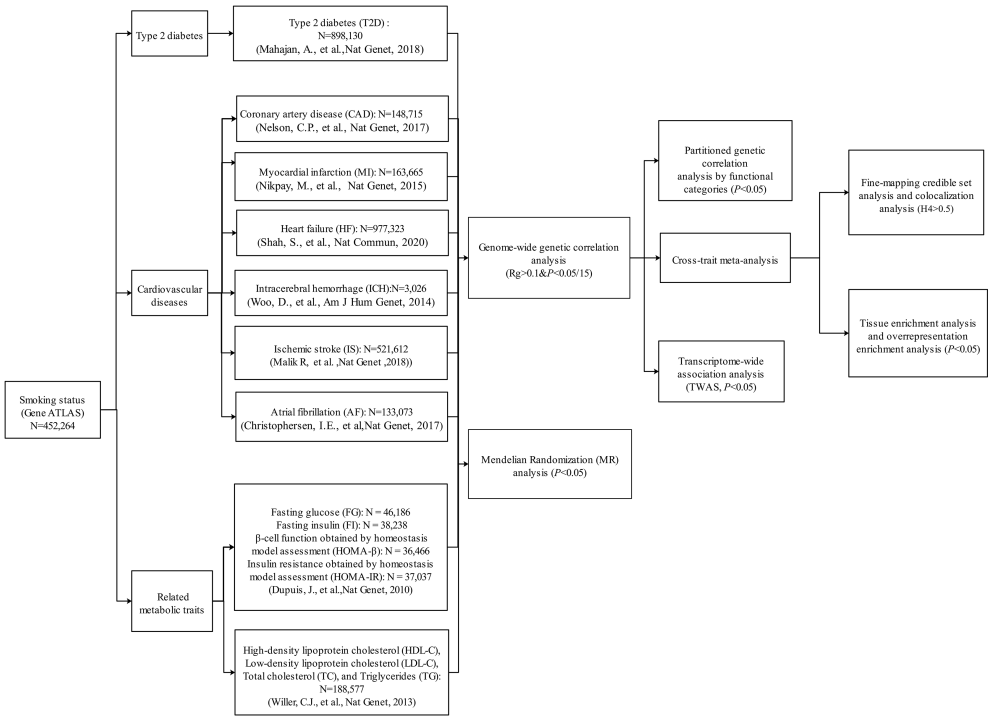


FIGURE 1 | Overall study design. Multiple genome-wide association study (GWAS) data sources were first retrieved. We first conducted a genome-wide genetic correlation analysis between smoking status and type 2 diabetes (T2D), six cardiovascular diseases (CVDs), and eight related metabolic traits. For the traits that showed significant genetic correlation with smoking status, we further conducted post-GWAS analyses to investigate the genetic overlap between them (variant/region/functional levels). Then, we also explored the causal relationship between smoking status and T2D, six CVDs, and eight related metabolic traits.

disease (IHD), and angina (20). More details of these datasets can be seen in the original publications or related websites (18, 20, 32–40). In this study, our analyses were restricted to autosomal chromosomes.

Statistical Analysis

Linkage Disequilibrium Score Regression

We used LDSC, a method requiring only GWAS summary statistics and having no bias by sample overlap, to estimate genetic correlations between smoking status and T2D, 6 major CVDs, and 8 related metabolic traits (41). This method relies on an algorithm that multiplies the Z score of the same SNP and two different phenotypes and then regresses the product of the Z scores from two phenotypes on the LD that the SNP has with all neighboring SNPs (25, 41). The Bonferroni correction was used to adjust multiple testing (two-tailed $p < 0.05/15$).

Partitioned Genetic Correlation

Genetic correlations within functional categories between smoking status and T2D, CAD, MI, and HF were estimated using partitioned LDSC to further describe the genetic overlap at the level of functional categories (42). Eleven functional categories were involved, including the DNase I digital genomic footprinting (DGF) region, DNase I hypersensitivity sites (DHSs), fetal DHS, intron, super-enhancer, transcription factor-binding sites (TFBS), transcribed regions, and histone marks H3K4me1, H3K27ac, H3K4me3, and H3K9ac. This method recalculated LD scores for SNPs partitioned in each particular functional category to estimate the genetic correlation within that functional group.

Cross-Trait Meta-Analysis

We applied a cross-trait GWAS meta-analysis by the R package Cross-Phenotype Association (CPASSOC) to further identify shared loci of the above four trait pairs with strong and significant genetic correlation (26). This method is robust to sample overlap and accommodates different types of phenotypic traits, correlated, independent, continuous, or binary traits. In addition, the effects of trait heterogeneity, population structure, and cryptic relatedness can be controlled by CPASSOC (26). S_{Het} was chosen as the main statistics. SNPs with $P_{SHet} < 5 \times 10^{-8}$ and trait-specific $p < 0.01$ were considered to have effects on both traits.

Fine-Mapping Credible Set Analysis

To identify the regions of shared loci more precisely, fine-mapping credible set analysis based on a Bayesian algorithm was performed to determine credible sets of causal variants at each of the shared loci (43–45). The identified credible sets of causal variants were 99% likely to contain causal disease-associated SNPs by extracting variants that were highly linked ($r^2 > 0.4$) with the index SNP and within 500 kb of the index SNP (46).

Colocalization Analysis

A colocalization analysis by the R package coloc was applied to determine whether the association signals of trait pairs colocalized at

the same locus (47, 48). The probability that both traits are associated and share a single causal variant (Coloc H4 Prob) was calculated with variants extracted within 500 kb of the index SNP at each of the shared loci. Loci with Coloc H4 Prob greater than 0.5 were considered to colocalize (49).

Tissue Enrichment Analysis, Overrepresentation Enrichment Analysis, and Transcriptome–Wide Association Study Analysis

To further understand the biological insights of the identified shared genes between smoking status and T2D, CAD, MI, and HF, we conducted multiple post-GWAS functional analyses. Based on RNA-Seq data from the Human Protein Atlas (HPA) across 35 human tissues (50), we used the TissueEnrich web application to calculate the tissue-specific gene enrichment and further understand whether identified shared genes of each trait pair were enriched in disease-relevant tissues (51). We applied the WebGestalt application (52) to determine overrepresentation enrichment of the identified shared gene set in Gene Ontology (GO) biological processes (53, 54). Furthermore, we conducted TWAS using the FUSION software package and 48 Genotype-Tissue Expression (GTEx) (version 7) reference weights (27) to explore the gene expression association in different tissues between smoking status and T2D, CAD, MI, and HF. The false discovery rate (FDR) Benjamini–Hochberg procedure was applied to correct for multiple testing, and FDR < 0.05 was regarded as significant.

Bidirectional Mendelian Randomization Analysis

Finally, we used the TwoSampleMR package to perform a bidirectional MR analysis to explore the causality between smoking status and T2D, 6 major CVDs, and 8 related metabolic traits (28–31). Bidirectional MR is a form of causal inference analysis that can estimate causal directions and effects by employing genetic instruments selected from large-scale GWASs (55), even in the presence of unmeasured confounders. Three basic assumptions must be fulfilled to yield unbiased causal estimates in the MR analysis: 1) the genetic instruments used must be associated with the exposure, 2) the genetic instruments should be independent of the confounders between the exposure and outcome, and 3) the genetic instruments affect the outcome only through the exposure (46, 56). In this study, we extracted genetic instruments (SNPs) with $p < 5 \times 10^{-8}$ from the GWAS summary statistics of the exposure of interest, conducted the horizontal pleiotropy test, and selected independent genetic instruments at $r^2 < 0.001$ to satisfy these three assumptions. For each potential causality, the inverse variance-weighted (IVW) method was used to obtain the primary causal estimates. The FDR Benjamini–Hochberg procedure was applied to correct for multiple testing (FDR < 0.05).

Notably, the T2D, CAD, and HF GWASs contained UK Biobank participants, which may overlap to some extent with smoking status GWAS from the UK Biobank. Therefore, we additionally extracted T2D, CAD, and HF GWAS summary statistics from earlier or smaller-scale GWASs (57–59) that did not contain UK Biobank participants to further confirm the potential causal associations between smoking status and T2D, CAD, and HF. The details of these GWASs are presented in **Supplementary Table 2**.

RESULTS

Genome-Wide Genetic Correlation

Understanding the genetic correlations of different complex traits or diseases can provide preliminary insights into genetic etiology. Therefore, we firstly estimated genetic correlations between smoking status and T2D, 6 major CVDs, and 8 related metabolic traits by LDSC. Among these traits, T2D (Rg = 0.170, $p = 9.39 \times 10^{-22}$), CAD (Rg = 0.234, $p = 1.96 \times 10^{-27}$), MI (Rg = 0.226, $p = 1.08 \times 10^{-17}$), and HF (Rg = 0.276, $p = 8.43 \times 10^{-20}$) showed strong and significant positive genetic correlations with smoking status (**Table 1**). In addition, we found nominally significant positive genetic correlations with smoking status for IS, ICH, and FG (**Table 1**). Genetic correlations between smoking status and HDL-C or TG reached statistical significance, but the magnitude of genetic correlation was less than 10% (**Table 1**). However, we did not find evidence of genetic correlations with smoking status for AF, FI, HOMA-B, HOMA-IR, LDL-C, and TC (**Table 1**).

TABLE 1 | Genetic correlations between smoking status and T2D, CVDs, and related metabolic traits (α = 0.05/15).

Phenotype 1	Phenotype 2		Rg	Rg_SE	*p*-Value
Smoking status	T2D		0.170	0.018	9.39E–22*
	CVDs	CAD	0.234	0.022	1.96E–27*
		MI	0.226	0.026	1.08E–17*
		HF	0.276	0.030	8.43E–20*
		IS	0.164	0.057	3.70E–03
		ICH	0.188	0.080	1.80E–02
		AF	0.029	0.029	3.17E–01
	Glycemic traits	FG	0.105	0.042	1.31E–02
		FI	0.048	0.055	3.84E–01
		HOMA–β	–0.012	0.052	8.13E–01
		HOMA–IR	0.064	0.058	2.72E–01
	Blood lipids	LDL–C	0.022	0.030	4.77E–01
		HDL–C	–0.094	0.024	6.14E–05
		TC	0.032	0.026	2.11E–01
		TG	0.096	0.026	2.00E–04

Rg, genetic correlation estimate; SE, standard error of genetic correlation estimate; T2D, type 2 diabetes; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure; IS, ischemic stroke; ICH, intracerebral hemorrhage; AF, atrial fibrillation; FG, fasting glucose; FI, fasting insulin; HOMA-β, β-cell function obtained by homeostasis model assessment; HOMA-IR, insulin resistance obtained by homeostasis model assessment; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride.

**A significant p-value after Bonferroni correction.*

Partitioned Genetic Correlation

We used partitioned LDSC analysis to further evaluate genetic correlations between smoking status and T2D, CAD, MI, and HF in 11 functional annotations. Almost all the partitioned genetic correlations in each trait pair were positive (**Figure 2** and **Supplementary Table 3**). Large and statistically significant genetic correlations in many functional categories were observed, and a few categories stood out in particular. The highest magnitude of significant genetic correlation between smoking status and T2D (Rg = 0.167), MI (Rg = 0.164), and HF (Rg = 0.227) was in transcribed regions, where this region can transcribe DNA sequence to mRNA (**Figure 2** and **Supplementary Table 3**). Smoking status and CAD (Rg = 0.162) showed the highest magnitude of significant genetic correlation in DHSs, which are regions of chromatin that are sensitive to cleavage by the DNase I enzyme (**Figure 2** and **Supplementary Table 3**).

Cross-Trait Meta-Analysis

The strong genetic correlations for smoking status and T2D, CAD, MI, and HF encouraged the exploration of common genetic architecture; therefore, we performed a genome-wide cross-trait meta-analysis to identify shared genetic loci between them (meta-analysis $p < 5 \times 10^{-8}$; trait-specific $p < 0.01$). The lists of shared loci of each trait pair are provided in **Tables 2**, **3** and **Supplementary Tables 4–7**.

We found 210 loci significantly associated with both smoking status and T2D, and of these, 32 loci were novel. The most significant locus (index SNP rs9937053, $p_{meta} = 6.72 \times 10^{-81}$) was mapped to *FTO* (**Supplementary Table 4**), the first gene contributing to common forms of human obesity (60). Previous studies have indicated that *FTO* is an essential regulator in the development of obesity-induced metabolic and vascular changes (61) and that adiposity-related risk alleles at *FTO* may predispose individuals to diabetes and cardiovascular events (62, 63). A total

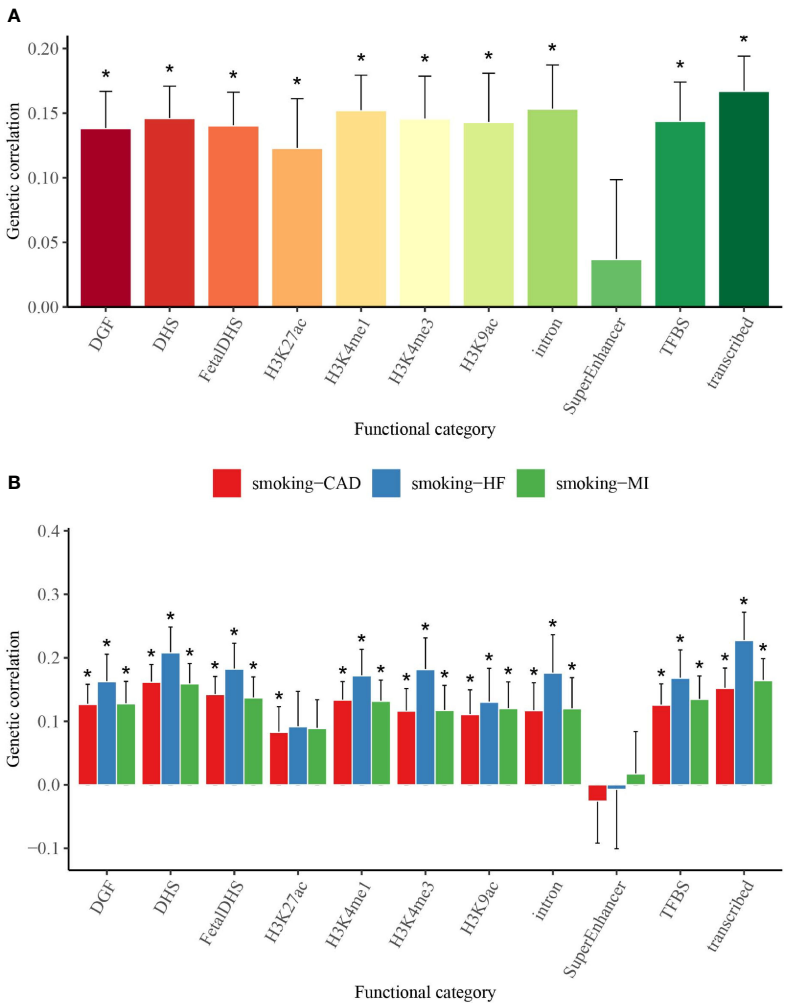


FIGURE 2 | Partitioned genetic correlations of smoking status and T2D, CAD, MI, and HF. **(A)** Partitioned genetic correlations of smoking status and T2D. **(B)** Partitioned genetic correlations of smoking status and CAD, MI, and HF. The vertical axis represents the genetic correlation estimate; the horizontal axis represents 11 functional categories. The asterisk represents significance after Benjamini–Hochberg correction (FDR < 0.05); error bars represent the SE of the genetic correlation estimate. T2D, type 2 diabetes; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure; FDR, false discovery rate.

TABLE 2 | Novel shared loci in the cross-trait meta-analysis of smoking status and T2D ($p_{meta} < 5 \times 10^{-}8$; single trait $p < 0.01$).

SNP	CHR	N	Position	kb	p_{meta}	Variant annotation	Genes within clumping region
rs10093628	8	6	chr8:9393379.9452088	58.71	2.72E–10	Intergenic variant	TNKS
rs7650482	3	4	chr3:12840934.12848822	7.889	5.27E–10	Coding transcript intron variant	CAND2
rs2608280	11	3	chr11:93209472.93264680	55.209	3.60E–09	Downstream gene variant	SMCO4
rs4804414	19	3	chr19:7222655.7223848	1.194	5.98E–09	Coding transcript intron variant	INSR
rs181110840	10	1	chr10:114645185.114645185	0.001	6.00E–09	Intergenic variant	TCF7L2*
rs72682256	14	21	chr14:43069125.43122091	52.967	6.09E–09	Intergenic variant	RP11-90P16.1*
rs8009520	14	20	chr14:103261300.103280657	19.358	6.20E–09	Coding transcript intron variant	TRAF3
rs17412920	22	14	chr22:28628209.28947631	319.423	7.68E–09	Coding transcript intron variant	MIR5739, TTC28
rs7944490	11	20	chr11:17001934.17017622	15.689	8.59E–09	Coding transcript intron variant	PLEKHA7
rs269267	7	1	chr7:140372299.140372299	0.001	9.15E–09	Five prime utr intron variant	DENND2A*
rs7003385	8	4	chr8:41558269.41586773	28.505	1.07E–08	Coding transcript intron variant	ANK1
rs61915371	12	2	chr12:27893972.27896264	2.293	1.14E–08	Coding transcript intron variant	MRPS35
rs62064490	17	12	chr17:9800979.9804724	3.746	1.16E–08	Coding transcript intron variant	RCVRN
rs4841432	8	1	chr8:10583506.10583506	0.001	1.24E–08	Synonymous variant	SOX7
rs2193261	7	2	chr7:117478028.117486934	8.907	1.32E–08	Coding transcript intron variant	CTTNBP2
rs10985975	9	5	chr9:126101008.126123009	22.002	1.44E–08	Intergenic variant	CRB2
rs583887	11	26	chr11:65575263.65663547	88.285	1.52E–08	Upstream gene variant	CCDC85B, CFL1, CTSW, EFEMP2, FIBP, FOSL1, MUS81, SNX32
rs17684514	8	2	chr8:8547642.8574282	26.641	1.59E–08	Intergenic variant	CLDN23
rs1362910	8	2	chr8:30856464.30857668	1.205	2.27E–08	Coding transcript intron variant	PURG
rs12891360	14	3	chr14:104008159.104011429	3.271	2.32E–08	Downstream gene variant	TRMT61A*
rs34954697	2	1	chr2:226918363.226918363	0.001	2.85E–08	Intergenic variant	IRS1*
rs1669801	14	5	chr14:46921092.46936747	15.656	2.98E–08	Intergenic variant	LINC00871
rs2536951	9	1	chr9:126646519.126646519	0.001	3.13E–08	Coding transcript intron variant	DENND1A
rs112583287	6	1	chr6:160919184.160919184	0.001	3.16E–08	Non-coding transcript intron variant	LPAL2
rs4488763	22	1	chr22:32380164.32380164	0.001	3.49E–08	Intergenic variant	YWHAH*
rs117981235	11	1	chr11:9820342.9820342	0.001	3.62E–08	Coding transcript intron variant	SBF2, SBF2-AS1
rs6059938	20	4	chr20:33178324.33187130	8.807	4.18E–08	Coding transcript intron variant	PIGU
rs536445	3	1	chr3:173120103.173120103	0.001	4.30E–08	Five prime utr intron variant	NLGN1
rs117471638	10	1	chr10:93158084.93158084	0.001	4.39E–08	Intergenic variant	LOC100188947
rs1465573	5	1	chr5:157985730.157985730	0.001	4.51E–08	Intergenic variant	EBF1*
rs3735260	7	1	chr7:69064637.69064637	0.001	4.81E–08	Five prime utr exon variant	AUTS2
rs2249850	10	1	chr10:104512006.104512006	0.001	4.87E–08	Coding transcript intron variant	WBP1L

CHR, chromosome; SNP, single-nucleotide polymorphism; T2D, type 2 diabetes.
**The nearest gene to this locus.*

of 63 genome-wide significant loci were identified in the meta-analysis of smoking status and CAD, of which 12 loci were novel (**Supplementary Table 5**). The most significant locus (index SNP rs1412830, $p_{meta} = 3.03 \times 10^{-34}$) was mapped to the *CDKN2B-AS1* region, which was also found to be significant in the cross-trait meta-analysis for smoking status and T2D ($p_{meta} = 2.63 \times 10^{-17}$) or MI ($p_{meta} = 1.45 \times 10^{-23}$) (**Figure 3**). *CDKN2B-AS1* is a significant genetic susceptibility locus for CVDs and has also been linked to several other pathologies, such as several cancers, T2D, periodontitis, Alzheimer's disease,

and glaucoma (64, 65). A sum of 38 loci, including 6 novel loci, were found to be significantly associated with both smoking status and MI (**Supplementary Table 6**). The top two significant loci (index SNP rs12617922, $p_{meta} = 4.36 \times 10^{-25}$; index SNP rs12244388, $p_{meta} = 7.40 \times 10^{-24}$) were located at *RPL6P5* and *AS3MT*. *AS3MT* encodes arsenite methyltransferase and plays a role in arsenic metabolism by catalyzing the transfer of a methyl group from *S*-adenosyl-L-methionine (AdoMet) to trivalent arsenical (66). Cigarette smoke contains arsenic with adverse effects and arsenic exposure has been proven to be linked with the risk of acute MI (67). The genome-wide cross-trait meta-analysis between smoking status and HF identified 28 genome-wide significant loci, of which 3 loci were novel (**Supplementary Table 7**). The strongest signal was observed on chromosome 3 at the *CADM2* region (index SNP rs34495106, $p_{meta} = 3.02 \times 10^{-19}$), a critical gene associated with a range of behavioral and metabolic traits, including physical activity, alcohol and cannabis use, and obesity (68).

Notably, some shared loci overlapped in the cross-trait meta-analysis of smoking status–T2D and smoking status–CVDs (**Figures 3**, **4**). In addition to the SNP rs1412830 located at the *CDKN2B-AS1* region, we observed four overlapping significant loci (index SNPs: rs12453682, rs1381274, rs2867112, and rs4790874) in the genome-wide cross-trait meta-analysis of smoking status–T2D and smoking status–CAD. Of these, the SNP rs2867112 is near the protein-coding gene body *TMEM18*, and genetic variants in the proximity of the gene have been linked to obesity (69), insulin levels, and blood glucose levels (70). In addition, two loci (index SNPs: rs72712556 and rs10030552) mapped to *MAML3* were found to be genome-wide significant in the meta-analysis of smoking status–T2D and smoking status–HF. These two loci reached genome-wide significance in the single-trait GWAS of smoking status, but their association with T2D or HF remains unknown. More importantly, genes *AS3MT* and *SMG6* were identified in the cross-trait meta-analysis of all four trait pairs (smoking status–T2D, smoking status–CAD, smoking status–MI, and smoking status–HF). Gene *AS3MT* is known to act in arsenic metabolism (66), and polymorphisms in the *AS3MT* have been reported to be associated with CVDs (71) and T2D risks (72, 73). *SMG6* is ubiquitously expressed in many tissues and cell types and has dual functions in telomere maintenance and RNA surveillance pathways (74). Multiple loci in *SMG6* have been proven to be associated with smoking behavior (17) and CAD (75, 76). However, its role in T2D remains to explore.

Fine-Mapping Credible Set Analysis and Colocalization Analysis

Based on Bayesian fine-mapping, we identified the 99% credible set of causal variants at each of the shared loci. The lists of credible sets of causal variants for each shared locus are provided in **Supplementary Tables 8–11**. In addition, a colocalization analysis was applied to determine whether the two traits were associated and shared the same causal variant at each shared locus. The number of the shared loci considered to colocalize in each trait pair was 20 (smoking status–T2D), 7 (smoking status–CAD), 4 (smoking status–MI), and 4 (smoking status–HF) (**Supplementary Tables 12–15**). Among these, 3 loci (index SNPs: rs329122, rs3742305, and rs1443750) reached a great probability (>95%) of having shared causal variants of smoking status and T2D, in addition to 2 loci (index SNPs: rs11556924 and rs10774625) for smoking status–CAD, 2 loci (index SNPs: rs11556924 and rs653178) for smoking status–MI, and one locus (index SNP: rs4766578) for smoking status–HF.

TABLE 3 | Novel shared loci in the cross-trait meta-analysis of smoking status and CAD, MI, and HF ($p_{meta} < 5 \times 10^{-8}$; single trait $p < 0.01$).

Phenotype	SNP	CHR	N	Position	kb	p_{meta}	Variant annotation	Genes within clumping region
CAD	rs715694	15	2	chr15:47488977.47489021	0.045	5.07E−09	Five prime utr intron variant	SEMA6D
	rs7868608	9	1	chr9:128746044.128746044	0.001	6.16E−09	Intergenic variant	PBX3*
	rs1603985	3	1	chr3:25148868.25148868	0.001	1.25E−08	Intergenic variant	RARB*
	rs530324	8	1	chr8:27491186.27491186	0.001	1.29E−08	Upstream gene variant	SCARA3*
	rs62263602	3	3	chr3:49991060.50152491	161.432	1.59E−08	Coding transcript intron variant	RBM5, BM5-AS1, BM6
	rs10818125	9	12	chr9:120986288.121008326	22.039	2.29E−08	Intergenic variant	TUBB4BP6*
	rs56399143	4	1	chr4:147630649.147630649	0.001	2.59E−08	Coding transcript intron variant	TTC29
	rs7546040	1	13	chr1:44202991.44247233	44.243	2.77E−08	Coding transcript intron variant	ST3GAL3
	rs6734603	2	1	chr2:182038729.182038729	0.001	2.81E−08	Intergenic variant	ITGA4*
	rs10183073	2	1	chr2:146408408.146408408	0.001	4.07E−08	Intergenic variant	RPL6P5*
	rs2107109	12	1	chr12:113212371.113212371	0.001	4.72E−08	Five prime utr intron variant	RPH3A*
	rs1362727	18	1	chr18:25235351.25235351	0.001	4.84E−08	Intergenic variant	CDH2*
MI	rs62216572	21	2	chr21:46488032.46491155	3.124	5.34E−09	Downstream gene variant	SSR4P1
	rs10490563	2	2	chr2:161914168.161915361	1.194	9.99E−09	Intergenic variant	TANK*
	rs10067365	5	3	chr5:125401016.125432585	31.57	1.47E−08	Intergenic variant	GRAMD3*
	rs530324	8	1	chr8:27491186.27491186	0.001	1.87E−08	Upstream gene variant	SCARA3*
	rs56399143	4	1	chr4:147630649.147630649	0.001	3.17E−08	Coding transcript intron variant	TTC29
	rs288159	5	1	chr5:107364363.107364363	0.001	4.85E−08	Coding transcript intron variant	FBXL17
HF	rs4697140	4	7	chr4:20092322.20114221	21.9	4.03E−08	Intergenic variant	SLIT2*
	rs2680705	17	1	chr17:56495584.56495584	0.001	4.52E−08	Upstream gene variant	RNF43*
	rs6917970	6	3	chr6:129428104.129428850	0.747	4.76E−08	Coding transcript intron variant	LAMA2

CHR, chromosome; SNP, single-nucleotide polymorphism; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure.
**The nearest gene to this locus*

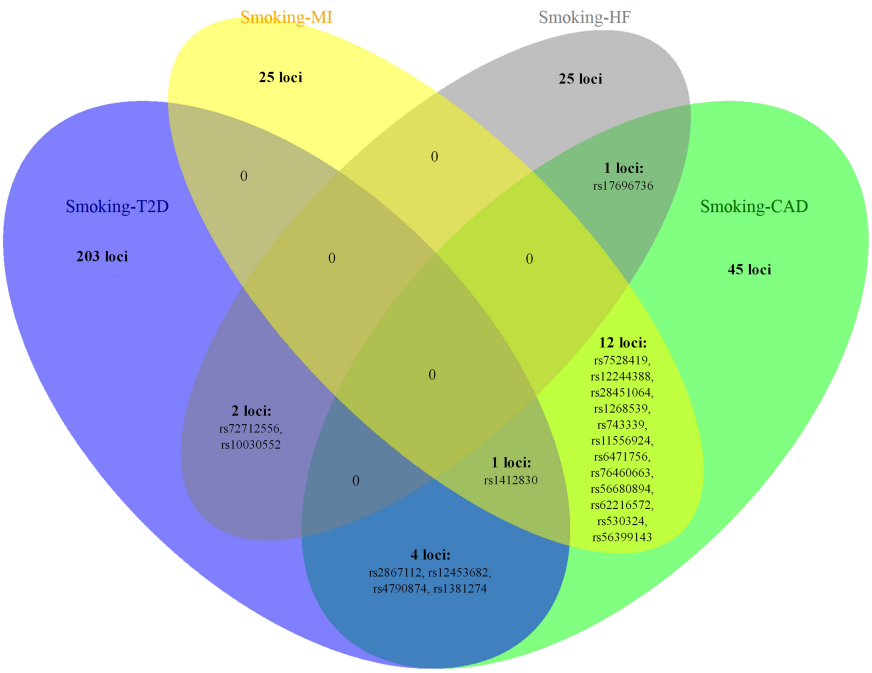


FIGURE 3 | The overlapping loci at the SNP level identified by the cross-trait meta-analysis across different trait pairs. The Venn diagram illustrates the overlapping loci at the SNP level identified by the cross-trait meta-analysis across different trait pairs. T2D, type 2 diabetes; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure; SNP, single-nucleotide polymorphism.

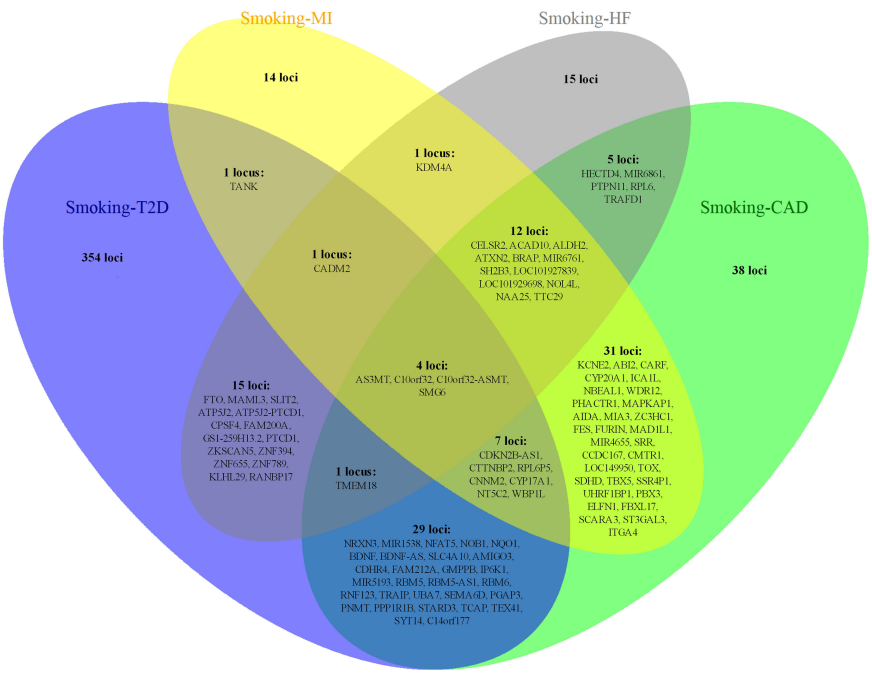


FIGURE 4 | The overlapping loci at the gene level identified by the cross-trait meta-analysis across different trait pairs. The Venn diagram illustrates the overlapping loci at the gene level identified by the cross-trait meta-analysis across different trait pairs. T2D, type 2 diabetes; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure.

Tissue Enrichment Analysis

To determine whether shared genes from cross-trait meta-analysis between smoking status and T2D, CAD, MI, and HF were enriched for expression in the disease-relevant tissues, we performed a tissue enrichment analysis using the TissueEnrich web application and tissue-specific genes from RNA-Seq data of the HPA. We found that the shared genes of smoking status with T2D, CAD, MI, and HF were all most strongly enriched in the adrenal gland (**Figure 5**). The stomach was another strongly enriched tissue for the shared genetic genes of smoking status–CAD and smoking status–MI, in addition to the cerebral cortex for the shared genetic genes of smoking status–HF (**Figure 5**).

Overrepresentation Enrichment Analysis

The overrepresentation enrichment analysis of the GO biological processes highlighted several significantly enriched biological processes for the shared genes between smoking status and T2D, mainly involving regulation of insulin secretion and regulation of peptide hormone secretion (**Supplementary Table 16**). In addition, the shared genes between smoking status and CAD were significantly enriched in the positive regulation of leukocyte adhesion to vascular endothelial cells, axon development, cell morphogenesis involved in neuron differentiation, and neuron projection morphogenesis (**Supplementary Table 17**). However, no significantly enriched biological process for the shared genes of smoking status–MI and smoking status–HF was found.

Transcriptome-Wide Association Analysis

We conducted a TWAS analysis to explore the genes whose expression in different tissues was associated with smoking status, T2D, CAD, MI, and HF, and to determine if these genes were common among these traits. The lists of gene–tissue pairs significantly associated with each trait are shown in **Supplementary Tables 18–22**. Among these gene–tissue pairs, 354 gene–tissue pairs overlapped between smoking status and T2D, in addition to 37 gene–tissue pairs for smoking status–CAD, 17 gene–tissue pairs for smoking status–MI, and one gene–tissue pair for smoking status–HF (**Supplementary Table 23**). Notably, 17 gene–tissue pairs involving four genes (*FAM117B*, *FES*, *ICA1L*, and *NBEAL1*) for smoking status–MI were contained in gene–tissue pairs for smoking status–CAD, most of which were observed in the nervous, cardiovascular, exo-/endocrine, and digestive systems. *C2orf69*–Brain Caudate basal ganglia gene–trait pair was the only one observed overlapping gene–tissue pair between smoking status and HF. Moreover, the enrichment of smoking status and T2D genes expressed were across multiple tissues, not only including nervous, cardiovascular, exo-/endocrine, and digestive systems but also involving the genital system.

Mendelian Randomization Analysis

We performed a bidirectional MR analysis to explore the causal relationship between smoking status and T2D, 6 major CVDs, and 8 related metabolic traits. In the detection of the causal effect of smoking status on cardiometabolic traits, we found that smoking status had significant positive causal effects on T2D ($\beta = 0.385$, $p = 3.31 \times 10^{-3}$), CAD ($\beta = 0.670$, $p = 7.86 \times 10^{-11}$), MI ($\beta = 0.725$, $p = 2.32 \times 10^{-9}$), and HF ($\beta = 0.520$, $p = 1.53 \times 10^{-6}$) (**Table 4**). However, the causal effects of smoking status on other traits (IS, ICH, AF, FG, FI, HOMA-B, HOMA-IR, HDL-C, LDL, TC, and TG) were not identified (**Table 4**). In addition, we did not observe any significant causal effect of cardiometabolic traits on smoking status (**Table 4**). Consistent findings that smoking status had significant positive causal effects on T2D, CAD, and HF were observed using additional GWAS data (**Supplementary Table 24**). MR-Egger regression analysis showed that none of the results were affected by horizontal pleiotropy (**Table 4**). These results corroborated each other and supported the robustness of our primary findings.

DISCUSSION

To our knowledge, this is the first study to systematically explore shared genetic etiology and the causal relationship between smoking status and T2D and CVDs. First, we found strong positive genetic correlations and further identified shared genetic loci between smoking status and T2D, CAD, MI, and HF. Second, we found that the shared genetic loci were mainly enriched in the adrenal gland and stomach tissues and the biological pathways of nervous system development and regulation of peptide hormone secretion. Third, our TWAS further provided evidence that the enrichment of shared genes expressed was across multiple tissues, including exo-/endocrine, cardiovascular, nervous, digestive, and genital systems. Finally, we identified the causal associations of smoking status with T2D, CAD, MI, and HF. In general, exploration of the shared genetic architecture and causality between smoking status and T2D or CVDs furthers the understanding of the biological mechanisms underlying this comorbidity.

The strong genetic correlations consistent with previous studies (21, 77) suggested that the phenotypic correlations between smoking status and T2D, CAD, MI, and HF were due to a common genetic predisposition base, and we further identified 210 shared genetic loci for smoking status–T2D, in addition to 63 loci for smoking status–CAD, 38 loci for smoking status–MI, and 28 loci for smoking status–HF in the genome-wide cross-trait meta-analysis. Among these shared genetic variants, 32 novel loci were found for smoking status–T2D, along with 12 novel loci for smoking status–CAD, 6 novel loci for smoking status–MI, and 3 novel loci for smoking status–HF, demonstrating the great power of cross-trait meta-analysis in identifying specific shared loci. We highlight several overlapping loci or genes in different trait pairs, which may provide more effective genetic targets for the timely prevention, diagnosis, and treatment of smoking-related T2D and CVDs. The only top locus common to the smoking status–T2D, smoking status–CAD, and smoking status–MI meta-analysis was rs1412830 mapped to *CDKN2B-AS1*. *CDKN2B-AS1* gene is an indispensable long non-coding RNA in multiple diseases (65). In addition to T2D and CVDs (64), *CDKN2B-AS1* has been shown to be aberrantly

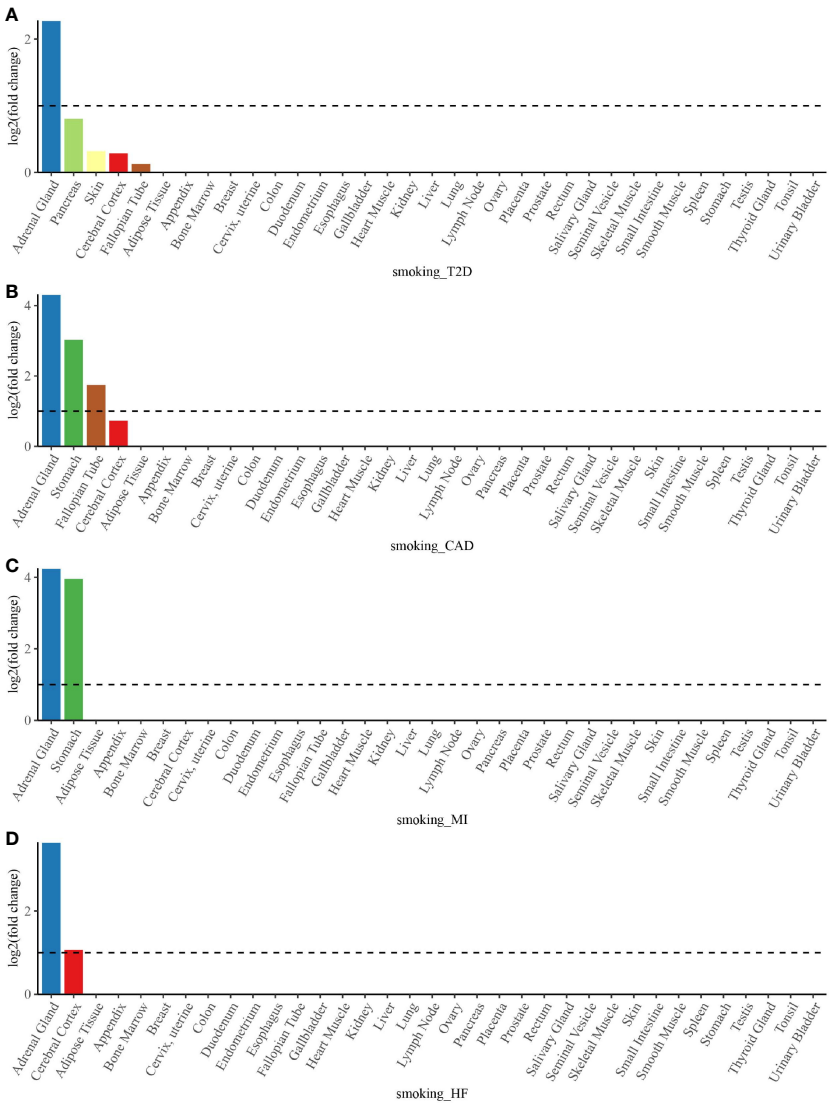


FIGURE 5 | Tissue enrichment analysis for the expression of cross-trait-associated genes between smoking status and T2D **(A)**, CAD **(B)**, MI **(C)**, and HF **(D)**. The vertical axis illustrates the logarithm of tissue expression enrichment fold change based on two. The horizontal axis illustrates 35 independent tissue types. T2D, type 2 diabetes; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure.

expressed in various malignancies, idiopathic pulmonary fibrosis, endometriosis, inflammatory bowel disease, and primary open-angle glaucoma and to participate in the progression of lipids, carbohydrate metabolism, and inflammation regulation (65), which is likely to serve as a promising therapeutic target or prognostic biomarker in multiple human diseases. The SNP rs2867112 near the protein-coding gene body *TMEM18* was found to be significant in the meta-analysis for smoking status–T2D and smoking status–CAD. *TMEM18* is an important susceptibility locus for obesity (69), which is an independent risk factor for the development and progression of T2D and CVDs. A previous study provided evidence that smoking might modify the genetic effects of *TMEM18* on body mass index (BMI), a proxy for adiposity (78). In addition, two loci (index SNPs: rs72712556 and rs10030552) mapped to *MAML3* were found to have genome-wide significance in the meta-analysis of smoking status–T2D and smoking status–HF, which reached genome-wide significance in the single-trait GWAS of smoking status, but its association with HF or T2D remains unknown and deserves in-depth study. *AS3MT* and *SMG6* are two important genes that were identified in the cross-trait meta-analysis of all four trait pairs (smoking status–T2D, smoking status–CAD, smoking status–MI, and smoking status–HF). Cigarette smoke is a vital source of ingested low-level arsenic, and chronic arsenic exposure is associated with increased morbidity and mortality from CVDs (71, 79) and an increased risk of T2D (72, 73). Polymorphisms in *AS3MT* gene are associated with the efficiency

TABLE 4 | Bidirectional MR analysis of smoking status and T2D, CVDs, and related metabolic traits.

Exposure	Outcome	SNPs, n	Inverse variance weighted			MR-Egger		MR-Egger	
			β	*p*-Value	FDR	β	*p*-Value	Intercept	*p*-Value
Smoking status	T2D	127	0.385	3.31E–03	2.48E–02*	–0.022	0.969	0.004	0.470
	CAD	127	0.670	7.86E–11	2.36E–09*	0.195	0.669	0.005	0.286
	MI	127	0.725	2.32E–09	3.48E–08*	0.288	0.596	0.004	0.410
	HF	127	0.520	1.53E–06	1.53E–05*	0.589	0.222	–0.001	0.884
	IS	59	0.573	5.26E–02	1.81E–01	–0.007	0.996	0.005	0.665
	ICH	87	–0.202	7.73E–01	9.41E–01	–2.175	0.469	0.019	0.499
	AF	127	0.018	9.09E–01	9.41E–01	–0.131	0.853	0.001	0.828
	FG	58	0.027	6.65E–01	9.41E–01	–0.240	0.391	0.003	0.328
	FI	58	–0.017	8.02E–01	9.41E–01	0.369	0.224	–0.004	0.193
	HOMA-β	58	0.016	8.07E–01	9.41E–01	0.503	0.082	–0.005	0.084
	HOMA-IR	58	0.014	8.47E–01	9.41E–01	0.328	0.307	–0.003	0.315
	HDL-C	57	–0.145	8.43E–02	2.13E–01	–0.668	0.077	0.005	0.153
	LDL-C	57	0.116	1.95E–01	4.19E–01	0.318	0.430	–0.002	0.606
	TG	57	0.169	5.96E–02	1.81E–01	0.709	0.077	–0.005	0.164
	TC	57	0.158	6.05E–02	1.81E–01	0.203	0.593	0.000	0.903
T2D	Smoking status	202	0.003	2.93E–01	5.17E–01	–0.002	0.727	0.000	0.363
CAD		47	–0.001	9.03E–01	9.41E–01	–0.008	0.522	0.001	0.517
MI		25	0.003	6.40E–01	9.41E–01	–0.024	0.105	0.003	0.048
HF		12	–0.003	8.30E–01	9.41E–01	–0.035	0.546	0.002	0.568
IS[a]		19	0.010	1.11E–01	2.55E–01	–0.029	0.248	0.004	0.110
ICH[a]		13	0.001	7.34E–01	9.41E–01	–0.007	0.457	0.002	0.404
AF		24	0.006	8.52E–02	2.13E–01	0.005	0.571	0.000	0.815
FG		14	0.025	4.97E–02	1.81E–01	0.048	0.108	–0.001	0.361
FI[a]		11	–0.026	2.52E–01	4.72E–01	0.070	0.328	–0.003	0.170
HOMA-β		4	–0.049	2.17E–01	4.34E–01	–0.240	0.222	0.006	0.287
HOMA-IR[a]		13	–0.004	8.38E–01	9.41E–01	0.025	0.735	–0.001	0.678
HDL-C		87	0.000	9.72E–01	9.72E–01	0.004	0.662	0.000	0.619
LDL-C		77	–0.007	5.77E–02	1.81E–01	–0.002	0.636	0.000	0.263
TG		55	–0.001	8.43E–01	9.41E–01	0.004	0.642	0.000	0.463
TC		88	–0.009	1.27E–02	7.61E–02	–0.005	0.415	0.000	0.393

False discovery rate (FDR) Benjamini–Hochberg procedure was used to correct for multiple testing (FDR < 0.05).

MR, Mendelian randomization; SNPs, single-nucleotide polymorphisms; Rg, genetic correlation estimate; SE, standard error of genetic correlation estimate; T2D, type 2 diabetes; CAD, coronary artery disease; MI, myocardial infarction; HF, heart failure; IS, ischemic stroke; ICH, intracerebral hemorrhage; AF, atrial fibrillation; FG, fasting glucose; FI, fasting insulin; HOMA-β, β-cell function obtained by homeostasis model assessment; HOMA-IR, insulin resistance obtained by homeostasis model assessment; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride.

[a]Few SNPs achieved genome-wide significance in the original GWAS; in order to obtain valid and reliable instrumental variables for MR analysis, we set the p-value threshold to 1×10^{-5}.

**A significant p-value after Benjamini-Hochberg correction.*

of arsenic biotransformation (66, 72), suggesting that the mechanisms of arsenic metabolism and biotransformation may play an important role in smoking-related T2D and CVDs. Multiple loci in *SMG6* have been proven to be associated with smoking behavior (17) and CAD (75, 76). Moreover, a previous study has shown that tobacco smoking is associated with the methylation of genes related to CAD, which includes *SMG6* gene (75). These findings provide novel insights into the pathways that link tobacco smoking to the risk of CVDs. However, the role of *SMG6* gene in smoking-related T2D remains to be explored.

In addition to the significant findings in the shared genes related to both smoking and T2D or CVDs, we identified the relevant tissues and biological processes that the shared genes enriched in which suggests the potential biological mechanisms that confer comorbid effects. Tissue enrichment analysis showed that the shared genes of smoking status with T2D, CAD, MI, and HF were all most strongly enriched in the adrenal gland. A previous study has reported that cigarette smoking is a strong activator of the hypothalamus–pituitary–adrenal (HPA) axis followed by significant elevations in the adrenal hormone cortisol (80). Cortisol plays an important role in lipid and glucose metabolism; and elevated cortisol levels, if prolonged, lead to a redistribution of body fat characterized by truncal obesity, which is a risk factor for T2D and CVDs (81). Activation of the HPA axis is also thought to contribute to drug abuse during the addictive process, which may also contribute to the abuse-related effects of cigarette smoking (82). In the overrepresentation enrichment analysis, the biological pathway of insulin secretion was found to be significant for the shared genes of smoking status and T2D, indicating that smoking can affect pancreatic islet cell function. Many studies have found neuronal nicotinic acetylcholine receptors (nAChRs) expressed on pancreatic islet cells (83), and these functional nAChRs sensitive to nicotine in pancreatic cells may be a switch to modulate pancreatic cell physiological function and involved in tobacco toxicity (84). Furthermore, several studies in animal models have shown that nicotine can increase apoptosis of islet β-cells, thus reducing insulin secretion (85–88). Mitochondrial dysfunction, oxidative stress, and inflammation are involved as underlying mechanisms for the direct toxicity induced by

nicotine *via* nAChRs (84). The stomach was another strongly enriched tissue for the shared genetic loci of smoking status–CAD and smoking status–MI. Relevant studies have shown that smoking can increase the probability of getting heartburn and peptic ulcers (89), and gastrointestinal diseases may trigger myocardial ischemia-related chest pain probably through the afferent vagal fibers shared by the esophagus and the heart to induce a coronary spasm (90, 91). In addition, the shared genes for smoking status–HF/CAD were enriched in cerebral cortex tissue and the biological pathways of nervous system development, indicating the important role of the nervous system on the comorbidity of smoking and CVDs. Nicotine and fine particulate matter in tobacco smoke can lead to increased sympathetic nerve activity (92), which is one of the hallmarks of chronic congestive HF (93) and plays a role in the process of atherosclerosis (94).

Our TWAS further provided evidence that the shared genes were mostly from the exo-/endocrine, cardiovascular, nervous, and digestive systems. In addition, the TWAS result reported the enrichment of the shared genes between smoking status and T2D from the genital system. Smoking and T2D have a variety of adverse effects on the genital system (95, 96). More importantly, smoking and diabetes may influence the epigenetic modification during the production of germ cells, and these epigenetic dysregulations may be inherited through the germ line and passed onto more than one generation, which in turn may increase the risk of related diseases in offspring (97). A total of 58 significant genes in TWAS were also found to be genome-wide significant in cross-trait meta-analysis for smoking status–T2D, in addition to 13 genes for smoking status–CAD and 3 genes for smoking status–MI, which further indicated the fact that a significant portion of shared genetic loci we identified in the cross-trait meta-analysis were indeed functional variants of modulating gene expression on influencing both phenotypes. Among these, we highlight the importance of the gene *TCF7L2*, which showed significance in the cross-trait meta-analysis and TWAS of smoking status and T2D. SNPs in *TCF7L2* are especially known to be associated with a higher risk of developing T2D (98). Recently, a study has suggested that *TCF7L2* links nicotine addiction to diabetes in animal models. This study has revealed that *TCF7L2* is densely expressed in the medial habenula and plays an important role in regulating the function of nAChRs in the habenula and in controlling nicotine intake (22). Habenular neurons provide polysynaptic input to the pancreas, and nicotine acts on this habenula–pancreas circuit, in a *TCF7L2*-dependent manner and *via* the autonomic nervous system, to increase blood glucose levels (22). Furthermore, *FES*, *ICA1L*, and *NBEAL1* genes showed significance in the cross-trait meta-analysis and TWAS of smoking status–CAD and smoking status–MI and expressed in multiple tissues, including the brain, nerve, artery, adipose, pancreas, and thyroid tissues. Gene *FES*, which encodes the human cellular counterpart of a feline sarcoma retrovirus protein with transforming capabilities, is well known to be associated with myeloid leukemia (99), but recent studies observed the function of *FES* in modulating atherosclerotic plaque vulnerability (100) and the effect of tobacco smoking on DNA methylation of *FES* (75). Genes *ICA1L* and *NBEAL1* were mapped by the same locus (index SNP: rs114123510), and both are related to cholesterol metabolism, in which dysregulation promotes the pathology of atherosclerosis, MI, and strokes (101). Notably, *C2orf69*–Brain Caudate basal ganglia gene–trait pair was the only one observed overlapping gene–tissue pair between smoking status and HF. *C2orf69* is an evolutionarily conserved gene whose function needs to be further clarified, but recent studies have shown its association with a fatal autoinflammatory syndrome that disrupts the development/homeostasis of the immune and central nervous systems (102, 103), which may contribute to the link between smoking and HF.

In addition to pleiotropy, the associations between smoking status and these cardiometabolic traits may be due to causality. Consistent with previous large cohort (4, 5) and MR studies (104–106), our exploratory bidirectional MR analysis found that smoking status had significant positive causal effects on T2D, CAD, MI, and HF, which suggests that the genetic correlations of the above trait pairs are attributed to both shared genetic architecture and causality. However, we did not observe a significant causal association between smoking status and IS, which is inconsistent with two recent studies (104, 105). This may be due to the different definitions of smoking, involving different ancestry populations, and different sample sizes, which need further confirmation. Besides, we did not observe any causality in the detection of the causal effect of 15 cardiometabolic traits on smoking status, excluding the possibility of reverse causation between smoking status and T2D or CVDs. The potential mechanisms underlying the causal relationship between smoking and T2D or CVDs require further investigation, but the shared loci and related pathways could provide new insights and directions.

In addition, we explored the genetic correlations between smoking status and T2D/CVD-related metabolic traits and observed a nominal positive correlation of smoking status with FG, a weak negative correlation of smoking status with HDL-C, and a weak positive correlation of smoking status with TG. Lipid and glycemic traits, resulting from complex and interwoven physiological mechanisms, are indicators of T2D and CVD risks, and understanding their associations with smoking can provide better insight into the pathophysiological intersect of T2D and CVDs. Previous studies have proven the role of smoking in elevating plasma TG concentration, decreasing plasma HDL-C concentration (107), and increasing the risk of impaired FG (108) and insulin resistance (109), which enhance the increased risk of T2D and CVDs. Although smoking cessation can ameliorate these changes, it is worth noting that smoking cessation is frequently followed by weight gain, which can contribute to the increased short-term risk of T2D (5, 110). Therefore, for smokers at risk for T2D, smoking cessation should be coupled with strategies for T2D prevention and early detection (5).

We acknowledge the limitations of our study. Despite the large sample sizes and high power of the GWAS summary statistics coming from meta-analysis studies, the homogeneity

among different summary statistics was reduced. However, each study conducted study-specific quality control to ensure data quality. In addition, simulations have confirmed that the effect of population structure and cryptic relatedness could be controlled well by our cross-trait meta-analysis method CPASSOC. Second, because of the concerns on sample size, accuracy, and availability of the GWAS data, we only analyzed smoking status in this study and did not consider quantitative smoking phenotypes such as cigarettes smoked per day or the years of smoking. Besides, smokeless tobacco products such as snuff tend to show different associations with T2D or CVDs as compared to cigarette smoking (111–113). It is important to consider these phenotypes in future investigations to shed light on the relationship between smoking and T2D or CVDs. Third, limited to the existing original GWASs, the sample sizes of some original trait-specific GWASs, especially ICH, were relatively small, which resulted in limited statistical power (**Supplementary Table 25**). Fourth, to yield reliable results, we used the data from the largest or latest GWASs, but there may be sample overlap between smoking status and T2D, CAD, and HF, which can influence the inference of causality in MR analysis. However, we used additional GWAS data of these traits with no sample overlap with smoking status GWAS to further confirm our primary findings and observed highly consistent results. Such consistency reinforced the robustness of our findings. Fifth, additional appropriate data were not available for us to replicate our findings. However, we used the data from the largest or latest GWASs for these traits to yield reliable results, and if possible, we will perform replication analysis in the future. Finally, our study was limited to assessing the shared genetic etiology between smoking status and T2D or CVDs. The effects of environmental factors and gene–environment interactions between smoking status and T2D or CVDs still need to be explored in further studies.

In summary, our findings provide strong evidence on shared genetic etiology and causal associations between smoking status and T2D or CVDs, underscoring the potential shared biological mechanisms underlying the link between smoking and T2D or CVDs. This work is important and opens up a new way for more effective and timely prevention, diagnosis, and treatment of smoking-related T2D or CVDs.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**. The download links for all the data relevant to the study can be found in the **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the relevant institutional review boards. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YC, XW, JJ, and TH designed the research. JJ and TH had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. YC and XW wrote the paper and performed the data analysis. All authors contributed to the statistical analysis, critically reviewed the manuscript during the writing process, and approved the final version to be published. YC, XW, JJ, and TH are the guarantors for the study.

FUNDING

The study was supported by grants from the National Key R&D Program of China (2019YFC2003400), the Peking University Start-up Grant (BMU2018YJ002), and the high-performance Computing Platform of Peking University. The funding organization had no role in the preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.809445/full#supplementary-material

REFERENCES

1. World Health Organization. *WHO Report on the Global Tobacco Epidemic, 2017: Monitoring Tobacco Use and Prevention Policies*. Available at: https://www.who.int/publications/i/item/9789241512824 (Accessed January 13, 2022).
2. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General*. Atlanta (GA: Centers for Disease Control and Prevention (US (2014).
3. Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st Century Hazards of Smoking and Benefits of Stopping: A Prospective Study of One Million Women in the UK. *Lancet* (2013) 381(9861):133–41. doi: 10.1016/s0140-6736(12)61720-6
4. Banks E, Joshy G, Korda RJ, Stavreski B, Soga K, Egger S, et al. Tobacco Smoking and Risk of 36 Cardiovascular Disease Subtypes: Fatal and Non-Fatal Outcomes in a Large Prospective Australian Study. *BMC Med* (2019) 17(1):128. doi: 10.1186/s12916-019-1351-4
5. Yeh HC, Duncan BB, Schmidt MI, Wang NY, Brancati FL. Smoking, Smoking Cessation, and Risk for Type 2 Diabetes Mellitus: A Cohort Study. *Ann Intern Med* (2010) 152(1):10–7. doi: 10.7326/0003-4819-152-1-201001050-00005
6. Wannamethee SG, Shaper AG, Perry IJ. Smoking as a Modifiable Risk Factor for Type 2 Diabetes in Middle-Aged Men. *Diabetes Care* (2001) 24 (9):1590–5. doi: 10.2337/diacare.24.9.1590
7. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active Smoking and the Risk of Type 2 Diabetes: A Systematic Review and Meta-Analysis. *Jama* (2007) 298(22):2654–64. doi: 10.1001/jama.298.22.2654
8. Pan A, Wang Y, Talaei M, Hu FB, Wu T. Relation of Active, Passive, and Quitting Smoking With Incident Type 2 Diabetes: A Systematic Review and Meta-Analysis. *Lancet Diabetes Endocrinol* (2015) 3(12):958–67. doi: 10.1016/s2213-8587(15)00316-2

9. Mons U, Müezzinler A, Gellert C, Schöttker B, Abnet CC, Bobak M, et al. Impact of Smoking and Smoking Cessation on Cardiovascular Events and Mortality Among Older Adults: Meta-Analysis of Individual Participant Data From Prospective Cohort Studies of the CHANCES Consortium. *BMJ (Clin Res Ed)* (2015) 350:h1551. doi: 10.1136/bmj.h1551
10. Huxley RR, Woodward M. Cigarette Smoking as a Risk Factor for Coronary Heart Disease in Women Compared With Men: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. *Lancet (Lond Engl)* (2011) 378 (9799):1297–305. doi: 10.1016/s0140-6736(11)60781-2
11. World Health Organization. *Tobacco Responsible for 20% of Deaths From Coronary Heart Disease* (2020). Available at: https://www.who.int/news/item/22-09-2020-tobacco-responsible-for-20-of-deaths-from-coronary-heart-disease (Accessed January 13, 2022).
12. Ockene IS, Miller NH. Cigarette Smoking, Cardiovascular Disease, and Stroke: A Statement for Healthcare Professionals From the American Heart Association. American Heart Association Task Force on Risk Reduction. *Circulation* (1997) 96(9):3243–7. doi: 10.1161/01.cir.96.9.3243
13. Li MD, Cheng R, Ma JZ, Swan GE. A Meta-Analysis of Estimated Genetic and Environmental Effects on Smoking Behavior in Male and Female Adult Twins. *Addict (Abingdon Engl)* (2003) 98(1):23–31. doi: 10.1046/j.1360-0443.2003.00295.x
14. Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, et al. Heritability and Familiality of Type 2 Diabetes and Related Quantitative Traits in the Botnia Study. *Diabetologia* (2011) 54(11):2811–9. doi: 10.1007/s00125-011-2267-5
15. Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. Heritability of Death From Coronary Heart Disease: A 36-Year Follow-Up of 20 966 Swedish Twins. *J Intern Med* (2002) 252(3):247–54. doi: 10.1046/j.1365-2796.2002.01029.x
16. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association Studies of Up to 1.2 Million Individuals Yield New Insights Into the Genetic Etiology of Tobacco and Alcohol Use. *Nat Genet* (2019) 51(2):237–44. doi: 10.1038/s41588-018-0307-5
17. Xu K, Li B, McGinnis KA, Vickers-Smith R, Dao C, Sun N, et al. Genome-Wide Association Study of Smoking Trajectory and Meta-Analysis of Smoking Status in 842,000 Individuals. *Nat Commun* (2020) 11(1):5302. doi: 10.1038/s41467-020-18489-3
18. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps. *Nat Genet* (2018) 50(11):1505–13. doi: 10.1038/s41588-018-0241-6
19. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res* (2018) 122(3):433–43. doi: 10.1161/circresaha.117.312086
20. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association Analyses Based on False Discovery Rate Implicate New Loci for Coronary Artery Disease. *Nat Genet* (2017) 49(9):1385–91. doi: 10.1038/ng.3913
21. Matoba N, Akiyama M, Ishigaki K, Kanai M, Takahashi A, Momozawa Y, et al. GWAS of Smoking Behaviour in 165,436 Japanese People Reveals Seven New Loci and Shared Genetic Architecture. *Nat Hum Behav* (2019) 3 (5):471–7. doi: 10.1038/s41562-019-0557-y
22. Duncan A, Heyer MP, Ishikawa M, Caligiuri SPB, Liu XA, Chen Z, et al. Habenular TCF7L2 Links Nicotine Addiction to Diabetes. *Nature* (2019) 574 (7778):372–7. doi: 10.1038/s41586-019-1653-x
23. Bentley AR, Sung YJ, Brown MR, Winkler TW, Kraja AT, Ntalla I, et al. Multi-Ancestry Genome-Wide Gene–Smoking Interaction Study of 387,272 Individuals Identifies New Loci Associated With Serum Lipids. *Nat Genet* (2019) 51(4):636–48. doi: 10.1038/s41588-019-0378-y
24. Wang XL, Greco M, Sim AS, Duarte N, Wang J, Wilcken DE. Effect of CYP1A1 MspI Polymorphism on Cigarette Smoking Related Coronary Artery Disease and Diabetes. *Atherosclerosis* (2002) 162(2):391–7. doi: 10.1016/s0021-9150(01)00723-7
25. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang JSchizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score Regression Distinguishes Confounding From Polygenicity in Genome-Wide Association Studies. *Nat Genet* (2015) 47(3):291–5. doi: 10.1038/ng.3211
26. Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, et al. Meta-Analysis of Correlated Traits *via* Summary Statistics From GWASs With an Application in Hypertension. *Am J Hum Genet* (2015) 96(1):21–36. doi: 10.1016/j.ajhg.2014.11.011
27. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative Approaches for Large-Scale Transcriptome-Wide Association Studies. *Nat Genet* (2016) 48(3):245–52. doi: 10.1038/ng.3506
28. Smith GD, Ebrahim S. 'Mendelian Randomization': Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease? *Int J Epidemiol* (2003) 32(1):1–22. doi: 10.1093/ije/dyg070
29. Davey Smith G, Hemani G. Mendelian Randomization: Genetic Anchors for Causal Inference in Epidemiological Studies. *Hum Mol Genet* (2014) 23(R1): R89–98. doi: 10.1093/hmg/ddu328
30. Pierce BL, Burgess S. Efficient Design for Mendelian Randomization Studies: Subsample and 2-Sample Instrumental Variable Estimators. *Am J Epidemiol* (2013) 178(7):1177–84. doi: 10.1093/aje/kwt084
31. Bowden J, Davey Smith G, Burgess S. Mendelian Randomization With Invalid Instruments: Effect Estimation and Bias Detection Through Egger Regression. *Int J Epidemiol* (2015) 44(2):512–25. doi: 10.1093/ije/dyv080
32. GeneATLAS. Available at: http://geneatlas.roslin.ed.ac.uk/trait/?traits=675 (Accessed January 13, 2022).
33. Canela-Xandri O, Rawlik K, Tenesa A. An Atlas of Genetic Associations in UK Biobank. *Nat Genet* (2018) 50(11):1593–9. doi: 10.1038/s41588-018-0248-z
34. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A Comprehensive 1,000 Genomes-Based Genome-Wide Association Meta-Analysis of Coronary Artery Disease. *Nat Genet* (2015) 47(10):1121–30. doi: 10.1038/ng.3396
35. Shah S, Henry A, Roselli C, Lin H, Sveinbjornsson G, Fatemifar G, et al. Genome-Wide Association and Mendelian Randomisation Analysis Provide Insights Into the Pathogenesis of Heart Failure. *Nat Commun* (2020) 11 (1):163. doi: 10.1038/s41467-019-13690-5
36. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry Genome-Wide Association Study of 520,000 Subjects Identifies 32 Loci Associated With Stroke and Stroke Subtypes. *Nat Genet* (2018) 50(4):524–37. doi: 10.1038/s41588-018-0058-3
37. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, et al. Meta-Analysis of Genome-Wide Association Studies Identifies 1q22 as a Susceptibility Locus for Intracerebral Hemorrhage. *Am J Hum Genet* (2014) 94(4):511–21. doi: 10.1016/j.ajhg.2014.02.012
38. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J, et al. Large-Scale Analyses of Common and Rare Variants Identify 12 New Loci Associated With Atrial Fibrillation. *Nat Genet* (2017) 49(6):946–52. doi: 10.1038/ng.3843
39. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New Genetic Loci Implicated in Fasting Glucose Homeostasis and Their Impact on Type 2 Diabetes Risk. *Nat Genet* (2010) 42(2):105–16. doi: 10.1038/ng.520
40. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and Refinement of Loci Associated With Lipid Levels. *Nat Genet* (2013) 45(11):1274–83. doi: 10.1038/ng.2797
41. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An Atlas of Genetic Correlations Across Human Diseases and Traits. *Nat Genet* (2015) 47(11):1236–41. doi: 10.1038/ng.3406
42. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, et al. Partitioning Heritability by Functional Annotation Using Genome-Wide Association Summary Statistics. *Nat Genet* (2015) 47(11):1228–35. doi: 10.1038/ng.3404
43. Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, et al. Meta-Analysis of 375,000 Individuals Identifies 38 Susceptibility Loci for Migraine. *Nat Genet* (2016) 48(8):856–66. doi: 10.1038/ng.3598
44. Wellcome Trust Case Control C, Maller JB, McVean G, Byrnes J, Vukcevic D, Palin K, et al. Bayesian Refinement of Association Signals for 14 Loci in 3 Common Diseases. *Nat Genet* (2012) 44(12):1294–301. doi: 10.1038/ng.2435
45. Schizophrenia Working Group of the Psychiatric Genomics C. Biological Insights From 108 Schizophrenia-Associated Genetic Loci. *Nature* (2014) 511(7510):421–7. doi: 10.1038/nature13595

46. Wang X, Jia J, Huang T. Shared Genetic Architecture and Casual Relationship Between Leptin Levels and Type 2 Diabetes: Large-Scale Cross-Trait Meta-Analysis and Mendelian Randomization Analysis. *BMJ Open Diabetes Res Care* (2020) 8(1):e001140. doi: 10.1136/bmjdrc-2019-001140
47. Wallace C. Statistical Testing of Shared Genetic Control for Potentially Related Traits. *Genet Epidemiol* (2013) 37(8):802–13. doi: 10.1002/gepi.21765
48. Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical Independence of the Colocalized Association Signals for Type 1 Diabetes and RPS26 Gene Expression on Chromosome 12q13. *Biostatistics* (2009) 10(2):327–34. doi: 10.1093/biostatistics/kxn039
49. Zhu Z, Lee PH, Chaffin MD, Chung W, Loh PR, Lu Q, et al. A Genome-Wide Cross-Trait Analysis From UK Biobank Highlights the Shared Genetic Architecture of Asthma and Allergic Diseases. *Nat Genet* (2018) 50 (6):857–64. doi: 10.1038/s41588-018-0121-0
50. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-Based Map of the Human Proteome. *Science* (2015) 347(6220):1260419. doi: 10.1126/science.1260419
51. Jain A, Tuteja G. TissueEnrich: Tissue-Specific Gene Enrichment Analysis. *Bioinformatics* (2019) 35(11):1966–7. doi: 10.1093/bioinformatics/bty890
52. Wang J, Duncan D, Shi Z, Zhang B. WEB-Based GEne SeT AnaLysis Toolkit (WebGestalt): Update 2013. *Nucleic Acids Res* (2013) 41(Web Server issue): W77–83. doi: 10.1093/nar/gkt439
53. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: Tool for the Unification of Biology. The Gene Ontology Consortium. *Nat Genet* (2000) 25(1):25–9. doi: 10.1038/75556
54. Gene Ontology Consortium. Gene Ontology Consortium: Going Forward. *Nucleic Acids Res* (2015) 43(Database issue):D1049–56. doi: 10.1093/nar/gku1179
55. Choi KW, Chen CY, Stein MB, Klimentidis YC, Wang MJ, Koenen KC, et al. Assessment of Bidirectional Relationships Between Physical Activity and Depression Among Adults: A 2-Sample Mendelian Randomization Study. *JAMA Psychiatry* (2019) 76(4):399–408. doi: 10.1001/jamapsychiatry.2018.4175
56. Benn M, Nordestgaard BG. From Genome-Wide Association Studies to Mendelian Randomization: Novel Opportunities for Understanding Cardiovascular Disease Causality, Pathogenesis, Prevention, and Treatment. *Cardiovasc Res* (2018) 114(9):1192–208. doi: 10.1093/cvr/cvy045
57. FinnGen. *FinnGen (Documentation of R5 Release)* (2021). Available at: https://finngen.gitbook.io/documentation/ (Accessed January 13, 2022).
58. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* (2017) 66(11):2888–902. doi: 10.2337/db16-1253
59. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-Scale Association Analysis Identifies 13 New Susceptibility Loci for Coronary Artery Disease. *Nat Genet* (2011) 43(4):333–8. doi: 10.1038/ng.784
60. Loos RJF, Bouchard C. FTO: The First Gene Contributing to Common Forms of Human Obesity. *Obes Rev* (2008) 9(3):246–50. doi: 10.1111/j.1467-789X.2008.00481.x
61. Krüger N, Biwer LA, Good ME, Ruddiman CA, Wolpe AG, DeLalio LJ, et al. Loss of Endothelial FTO Antagonizes Obesity-Induced Metabolic and Vascular Dysfunction. *Circ Res* (2020) 126(2):232–42. doi: 10.1161/circresaha.119.315531
62. Freathy RM, Timpson NJ, Lawlor DA, Pouta A, Ben-Shlomo Y, Ruokonen A, et al. Common Variation in the FTO Gene Alters Diabetes-Related Metabolic Traits to the Extent Expected Given its Effect on BMI. *Diabetes* (2008) 57(5):1419–26. doi: 10.2337/db07-1466
63. Ahmad T, Chasman DI, Mora S, Pare G, Cook NR, Buring JE, et al. The Fat-Mass and Obesity-Associated (FTO) Gene, Physical Activity, and Risk of Incident Cardiovascular Events in White Women. *Am Heart J* (2010) 160 (6):1163–9. doi: 10.1016/j.ahj.2010.08.002
64. Cunnington MS, Santibanez Koref M, Mayosi BM, Burn J, Keavney B. Chromosome 9p21 SNPs Associated With Multiple Disease Phenotypes Correlate With ANRIL Expression. *PloS Genet* (2010) 6(4):e1000899. doi: 10.1371/journal.pgen.1000899
65. Song C, Qi Y, Zhang J, Guo C, Yuan C. CDKN2B-AS1: An Indispensable Long Non-Coding RNA in Multiple Diseases. *Curr Pharm Design* (2020) 26 (41):5335–46. doi: 10.2174/1381612826666200806102424
66. Lin S, Shi Q, Nix FB, Styblo M, Beck MA, Herbin-Davis KM, et al. A Novel S-Adenosyl-L-Methionine:Arsenic(III) Methyltransferase From Rat Liver Cytosol. *J Biol Chem* (2002) 277(13):10795–803. doi: 10.1074/jbc.M110246200
67. Yuan Y, Marshall G, Ferreccio C, Steinmaus C, Selvin S, Liaw J, et al. Acute Myocardial Infarction Mortality in Comparison With Lung and Bladder Cancer Mortality in Arsenic-Exposed Region II of Chile From 1950 to 2000. *Am J Epidemiol* (2007) 166(12):1381–91. doi: 10.1093/aje/kwm238
68. Morris J, Bailey MES, Baldassarre D, Cullen B, de Faire U, Ferguson A, et al. Genetic Variation in CADM2 as a Link Between Psychological Traits and Obesity. *Sci Rep* (2019) 9(1):7339. doi: 10.1038/s41598-019-43861-9
69. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six New Loci Associated With Body Mass Index Highlight a Neuronal Influence on Body Weight Regulation. *Nat Genet* (2009) 41(1):25–34. doi: 10.1038/ng.287
70. Wiemerslage L, Gohel PA, Maestri G, Hilmarsson TG, Mickael M, Fredriksson R, et al. The Drosophila Ortholog of TMEM18 Regulates Insulin and Glucagon-Like Signaling. *J Endocrinol* (2016) 229(3):233–43. doi: 10.1530/JOE-16-0040
71. Gong G, O'Bryant SE. Low-Level Arsenic Exposure, AS3MT Gene Polymorphism and Cardiovascular Diseases in Rural Texas Counties. *Environ Res* (2012) 113:52–7. doi: 10.1016/j.envres.2012.01.003
72. Grau-Perez M, Kuo CC, Gribble MO, Balakrishnan P, Jones Spratlen M, Vaidya D, et al. Association of Low-Moderate Arsenic Exposure and Arsenic Metabolism With Incident Diabetes and Insulin Resistance in the Strong Heart Family Study. *Environ Health Perspect* (2017) 125(12):127004. doi: 10.1289/EHP2566
73. Drobná Z, Del Razo LM, García-Vargas GG, Sánchez-Peña LC, Barrera-Hernández A, Stýblo M, et al. Environmental Exposure to Arsenic, AS3MT Polymorphism and Prevalence of Diabetes in Mexico. *J Exposure Sci Environ Epidemiol* (2013) 23(2):151–5. doi: 10.1038/jes.2012.103
74. SMG6. *Nonsense Mediated mRNA Decay Factor [Homo Sapiens (Human)] - Gene - NCBI*. Available at: https://www.ncbi.nlm.nih.gov/gene/23293 (Accessed January 13, 2022).
75. Steenaard RV, Ligthart S, Stolk L, Peters MJ, van Meurs JB, Uitterlinden AG, et al. Tobacco Smoking is Associated With Methylation of Genes Related to Coronary Artery Disease. *Clin Epigenet* (2015) 7:54. doi: 10.1186/s13148-015-0088-y
76. Mega JL, Stitziel NO, Smith JG, Chasman DI, Caulfield M, Devlin JJ, et al. Genetic Risk, Coronary Heart Disease Events, and the Clinical Benefit of Statin Therapy: An Analysis of Primary and Secondary Prevention Trials. *Lancet (Lond Engl)* (2015) 385(9984):2264–71. doi: 10.1016/s0140-6736(14)61730-x
77. Carlsson S, Kuja-Halkola R, Magnusson C, Lagerros YT, Andersson T. Tobacco and Type 2 Diabetes: Is the Association Explained by Genetic Factors? *Int J Epidemiol* (2019) 48(3):926–33. doi: 10.1093/ije/dyz002
78. Fesinmeyer MD, North KE, Lim U, Bůžková P, Crawford DC, Haessler J, et al. Effects of Smoking on the Genetic Risk of Obesity: The Population Architecture Using Genomics and Epidemiology Study. *BMC Med Genet* (2013) 14:6. doi: 10.1186/1471-2350-14-6
79. States JC, Srivastava S, Chen Y, Barchowsky A. Arsenic and Cardiovascular Disease. *Toxicol Sci* (2009) 107(2):312–23. doi: 10.1093/toxsci/kfn236
80. Mendelson JH, Sholar MB, Goletiani N, Siegel AJ, Mello NK. Effects of Low- and High-Nicotine Cigarette Smoking on Mood States and the HPA Axis in Men. *Neuropsychopharmacology* (2005) 30(9):1751–63. doi: 10.1038/sj.npp.1300753
81. Rosmond R, Björntorp P. The Hypothalamic-Pituitary-Adrenal Axis Activity as a Predictor of Cardiovascular Disease, Type 2 Diabetes and Stroke. *J Intern Med* (2000) 247(2):188–97. doi: 10.1046/j.1365-2796.2000.00603.x
82. Mendelson JH, Goletiani N, Sholar MB, Siegel AJ, Mello NK. Effects of Smoking Successive Low- and High-Nicotine Cigarettes on Hypothalamic–Pituitary–Adrenal Axis Hormones and Mood in Men. *Neuropsychopharmacology* (2008) 33(4):749–60. doi: 10.1038/sj.npp.1301455
83. Yoshikawa H, Hellström-Lindahl E, Grill V. Evidence for Functional Nicotinic Receptors on Pancreatic Beta Cells. *Metabolism: Clin Exp* (2005) 54(2):247–54. doi: 10.1016/j.metabol.2004.08.020
84. Xie XT, Liu Q, Wu J, Wakui M. Impact of Cigarette Smoking in Type 2 Diabetes Development. *Acta Pharmacol Sin* (2009) 30(6):784–7. doi: 10.1038/aps.2009.49

85. Somm E, Schwitzgebel VM, Vauthay DM, Camm EJ, Chen CY, Giacobino JP, et al. Prenatal Nicotine Exposure Alters Early Pancreatic Islet and Adipose Tissue Development With Consequences on the Control of Body Weight and Glucose Metabolism Later in Life. *Endocrinology* (2008) 149 (12):6289–99. doi: 10.1210/en.2008-0361
86. Bruin JE, Gerstein HC, Morrison KM, Holloway AC. Increased Pancreatic Beta-Cell Apoptosis Following Fetal and Neonatal Exposure to Nicotine is Mediated *via* Mitochondria. *Toxicol Sci* (2008) 103(2):362–70. doi: 10.1093/toxsci/kfn012
87. Bruin JE, Petre MA, Raha S, Morrison KM, Gerstein HC, Holloway AC. Fetal and Neonatal Nicotine Exposure in Wistar Rats Causes Progressive Pancreatic Mitochondrial Damage and Beta Cell Dysfunction. *PloS One* (2008) 3(10):e3371. doi: 10.1371/journal.pone.0003371
88. Bruin JE, Petre MA, Lehman MA, Raha S, Gerstein HC, Morrison KM, et al. Maternal Nicotine Exposure Increases Oxidative Stress in the Offspring. *Free Radical Biol Med* (2008) 44(11):1919–25. doi: 10.1016/j.freeradbiomed.2008.02.010
89. Li LF, Chan RL, Lu L, Shen J, Zhang L, Wu WK, et al. Cigarette Smoking and Gastrointestinal Diseases: The Causal Relationship and Underlying Molecular Mechanisms (Review). *Int J Mol Med* (2014) 34(2):372–80. doi: 10.3892/ijmm.2014.1786
90. Liuzzo JP, Ambrose JA. Chest Pain From Gastroesophageal Reflux Disease in Patients With Coronary Artery Disease. *Cardiol Rev* (2005) 13(4):167–73. doi: 10.1097/01.crd.0000148844.13702.ce
91. Manfrini O, Bazzocchi G, Luati A, Borghi A, Monari P, Bugiardini R. Coronary Spasm Reflects Inputs From Adjacent Esophageal System. *Am J Physiol Heart Circ Physiol* (2006) 290(5):H2085–91. doi: 10.1152/ajpheart.00925.2005
92. Middlekauff HR, Park J, Moheimani RS. Adverse Effects of Cigarette and Noncigarette Smoke Exposure on the Autonomic Nervous System: Mechanisms and Implications for Cardiovascular Risk. *J Am Coll Cardiol* (2014) 64(16):1740–50. doi: 10.1016/j.jacc.2014.06.1201
93. Zucker IH, Wang W, Brändle M, Schultz HD, Patel KP. Neural Regulation of Sympathetic Nerve Activity in Heart Failure. *Prog Cardiovasc Dis* (1995) 37 (6):397–414. doi: 10.1016/s0033-0620(05)80020-9
94. Lichtor T, Davis HR, Johns L, Vesselinovitch D, Wissler RW, Mullan S. The Sympathetic Nervous System and Atherosclerosis. *J Neurosurg* (1987) 67 (6):906–14. doi: 10.3171/jns.1987.67.6.0906
95. Dechanet C, Anahory T, Mathieu Daude JC, Quantin X, Reyftmann L, Hamamah S, et al. Effects of Cigarette Smoking on Reproduction. *Hum Reprod Update* (2011) 17(1):76–95. doi: 10.1093/humupd/dmq033
96. Maresch CC, Stute DC, Alves MG, Oliveira PF, de Kretser DM, Linn T. Diabetes-Induced Hyperglycemia Impairs Male Reproductive Function: A Systematic Review. *Hum Reprod Update* (2018) 24(1):86–105. doi: 10.1093/humupd/dmx033
97. Ding GL, Liu Y, Liu ME, Pan JX, Guo MX, Sheng JZ, et al. The Effects of Diabetes on Male Fertility and Epigenetic Regulation During Spermatogenesis. *Asian J Andrology* (2015) 17(6):948–53. doi: 10.4103/1008-682x.150844
98. Jin T, Liu L. The Wnt Signaling Pathway Effector TCF7L2 and Type 2 Diabetes Mellitus. *Mol Endocrinol* (2008) 22(11):2383–92. doi: 10.1210/me.2008-0135
99. Cheng HY, Schiavone AP, Smithgall TE. A Point Mutation in the N-Terminal Coiled-Coil Domain Releases c-Fes Tyrosine Kinase Activity and Survival Signaling in Myeloid Leukemia Cells. *Mol Cell Biol* (2001) 21 (18):6170–80. doi: 10.1128/MCB.21.18.6170-6180.2001
100. Karamanavi E, Mcvey DG, van der Laan SW, Yang W, Chan K, Poston RN, et al. The FES Gene, Located at the Chromosome 15Q21.6 Coronary-Artery-Disease Locus, Modulates Atherosclerotic Plaque Vulnerability. *Atherosclerosis* (2020) 315:e19. doi: 10.1016/j.atherosclerosis.2020.10.070
101. Bindesboll C, Aas A, Ogmundsdottir MH, Pankiv S, Reine T, Zoncu R, et al. NBEAL1 Controls SREBP2 Processing and Cholesterol Metabolism and is a Susceptibility Locus for Coronary Artery Disease. *Sci Rep* (2020) 10(1):4528. doi: 10.1038/s41598-020-61352-0
102. Lausberg E, Gießelmann S, Dewulf JP, Wiame E, Holz A, Salvarinova R, et al. C2orf69 Mutations Disrupt Mitochondrial Function and Cause a Multisystem Human Disorder With Recurring Autoinflammation. *J Clin Invest* (2021) 131(12):e143078. doi: 10.1172/jci143078
103. Wong HH, Seet SH, Maier M, Gurel A, Traspas RM, Lee C, et al. Loss of C2orf69 Defines a Fatal Autoinflammatory Syndrome in Humans and Zebrafish That Evokes a Glycogen-Storage-Associated Mitochondriopathy. *Am J Hum Genet* (2021) 108(7):1301–17. doi: 10.1016/j.ajhg.2021.05.003
104. Larsson SC, Mason AM, Back M, Klarin D, Damrauer SM, Million Veteran P, et al. Genetic Predisposition to Smoking in Relation to 14 Cardiovascular Diseases. *Eur Heart J* (2020) 41(35):3304–10. doi: 10.1093/eurheartj/ehaa193
105. Levin MG, Klarin D, Assimes TL, Freiberg MS, Ingelsson E, Lynch J, et al. Genetics of Smoking and Risk of Atherosclerotic Cardiovascular Diseases: A Mendelian Randomization Study. *JAMA Netw Open* (2021) 4(1):e2034461. doi: 10.1001/jamanetworkopen.2020.34461
106. Yuan S, Larsson SC. A Causal Relationship Between Cigarette Smoking and Type 2 Diabetes Mellitus: A Mendelian Randomization Study. *Sci Rep* (2019) 9(1):19342. doi: 10.1038/s41598-019-56014-9
107. Chelland Campbell S, Moffatt RJ, Stamford BA. Smoking and Smoking Cessation – The Relationship Between Cardiovascular Disease and Lipoprotein Metabolism: A Review. *Atherosclerosis* (2008) 201(2):225–35. doi: 10.1016/j.atherosclerosis.2008.04.046
108. Nakanishi N, Nakamura K, Matsuo Y, Suzuki K, Tatara K. Cigarette Smoking and Risk for Impaired Fasting Glucose and Type 2 Diabetes in Middle-Aged Japanese Men. *Ann Intern Med* (2000) 133(3):183–91. doi: 10.7326/0003-4819-133-3-200008010-00009
109. Bergman BC, Perreault L, Hunerdosse D, Kerege A, Playdon M, Samek AM, et al. Novel and Reversible Mechanisms of Smoking-Induced Insulin Resistance in Humans. *Diabetes* (2012) 61(12):3156–66. doi: 10.2337/db12-0418
110. Hu Y, Zong G, Liu G, Wang M, Rosner B, Pan A, et al. Smoking Cessation, Weight Change, Type 2 Diabetes, and Mortality. *N Engl J Med* (2018) 379 (7):623–32. doi: 10.1056/NEJMoa1803626
111. Titova OE, Baron JA, Michaëlsson K, Larsson SC. Swedish Snuff (Snus) and Risk of Cardiovascular Disease and Mortality: Prospective Cohort Study of Middle-Aged and Older Individuals. *BMC Med* (2021) 19(1):111. doi: 10.1186/s12916-021-01979-6
112. Janzon E, Hedblad B. Swedish Snuff and Incidence of Cardiovascular Disease. A Population-Based Cohort Study. *BMC Cardiovasc Disord* (2009) 9:21. doi: 10.1186/1471-2261-9-21
113. Huhtasaari F, Lundberg V, Eliasson M, Janlert U, Asplund K. Smokeless Tobacco as a Possible Risk Factor for Myocardial Infarction: A Population-Based Study in Middle-Aged Men. *J Am Coll Cardiol* (1999) 34(6):1784–90. doi: 10.1016/s0735-1097(99)00409-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 01 March 2022
doi: 10.3389/fendo.2022.754401

Adjuvant Probiotics of *Lactobacillus salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *Bifidobacterium animalis* subsp. *lactis* CP-9 Attenuate Glycemic Levels and Inflammatory Cytokines in Patients With Type 1 Diabetes Mellitus

*Chung-Hsing Wang[1,2], Hung-Rong Yen[3,4,5,6], Wen-Li Lu[1], Hsieh-Hsun Ho[7], Wen-Yang Lin[7], Yi-Wei Kuo[7], Yen-Yu Huang[7], Shin-Yu Tsai[7] and Hung-Chih Lin[8,9,10]**

[1] Division of Medical Genetics, Pediatric Endocrinology & Metabolism, China Medical University Children's Hospital, China Medical University, Taichung, Taiwan, [2] School of Medicine, College of Medicine, China Medical University, Taichung, Taiwan, [3] Department of Chinese Medicine, China Medical University Hospital, Taichung, Taiwan, [4] Research Center for Traditional Chinese Medicine, China Medical University Hospital, Taichung, Taiwan, [5] Chinese Medicine Research Center, China Medical University, Taichung, Taiwan, [6] Department of Biotechnology, Asia University, Taichung, Taiwan, [7] Research and Development Department, Bioflag Biotech Co., Ltd., Tainan, Taiwan, [8] Division of Neonatology, China Medical University Children's Hospital, Taichung, Taiwan, [9] School of Chinese Medicine, China Medical University, Taichung, Taiwan, [10] Asia University Hospital, Asia University, Taichung, Taiwan

Introduction: Type 1 diabetes mellitus (T1DM) is characterized by autoimmune destruction of pancreatic β cells. Previous study has discovered that probiotic strains residing in the gut play essential roles in host immune regulation. However, few clinical results demonstrated probiotic would actually benefit in attenuating glycated hemoglobin (HbA1c) along with inflammatory cytokine levels of the T1DM patients and analyzed their gut microbiota profile at the same time. In this clinical trial, we evaluated the therapeutic efficacy of probiotics on HbA1c along with inflammatory cytokine levels of T1DM patients to determine an alternative administration mode for T1DM medication. The probiotics changed T1DM gut microbiota profile will be measured by next-generation sequencing (NGS).

Research Design and Methods: A randomized, double-blind, placebo-controlled trial was performed at China Medical University Hospital. T1DM patients between 6 and 18 years of age were enrolled. 27 patients were administered regular insulin therapy plus capsules containing probiotic strains *Lactobacillus salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *Bifidobacterium animalis* subsp. *lactis* CP-9 daily for 6 months, and 29 patients were administered insulin therapy without extra probiotic supplement as placebo group. The variations of fasting blood glucose and HbA1c in these patients were

OPEN ACCESS

Edited by:
Changwei Li,
Tulane University School of Public Health and Tropical Medicine, United States

Reviewed by:
Zhihong Sun,
Inner Mongolia Agricultural University, China
Pen-Hua Su,
Chung Shan Medical University, Taiwan

***Correspondence:**
Hung-Chih Lin
d0373@mail.cmuh.org.tw
orcid.org/0000-0002-8111-8371

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *09 September 2021*
Accepted: *31 January 2022*
Published: *01 March 2022*

Citation:
Wang C-H, Yen H-R, Lu W-L, Ho H-H, Lin W-Y, Kuo Y-W, Huang Y-Y, Tsai S-Y and Lin H-C (2022) Adjuvant Probiotics of Lactobacillus salivarius subsp. salicinius AP-32, L. johnsonii MH-68, and Bifidobacterium animalis subsp. lactis CP-9 Attenuate Glycemic Levels and Inflammatory Cytokines in Patients With Type 1 Diabetes Mellitus. Front. Endocrinol. 13:754401. doi: 10.3389/fendo.2022.754401

analyzed. In addition, serum levels of inflammatory cytokines and anti-inflammatory cytokine were assessed using enzyme-linked immunosorbent assay. Patients' stool microbiota were all subjects to NGS analysis.

Results: NGS data showed elevated populations of *Bifidobacterium animalis, Akkermansia muciniphila* and *Lactobacillus salivarius* in the gut of patients with T1DM who were taking probiotics. Patients with T1DM who were administered probiotics showed significantly reduced fasting blood glucose levels compared with the before-intervention levels. The HbA1c levels of the patients also improved after administration of probiotics. The concentrations of IL-8, IL-17, MIP-1β, RANTES, and TNF-α were significantly reduced and were associated with an increased TGF-β1 expression after probiotic intervention. The persistence effect of glycemic control and immunomodulation were observed even 3 months after discontinuation of the probiotics.

Conclusions: Here, we found that conventional insulin therapy plus probiotics supplementation attenuated T1DM symptoms than receiving insulin treatment only. Probiotics supplementation with insulin treatment changed gut microbiota and revealed better outcome in stabilizing glycemic levels and reducing HbA1c levels in patients with T1DM through beneficial regulation of immune cytokines.

Clinical Trial Registration: ClinicalTrials.gov, identifier NCT03880760.

Keywords: T1DM, probiotic, gut microbiota, glycemic levels, immune cytokines

INTRODUCTION

Type 1 diabetes mellitus (T1DM) is a chronic autoimmune disease, wherein pancreatic beta cells are attacked and disrupted by abnormal immune response (1). The pancreatic beta cells produce insulin and regulate glycemic homeostasis in the human body (2). Damage to the beta cells causes insulin insufficiency, and the elevation and accumulation of blood glucose leads to stroke, heart diseases, damage to the nervous system, impaired cognition, kidney failure, poor vision, and numbness in limbs (3, 4). An estimated 1,106,500 individuals aged 0–19 years worldwide are diagnosed as having T1DM, which accounts for 5%–10% of all patients with diabetes with >85% of them being young patients (5, 6). Insulin injection is the major clinical treatment for T1DM, but a cure for T1DM has not yet been reported (7). Therefore, understanding the pathogenesis of T1DM and prevention of its onset are the main strategies in the investigation of this disease (8).

Some researchers have reported that autoimmune reactions against beta cells may come from the activation of the immune system in genetically susceptible individuals, which is triggered by environmental factors that bear epitopes similar to those expressed by the beta cells. Several mechanisms, such as viral infections or other triggers (antigens from diet) were considered to initiate a hyperactive response toward β-cells (5, 9, 10). Other studies have publicized that 10-30% of patients with T1DM develop other autoimmune diseases (11). The clinicopathological mechanism of celiac disease is the one similar to T1DM, there is a high probability that both diseases have comorbidities. Interestingly, while celiac disease composes a proportion of the T1DM population in the Western countries (11), which is less common or nonexistent in the Asian population (12).

Studies have reported a link between gut microbiota and T1DM. Decreased population of Firmicutes and increased population of Bacteroidetes have been observed in most patients with T1DM. At the genus level, T1DM children contained higher population of *Veillonella*, *Clostridium*, and *Bacteroides* in intestine, whereas the healthy children had higher populations of *Prevotella, Blautia coccoides/Eubacterium* rectale group, *Lactobacillus*, and *Bifidobacterium* (13, 14). The connection between host immune cells and gut microbiota may also influence the pathogenesis of T1DM. Metabolites secreted by the gut microbes can regulate the immune response and slow down the development of process of T1DM (15, 16). In other words, dysbiosis of the gut microbiota may incite and accelerate the pathogenesis of T1DM (17).

The microbiota modulators and probiotics help to maintain a healthy homeostasis of gut microbiota, gut membrane integrity, and permeability and also upregulate anti-inflammatory cytokines such as transforming growth factor-β (TGF-β) while downregulating proinflammatory cytokines such as TNF-α (18). The functions of some microbial metabolites, including short-chain fatty acids and lactic acids, on modulation of innate immune response have also been reported (19). Thus, probiotics could be a potential regulator of T1DM and could help slow disease progression *via* the gut microbiota-immune axis.

Probiotic, which was defined as: "live microorganisms which when administered in adequate amounts confer a health benefit on the host" (20). However, few studies have demonstrated the

benefits of probiotics to patients with T1DM. Probiotic strains *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9 were screened by their ability of consuming glucose *in vitro* (21). In addition, *L. salivarius* subsp. *salicinius* AP-32 could enhance the expression of glucose transporter 2 in human intestinal epithelial cells and played a major role in regulating glucose metabolism and blood lipid levels in mice models of T2DM (21). Moreover, an *in vivo* study also indicated that *L. johnsonii* MH-68 and *L. salivarius* subsp. *salicinius* AP-32 reduced the levels of inflammatory chemokines (22).

The aim of this study is to improve glycemic control through reduction of inflammatory cytokines by means of daily use of probiotic strains. The strains used were *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9. Furthermore, the feces were tested for microbiota load to ensure that the consumption of the aforementioned probiotics was effectively colonized in the recipients' gut systems. The results of this study might provide a novel probiotic combination to ameliorate or even prevent the clinical symptoms of the chronic autoimmune disease, T1DM.

RESEARCH DESIGN AND METHODS

Patients with age of onset between 6 and 18 years and were diagnosed with T1DM (i.e., fulfilling 2 of the 3 characteristics of the following: 1) ketoacidemia, 2) insulin-related autoantibodies, and 3) insufficient insulin secreting capacity) were included in the study (23, 24). Given that these patients are longstanding diabetic patients who have already experienced the honeymoon periods, they have gone through thorough education regarding diabetes and diet. Patients with significant heart, kidney, and liver diseases; those diagnosed as having immunodeficiency and low immune function; those using probiotic-related products or those who had taken probiotics for more than 1 month; those who were taking antibiotics or stomach and intestinal drugs at the start of the trial; and those diagnosed as having allergies to probiotics were excluded. A total of 64 patients were initially recruited, 5 of whom withdrew from the study.

Thus, 59 patients were enrolled. *Via* computer-generated random numbering with double blinding, 27 were assigned to the probiotic group (female/male = 14/13) and 32 were assigned to the placebo group, out of which 3 withdrew due to loss to follow-up and extreme fluctuations in blood sugar (female/male = 10/19). Finally, a total of 56 patients were included for analysis (**Supplemental Figure S1**). Patients in the probiotic group were instructed to consume 1×10^{10} colony-forming units (CFU)/day of mixed probiotics (*L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9) for 6 months, during which their insulin regimens were given as per existing protocol (**Table 1**) (25), and 3-month follow-up (26). The fasting blood glucose (Glucose AC), glycated hemoglobin (HbA1c), immune cytokine concentration, and fecal bacterial phase changes before and after intervention were compared between the groups.

Probiotic Strains and Cultivation

Active probiotic strains *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9 were obtained from Bioflag Biotech Co., Ltd. (Tainan, Taiwan). AP-32 and MH-68 strains were isolated from human gut and CP-9 strain was isolated from breast milk. The deposition numbers for *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68 and *B. animalis* subsp. *lactis* CP-9 were CCTCC-M2011127, CCTCC-M2011128, and CCTCC-M2014588, respectively. The above-mentioned strains are preserved in China Center for Type Culture Collection (abbreviated as CCTCC) and China General Microbiological Culture Collection Center (abbreviated as CGMCC).

TABLE 1 | Baseline demographic and laboratory characteristics of patients with T1DM.

Parameter	Probiotics (n = 27)	Placebo (n = 29)	P-value
Age (years)	14.1 ± 5.1	14.3 ± 4.6	0.824
Sex (F/M)	14/13	10/19	0.280
T1DM duration (months)	74.4 ± 53.6	77.9 ± 46.9	0.640
Insulin regimen (N)			
BID AC	9	6	0.370
TID AC	2	1	0.605
QID AC	16	22	0.254
Insulin dosage (U/kg/day)	0.8 ± 0.3	0.8 ± 0.3	0.889
Glucose AC (mg/dL)	185.4 ± 41.5	172.2 ± 62.6	0.131
HbA1c (%) (mmol/mol)	9.3 ± 0.8	9.5 ± 1.9	0.883
	78.0 ± 8.9	79.9 ± 21.2	0.928
TNF-α (pg/mL)	52.5 ± 59.7	54.0 ± 56.8	0.812
IL-8 (pg/mL)	398.2 ± 233.4	477.7 ± 222.9	0.093
IL-17 (pg/mL)	26.8 ± 34.1	19.5 ± 24.5	0.441
MIP-1β (pg/mL)	113.9 ± 75.5	130.5 ± 86.9	0.549
RANTES (pg/mL)	447.9 ± 71.4	411.0 ± 69.0	0.068
TGF-β1 (pg/mL)	5624.2 ± 1984.6	6554.3 ± 1759.3	0.068

BID AC, Twice a day before meals; TID AC, Three times a day before meals; QID AC, Four times a day before meals and bedtime.
Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation.
Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.

De Man, Rogosa, and Sharpe (MRS) broth was used to culture *Lactobacillus* spp., and MRS broth supplemented with 0.05% cysteine was used to culture *Bifidobacterium* spp. The selected probiotics were incubated at 37°C under anaerobic conditions for 20 hours and subjected to lyophilization. The viability of dry bacteria was determined by analyzing the CFU. The dosage of the mixture of AP-32, MH-68, and CP-9 strains was 5×10^9 CFU/capsule. The mixing ratio of the three probiotics was 1:1:1.

Measurement of Serum Glucose (AC)

The patients underwent blood tests in the department of laboratory medicine in China Medical University Hospital, Taichung, Taiwan. The fasting blood glucose values were detected by Optium Xceed Glucometer (Abbott Diabetes Care Inc., Alameda, CA, USA) by endpoint detection colorimetry. Glucose AC was measured and analyzed before intervention, 6 months after intervention initiation, and 3 months after stopping the intervention in both the groups.

Measurement of Serum HbA1c Level

Glycated hemoglobin (HbA1c) value was measured using the HbA1c HPLC Assay Kit (Eagle Biosciences, Nashua, NH, USA, catalog number: A1C31-H100) with high-performance liquid chromatography (HPLC, PerkinElmer Series 200). The HbA1c was measured and analyzed before intervention, 6 months after intervention initiation, and 3 months after stopping the intervention in both the groups.

Measurement of Cytokine Levels by Enzyme-Linked Immunosorbent Assay (ELISA)

Extra blood left after the above tests was used to detect several inflammatory or anti-inflammatory cytokines before intervention, 6 months after intervention initiation, and 3 months after stopping the intervention. After centrifugation, the serum was refrigerated at −80°C until use for ELISA. The inflammatory cytokines used for ELISA test were as follows: TNF-α (Thermo Scientific, Carlsbad, CA, USA), IL-8 (Thermo Scientific, Carlsbad, CA, USA), IL-17 (R&D Systems, Minneapolis, Minnesota, USA), MIP-1β (PeproTech, Cranbury, NJ, USA), and RANTES (R&D Systems, Minneapolis, Minnesota, USA). The anti-inflammatory cytokine used was TGF-β1 (Thermo Scientific, Carlsbad, CA, USA). All samples were measured by the commercial protocol at least in triplicate.

Detection of Gut Microbiota Using NGS

The fecal samples of the patients were collected before and after taking probiotics, and immediately refrigerated at −80°C. Once the stool samples had thawed, the DNA (approximately 50–100 mg) was extracted by Quick-DNA™ fungal/bacterial micro-preparation reagent (ZYMO Research, Irvine, CA, USA). The DNA sample was then diluted to 5 ng/μL in sterile water and provided to Phalanx Biotech Co., Ltd., Taiwan for quality inspection on 1% agar gel electrophoresis. The DNA fragments (16S rRNA, 16S V3-V4) were amplified by PCR through special primers using the 2× KAPA HiFi HotStart ReadyMix Kit (KAPA Biosystems, USA). The DNA product of 460 base pair was selected for the next experiment. All amplified DNA samples were purified using AMPure XP beads (Beckman Coulter Genomics, USA). The purified DNA was used to generate a sequenced sample library using Illumina Nextera XT Index kit (Illumina, USA). The sequenced sample library was evaluated using Qubit 2.0 Fluorometer (Thermo Fisher Scientific, USA) and Agilent Bioanalyzer 2100 system (Agilent Technologies, Inc., USA). Finally, this sequencing library was sequenced on the Illumina MiSeq platform.

The adaptors were first trimmed from the sequencing reads. All paired-end reads were merged by PEAR v0.9.10. QIIME v.2 as the quality parameter was used to demultiplex and filter at Q20. PCR generated amplicons with chimeric sequences were detected by USEARCH. 99% similarity sequence of clustering operational taxonomic units (OTU) were analyzed by QIIME v.2 with SILVA v.132 database. Sequences that appeared once (singletons) or failed mapped to database were discarded. The taxonomy was classified after annotating OTU.

Statistical Analysis

Based on sample size calculation of previous study, the mean value of HbA1c at before-intervention was evaluated at 8.4% with a standard deviation of 1.3% for both groups (27). The sample size was calculated for two-tailed *t* test comparing two groups. In this per-protocol analysis, a clinically significant reduction of 1% in HbA1c was employed given that previous studies reported that every 1% reduction in HbA1c reduces the relative risk by 37% for microvascular complications, by 21% for diabetes-related deaths, and 14% for myocardial infarction (28, 29). For a power of 80% and a type 1 error of 0.05, the number of subjects needed for each arm of the study is 27. Assuming a drop-out rate of 10%, approximately 30 subjects was required for each arm in the study. The values of continuous variables are expressed as mean ± standard deviation. Fisher's exact test was used to compare categorical variables. Wilcoxon signed-rank test and Mann–Whitney *U* test were used to compare the differences between variables. For NGS analysis, the permutation method was used to test the taxonomic groups (phyla, phylum, order, family, genus, and species). Statistical analysis was performed using SPSS 12 (IBM, USA). $P < 0.05$ was considered statistically significant. Changes in the microbiota were generated by GraphPad Prism 8 (GraphPad Software, San Diego, CA, USA).

RESULTS

Patient Clinical Characteristics Before Probiotic Intervention

The clinical characteristics of all patients before intervention are shown in **Table 1**. A total of 27 patients with T1DM were assigned to the probiotic group and 29 were assigned to the placebo group (**Supplemental Figure S1**). The average age of the patients in the probiotic and placebo groups was 14.1 and 14.3 years, and T1DM duration were 74.4 and 77.9 months (**Table 1**). Among 27 patients under probiotic treatment, 9 patients were

injected insulin twice a day before meals (BID AC), 2 patients were injected insulin three times a day before meals (TID AC) and 16 patients were injected insulin four times a day before meals (QID AC). Among 29 patients in the placebo group, 6 patients were injected insulin before meals BID AC, 1 patient was TID AC, and 22 patients were QID AC. The dosage of injected insulin was analogous in two groups (0.8 U/kg/day). The average Glucose AC was 185.4 (mg/dL) in probiotic group and 172.2 (mg/dL) in placebo group, and the average HbA1c of the two groups was 9.3% (78.0 mmol/mol) and 9.5% (79.9 mmol/mol), respectively. The levels of six cytokines were measured for patients in the probiotic and placebo groups. The initial average values were as follows: the average TNF-α level was 52.5 and 54.0 pg/mL, IL-8 was 398.2 and 477.7 pg/mL, IL-17 was 26.8 and 19.5 pg/mL, MIP-1β was 113.9 and 130.5 pg/mL, RANTES was 447.9 and 411.0 pg/mL, and TGF-β1 was 5624.2 and 6554.3 pg/mL in the probiotic and placebo group, respectively. There was no statistically significant difference in any of the diagnostic items between the two groups before intervention (**Table 1**).

Elevated Population of *B. animalis* and *L. salivarius* in the Gut of Patients With T1DM After Taking Probiotics

NGS showed changes in the gut microbiota of the patients before and after intervention; the number of bacteria belonging to genus *Akkermansia*, phylum Verrucomicrobia reduced in both the groups after intervention (**Supplemental Tables S1**, **S2**). Next, we compared and analyzed the difference in the microbiota between the groups. Before the trial, the dominant gut microbiota phyla in both the groups were Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria, and Verrucomicrobia. Firmicutes accounted for the largest proportion in the probiotic (61.70%) and placebo groups (58.31%), and Actinobacteria had the second largest microbial population in the gut: 24.44% in the probiotic group and 27.05% in the placebo group. Verrucomicrobia accounted for only 4.33% in the probiotic group and 1.78% in the placebo group. However, six months after intervention, the populations of Firmicutes and Actinobacteria did not show significant differences between the groups. Compared with the placebo group, the probiotic group showed a significant difference in the population of Verrucomicrobia ($P = 0.035$) after the intervention (**Figure 1A**).

Changes in the gut microbiota at the genus level were further analyzed after the patients had taken the probiotics for 6 months. *Bifidobacterium* accounted for the major proportion in the probiotic group (24.59%) and in the placebo group (22.52%), *Lactobacillus* accounted for 0.35% in the probiotic group and 0.55% in the placebo group, and *Akkermansia* accounted for 0.28% in the probiotic group and 0.06% in the placebo group. Compared with the placebo group, the probiotic group showed no significant difference in the populations of *Bifidobacterium* and *Lactobacillus* after 6 months of intervention. At the genus level, the population of *Akkermansia* was significantly higher in the probiotic group than in the placebo group after intervention ($P = 0.035$; **Figure 1B**).

NGS analysis of the gut microbiota at the species level revealed that the populations of *B. animalis* ($P = 0.024$), *L. salivarius* ($P = 0.010$), and *Akkermansia* uncultured bacteria ($P = 0.035$) were significantly higher in the probiotic group than in the placebo group after probiotic intervention (**Figure 1C**). Compared with the population before intervention, the population of *L. salivarius* significantly increased after taking probiotics ($P = 0.001$; **Supplemental Table S3**).

Glycemic Control Declined in 6 Months After Consuming Probiotic Products

The average baseline fasting blood glucose (Glucose AC; mg/dL) before intervention was 185.4 mg/dL in the probiotic group and 172.2 mg/dL in the placebo group, and the average glycated hemoglobin (HbA1c) was 9.3% (78.0 mmol/mol) in the probiotic group and 9.5% (79.9 mmol/mol) in the placebo group **Table 2**). After 6-month probiotic intervention, the average Glucose AC decreased to 161.9 mg/dL ($P = 0.000$) and average glycated hemoglobin decreased to 8.5% (69.7mmol/mol, $P = 0.000$). Six months later, the average glycemic level of the placebo group was similar to the baseline (Glucose AC: 171.5 mg/dL; glycated hemoglobin 9.5% = 80.0 mmol/mol). Three months after stopping the intervention, the glycemic levels of the probiotic group were still lower than the baseline levels (Glucose AC: 163.9 mg/dL, $P = 0.002$; average glycated hemoglobin: 8.9% = 74.1 mmol/mol, $P = 0.005$), whereas the levels of Glucose AC (174.1 mg/dL) and average glycated hemoglobin (9.5% = 80.6 mmol/mol) remained the same as baseline in the placebo group (**Table 2**).

The rate of glycemic decline was evaluated by comparing the glycemic control between the probiotic group and placebo group in 6 months after intervention. Data were normalized to the baseline (before-intervention). The average fasting blood glucose rate in the probiotic group significantly dropped to 87.5%, compared with the placebo group, with a level of 101.5% ($P = 0.000$; **Table 2** and **Figure 2A**) and the average glycated hemoglobin (HbA1c) rate decreased to 91.8%, compared with the placebo group, with a rate of 100.1% ($P = 0.000$; **Table 2** and **Figure 2B**). Three months after stopping the intervention, the normalized glycemic levels of the probiotic group were still lower than those of the placebo group (Glucose AC: 89.0%, $P = 0.001$; glycated hemoglobin: 96.1%, $P = 0.038$; **Table 2**). This showed that probiotics can effectively reduce the excessively high blood glucose and glycated hemoglobin in patients with T1DM.

IL-8, TNF-α, IL-17, MIP-1β, and RANTES Levels Reduced After Taking Probiotics

The serum samples of the patients were collected for further measurements before and after intervention. The results showed that the average levels of the inflammatory cytokines IL-8 (350.4 pg/mL, $P = 0.000$), TNF-α (39.2 pg/mL, $P = 0.000$), IL-17 (19.7 pg/mL, $P = 0.001$), MIP-1β (99.6 pg/mL, $P = 0.006$), and RANTES (401.7 pg/mL, $P = 0.008$) in the probiotic group were significantly decreased after 6 months of probiotic intervention compared with the levels before intervention (**Table 3**). Three months after stopping the intervention, the IL-8 (359.7 pg/mL,

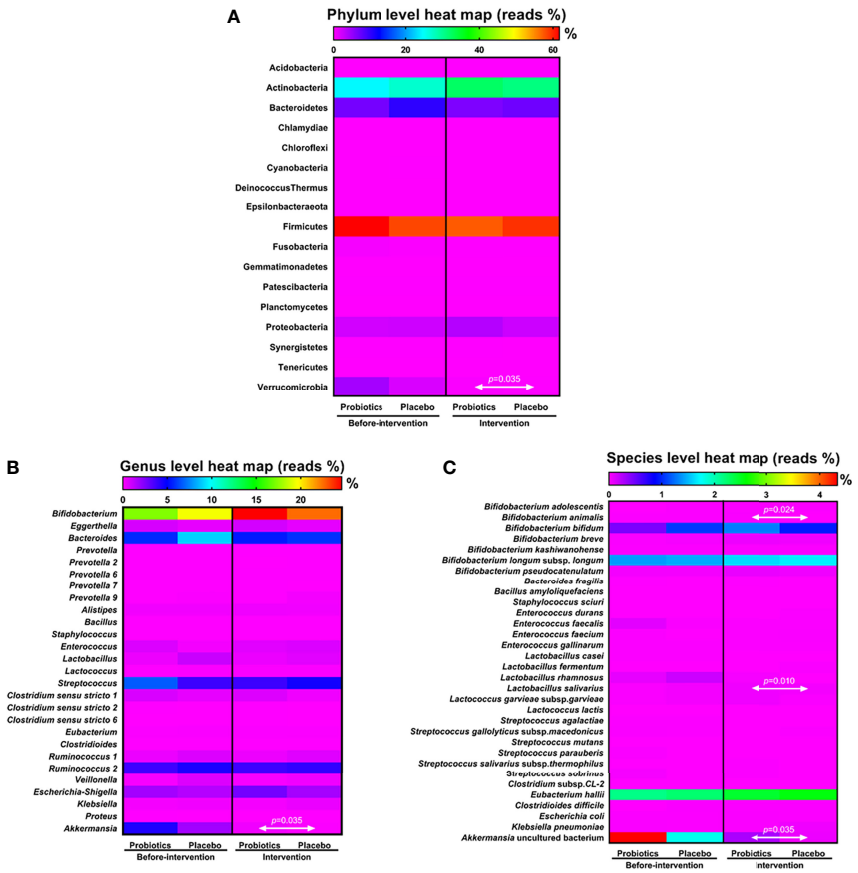


FIGURE 1 | Analysis of microbiota dispersion (%) among patients with T1DM. The data present microbiota changes at the phylum **(A)**, genus **(B)**, and species levels **(C)**. The amount of microbiota was normalized with that observed before intervention. Statistical difference between the probiotic group and placebo is indicated by a white two-way arrow. Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation. Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.

TABLE 2 | Glucose AC and HbA1c declined 6 months after consuming probiotic product.

(a.) Glycemic control.

	Before-intervention		6-month intervention		3-month after intervention	
	Probiotic group	**Placebo group**	**Probiotic group**	**Placebo group**	**Probiotic group**	**Placebo group**
Glucose AC (mg/dl)	185.4 ± 41.5	172.2 ± 62.6	161.9 ± 39.0$^{\#\#\#}$	171.5 ± 55.5	163.9 ± 39.0$^{\#\#}$	174.1 ± 58.8
HbA1c (%) (mmol/mol)	9.3 ± 0.8	9.5 ± 1.9	8.5 ± 0.9$^{\#\#\#}$	9.5 ± 2.1	8.9 ± 1.1$^{\#\#}$	9.5 ± 2.2
	78.0 ± 8.9	79.9 ± 21.2	69.7 ± 10.4$^{\#\#\#}$	80.0 ± 22.8	74.1 ± 11.6$^{\#\#}$	80.6 ± 24.5

Wilcoxon signed-rank test: when patients in each group were compared with their own before-intervention status, statistical difference was shown as $^{\#\#}P < 0.01$ and $^{\#\#\#}P < 0.001$. Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation. Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.

(b.) Glycemic control (normalized by individual basal level).

	6-month intervention/Before-intervention		3-month after intervention/Before-intervention	
	Probiotic group	Placebo group	Probiotic group	Placebo group
Glucose AC (%)	87.5 ± 10.3***	101.5 ± 11.7	89.0 ± 14.4**	103.5 ± 18.3
HbA1c (%)	91.8 ± 5.4***	100.1 ± 7.8	96.1 ± 6.3*	100.7 ± 8.9

*Mann–Whitney U test: when patients in the probiotic group were compared with those in the placebo group, statistical difference was shown as *P < 0.05, **P < 0.01 and ***P < 0.001. Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation. Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.*

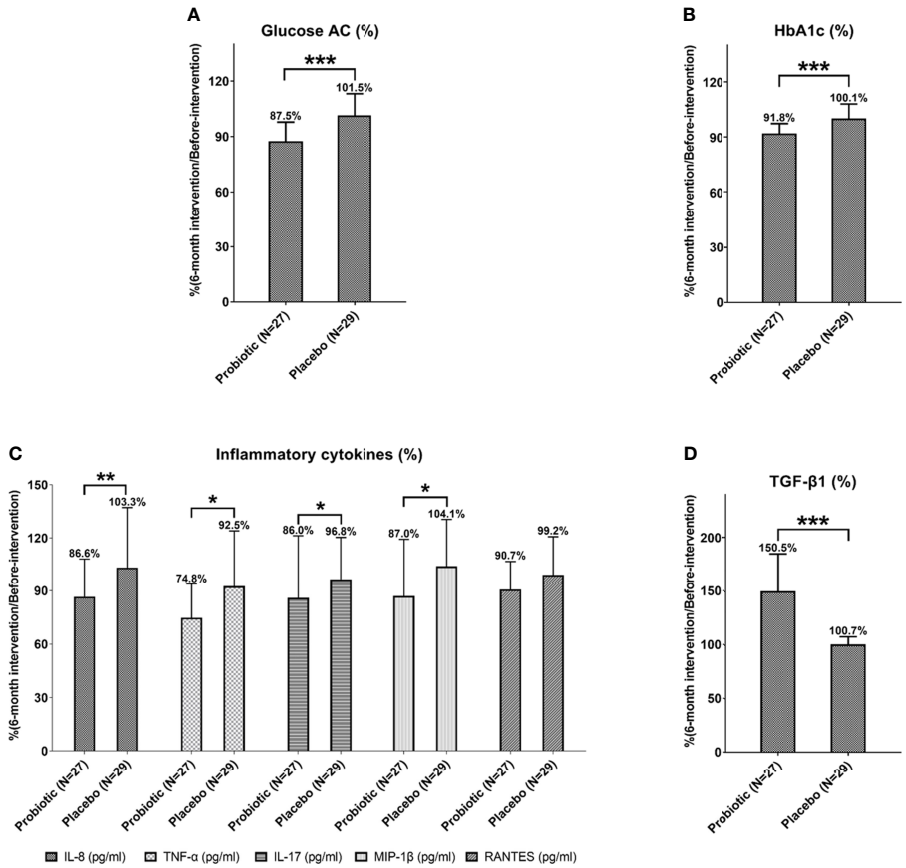


FIGURE 2 | The levels of Glucose AC **(A)**, HbA1c **(B)**, inflammatory cytokines **(C)** and anti-inflammatory cytokine TGF-β1 **(D)** of patients with T1DM after 6-month probiotic intervention were normalized with before-intervention levels and then compared with the placebo group. Statistical differences between probiotic and placebo groups were shown as $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation. Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.

$P = 0.004$), TNF-α (45.7 pg/mL, $P = 0.003$), and MIP-1β (98.1 pg/mL, $P = 0.004$) levels were still significantly lower than the before-intervention levels (**Table 3**).

Next, we compared the changes in inflammatory factors between the probiotic group and placebo group. Data were all normalized to the baseline level (before-intervention). The levels of IL-8 (86.6%, $P = 0.004$, TNF-α (74.8%, $P = 0.018$), IL-17 (86.0%, $P = 0.017$), and MIP-1β (87.0%, $P = 0.029$) significantly decreased after 6 months of probiotic intervention (**Figure 2C** and **Table 3**). Three months after stopping intervention, the IL-8 (89.6%, $P = 0.018$) and MIP-1β (85.7%, $P = 0.016$) levels were still significantly lower in the probiotic group than in the placebo group (**Table 3**). Overall, probiotic intervention lowered inflammation levels among patients with T1DM (**Supplemental Figure S2**).

TGF-β1 Levels Were Elevated After Taking Probiotics

After taking probiotics for 6 months, the level of anti-inflammatory cytokine TGF-β1 significantly increased compared with the before-intervention level (6-month after intervention: 8034.4 pg/mL; before-intervention 5624.2 pg/mL; $P = 0.000$). Three months after stopping the intervention, the TGF-β1 levels (7854.5 pg/mL, $P = 0.000$) were still significantly higher in the probiotic group (**Table 3**). Next, we compared the changes of anti-inflammatory factors between the groups. All data were normalized to the baseline levels (before-intervention). The average normalized TGF-β1 in the probiotic group increased to 150.5%, which was significantly different from the level in the placebo group ($P = 0.000$). The level of TGF-β1 still remained higher (147.0%, $P = 0.000$) than that in the placebo group of 3 months after stopping probiotics (**Figure 2D** and **Table 3**).

Correlation Between Cytokine Levels and Declined ΔHbA1c

ΔHbA1c indicated the HbA1c level after 6-month intervention normalized to the baseline level (before-intervention), ΔHbA1c <1 indicated that the HbA1c level had improved in the intervention period (**Supplemental Figure S3**). In the probiotic group, five inflammatory cytokines (IL-8, TNF-α, IL-17, MIP-1β and RANTES) were down-regulated together with declined ΔHbA1c level after consuming probiotic strains for 6

TABLE 3 | Analysis of inflammatory and anti-inflammatory cytokines after probiotic intervention.

(a.) Inflammatory factors

	Before intervention		6-month intervention		3-month after intervention	
	Probiotic group	**Placebo group**	**Probiotic group**	**Placebo group**	**Probiotic group**	**Placebo group**
IL-8 (pg/mL)	398.2 ± 233.4	477.7 ± 222.9	350.4 ± 225.1$^{\#\#\#}$	467.7 ± 225.3	359.7 ± 233.2$^{\#\#}$	468.6 ± 211.4
TGF-β1 (pg/mL)	5624.2 ± 1984.6	6554.3 ± 1759.3	8034.4 ± 2047.4$^{\#\#\#}$	6575.2 ± 1739.8	7854.5 ± 2030.7$^{\#\#\#}$	6482.8 ± 1665.6
TNF-α (pg/mL)	52.5 ± 59.7	54.0 ± 56.8	39.2 ± 47.5$^{\#\#\#}$	49.3 ± 57.2	45.7 ± 55.7$^{\#\#}$	54.7 ± 59.7
IL-17 (pg/mL)	26.8 ± 34.1	19.5 ± 24.5	19.7 ± 21.3$^{\#\#}$	19.1 ± 24.3	21.6 ± 22.5	18.7 ± 24.5
MIP-1β (pg/mL)	113.9 ± 75.5	130.5 ± 86.9	99.6 ± 76.2$^{\#\#}$	132.3 ± 85.4	98.1 ± 73.7$^{\#\#}$	126.8 ± 78.8
RANTES (pg/mL)	447.9 ± 71.4	411.0 ± 69.0	401.7 ± 75.2$^{\#\#}$	397.2 ± 60.5	428.0 ± 82.0	418.2 ± 54.0

Wilcoxon signed-rank test: when patients in each group were compared with their before-intervention status, statistical difference was shown as $^{\#\#}P < 0.01$ and $^{\#\#\#}P < 0.001$. Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation. Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.

(b.) Inflammatory factors (normalized to baseline levels).

	6-month intervention		3-month after intervention	
	Probiotic group	**Placebo group**	**Probiotic group**	**Placebo group**
IL-8 (%)	86.6 ± 21.5**	103.3 ± 33.8	89.6 ± 19.3*	107.0 ± 45.2
TGF-β1 (%)	150.5 ± 34.1***	100.7 ± 7.3	147.0 ± 33.8***	99.4 ± 4.5
TNF-α (%)	74.8 ± 19.1*	92.5 ± 31.6	87.2 ± 21.5	102.4 ± 35.8
IL-17 (%)	86.0 ± 35.3*	96.8 ± 23.6	100.8 ± 49.0	96.4 ± 27.5
MIP-1β (%)	87.0 ± 32.3*	104.1 ± 26.4	85.7 ± 25.4*	101.8 ± 23.2
RANTES (%)	90.7 ± 16.1	99.2 ± 21.5	96.2 ± 15.6	104.9 ± 23.0

Mann–Whitney U test: when patients in the probiotic group were compared with those in the placebo group, statistical difference was shown as $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. Placebo group: T1DM patients receive regular insulin treatment but no additional probiotic supplementation. Probiotic group: T1DM patients receive regular insulin treatment with additional probiotic supplementation.

months. The anti-inflammatory cytokines (TGF-β1) was up-regulated together with declined ΔHbA1c level after consuming probiotic strains for 6 months.

In the placebo group, 47.4% (N = 9) of the patients showed improved IL-8 levels along with improved ΔHbA1c levels, 71.4% (N = 10) of the patients showed improved TGF-β1 levels together with improved ΔHbA1c levels, 38.9% (N = 7) of the patients showed improved TNF-α levels together with improved ΔHbA1c levels, 47.1% (N = 8) of the patients showed improved IL-17 levels together with improved ΔHbA1c levels, 50.0% (N = 7) of the patients showed improved MIP-1β levels, and 56.3% (N = 9) of the patients showed improved RANTES levels. This result indicated that consumption of probiotics could help improve the inflammatory response and reduce excessive HbA1c in patients with T1DM (**Supplemental Figure S3**). In addition, the majority of patients with T1DM showed improved ΔGlucose AC along with reduced inflammatory cytokine levels after probiotic intervention (**Supplemental Figure S4**).

DISCUSSION

Glycemic management is a critical issue in diabetes (30). In addition to medical interventions (31), studies have investigated the role of probiotics in regulating hyperglycemia in diabetes mellitus (32). Several studies have reported about the effects of probiotics on downregulation of glycemia in patients with T2DM; Asemi et al. discovered that multispecies probiotics, namely *L. acidophilus, L. rhamnosus*, *L. casei*, *L. bulgaricus*, *B. longum*, and *Streptococcus thermophilus* could prevent the rise of fasting plasma glucose (FPG), reduce serum hs-C-reactive protein, and elevate plasma total glutathione in T2DM (33). Ejtahed et al. revealed that probiotic yogurt containing *L. acidophilus* and *B. lactis* could regulate FPG, hemoglobin A1c (HbA1c), and antioxidant levels in patients with T2DM (34). In addition, symbiotic bread containing viable and heat-treated *L. sporogenes* helped reduce serum triacylglycerol and very-low-density lipoprotein levels, while improving high-density lipoprotein levels in patients with T2DM (35). However, few studies have reported on their effects on T1DM. Whether probiotic supplementation altered gut microbiota profile in T1DM is still unclear. At present study, it is suggested that probiotics plus regular insulin treatment would present steadier control of Glucose AC level and HbA1c than insulin treatment only (placebo group) (**Table 2**).

In this study, NGS demonstrated that consumption of probiotic *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9 could successfully enrich the population of these gut microbiota, especially in the species of *Bifidobacterium animalis, Lactobacillus salivarius* and *Akkermansia muciniphila* (**Figure 1**). The colonized beneficial microbiota may play an important role in slowing down the levels of Glucose AC and HbA1c in patients with T1DM (**Table 2**). The Oral glucose tolerance test (OGTT) is a reliable method for measuring glycemic levels (36). However, the WHO mentioned several shortcomings of OGTT test in clinical practice including

inconvenience, greater cost and less reproducibility (37). Thus, the HbA1c and Glucose AC levels were selected as glycemic biomarkers of practical clinical practice at this study.

Furthermore, abundance of *Akkermansia muciniphila* has been reported to negatively correlate with HbA1c levels (38). Studies have revealed that obese children (above 3 years of age) had lower levels of *A. muciniphila* in their gut microbiota (39). Moreover, *A. muciniphila* can restore the gut barrier at the mucous layer, facilitate lipid metabolism, reduce fat mass accumulation, and reduce the incidence of fatty liver and hypercholesterol (40). This may account for the higher distribution of *A. muciniphila* among semi-supercentenarians (41). In our study, patients with T1DM who received the novel probiotic product slowed down the declined rate of *Akkermansia* uncultured bacterium in their gut, compared with patients who received the placebo. The molecular mechanism underlying the modulation of glycemic levels by the colonized microbiota warrants further investigation.

Previous clinical studies have discovered the relationship between human immunity and probiotic consumption. Groele et al. found that *L. rhamnosus* GG and *B. lactis* Bb12 had several advantages in T1DM, including improvement of the gut mucosal barrier, regulation of local and systemic immune responses, reduction of the risk of autoimmunity, restriction of the growth of pathogens, and preservation of β-cell function (42). Another clinical study involving 42 healthy individuals reported that *L. johnsonii* N6.2 could have beneficial effects on T1DM by regulating the kynurenine:tryptophan ratio, indoleamine 2,3-dioxygenase pathway, circulatory effector Th1 cells, and cytotoxic $CD8^+$ T cells (43). While previous researches might have explored the benefits of probiotics to T1DM, this is a pioneer study to include post-experimental follow-ups, where the persistence effect of immunomodulation by proinflammatory cytokine diminution and anti-inflammatory cytokine accretion was observed even 3 months after discontinuation of the probiotics.

Gut microbiome and their metabolites may exert function by affecting intestinal permeability, molecular mimicry, and modulating immune system (44). In this study, we revealed the ability of *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9 to downregulate the immune-related inflammatory cytokines IL-8, TNF-α, IL-17, MIP-1β, and RANTES and upregulate the anti-inflammatory cytokine TGF-β1 (**Table 3** and **Supplemental Figure S2**). Cytokines and chemokines have previously been implicated in the pathogenesis of T1DM (45), including IL-8 (46), RANTES (47), MIP-1β (48), TNF-α (49), and IL-17 (50). On the other hand, TGF-β secreted by regulatory T cells (Tregs) can suppress inflammation (51). Tregs help to prevent the development of T1DM; they delay the onset of T1DM by secretion of TGF-β (52). Next, the correlation between blood sugar levels, cytokine levels and the change of intestinal flora after treating probiotics were analyzed by Spearman rank-order correlation test. The results suggested that *Actinobacteria, Firmicutes, Bifidobacterium bifidum, Bacteoid fragilis*, and *Lactobacillus salivarius* showed the negative correlation with inflammatory cytokine TNF-α, IL-8, RANTES, MIP-1β, and Glucose AC, respectively. A negative correlation means that as one variable increases, the other tends to decline. Besides, *Lactobacillus salivarius* presented positive correlation with the anti-inflammatory cytokine TGF-β1 (**Supplemental Figure S5**).

Our results also demonstrated the correlation between improved cytokine levels and declined glycemic rate among patients with T1DM after probiotic intervention (**Supplemental Figures S3**, **S4**). Previous studies discovered that elevation in the concentrations of inflammatory cytokines such as TNF-α associated with glycemic control and cardiovascular risk factors among patients with T1DM (53). The results of this study suggested that probiotics may play an important role in reducing immune inflammatory cytokines and glycemic control in patients with T1DM. However, the mechanism by which the elevated inflammatory cytokines stimulate blood glucose (through which glucose transporters or insulin receptors) among patients with T1DM warrants further investigation.

The limitation of our study lies in the lack of knowledge to how the probiotics modulate the immunity to improve blood sugar. Firstly, while a correlation between our probiotics and improved blood sugar was observed, more work might be required to elucidate the underlying mechanism in order to explain for the causality. Secondly, as the current study utilizes a multi-strain supplement to make use of the synergic effect, more studies might be needed to clarify the individual effects of the strains contained in our probiotics. Thirdly, changes of c-peptide (54) and beta-cells (55) levels by probiotic intervention should be detected. Finally, the NGS sequencing method used in the article was to identify gut microbial bacteria by sequencing the V3-V4 region of 16S rRNA (56) with the annotations of SILVA database. The sequencing method still has some limitations in microbial classification. Therefore, the qPCR method with specific primers should be performed for further validating microbial taxonomy in the future.

CONCLUSION

The results of this study showed that conventional insulin treatment plus mixed probiotics strains of *L. salivarius* subsp. *salicinius* AP-32, *L. johnsonii* MH-68, and *B. animalis* subsp. *lactis* CP-9 showed better treatment outcome than insulin injection only (the placebo group). It's suggested that insulin treatment plus probiotic supplementation could enrich the population of beneficial gut microbiota in the gut (*Bifidobacterium animalis, Lactobacillus salivarius* and *Akkermansia muciniphila*), improve the glycemic control (Glucose AC and HbA1c), reduce the levels of inflammation-related cytokines (IL-8, TNF-α, IL-17, MIP-1β, and RANTES) and increase the levels of the anti-inflammatory cytokine TGF-β1 in patients with T1DM. The results of this study render a prospective therapeutic option for clinical T1DM treatment.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found below: "NCBI with Bioproject ID PRJNA798680 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA798680/).

ETHICS STATEMENT

This clinical trial complied with the Declaration of Helsinki, and it was reviewed and approved by China Medical University & Hospital Research Ethics Committee (CMUH107-REC2-036).

AUTHOR CONTRIBUTIONS

C-HW: conceptualization, reviewing, and editing. H-RY: reviewing and editing. W-LL: reviewing and editing. H-HH: supervision and project design. W-YL: original drafting. Y-WK: methodology and data analysis. Y-YH: methodology, data visualization, and project administration. S-YT: reviewing and editing. H-CL: supervision, reviewing, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from China Medical University Hospital (DMR-108-048, DMR-109-052 and DMR-110-209) and Asia University Hospital (ASIA-108-51010).

ACKNOWLEDGMENTS

Giving thanks to T1DM patients who participated this clinical trial as well as lab members of the research and design department of Bioflag Biotech Co., Ltd. (Tainan, Taiwan) who assisted to collect and analyze experimental data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.754401/full#supplementary-material

REFERENCES

1. Rosen CJ, Ingelfinger JR. Traveling Down the Long Road to Type 1 Diabetes Mellitus Prevention. *N Engl J Med* (2019) 381(7):666–7. doi: 10.1056/NEJMe1907458
2. Lacy PE. The Pancreatic Beta Cell. Structure and Function. *N Engl J Med* (1967) 276(4):187–95. doi: 10.1056/NEJM196701262760401
3. Kalyva E, Malakonaki E, Eiser C, Mamoulakis D. Health-Related Quality of Life (HRQoL) of Children With Type 1 Diabetes Mellitus (T1DM): Self and Parental Perceptions. *Pediatr Diabetes* (2011) 12(1):34–40. doi: 10.1111/j.1399-5448.2010.00653.x
4. Dokken BB. The Pathophysiology of Cardiovascular Disease and Diabetes: Beyond Blood Pressure and Lipids. *Diabetes Spectr* (2008) 21(3):160–5. doi: 10.2337/diaspect.21.3.160%JDiabetesSpectrum
5. Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of Type 1 Diabetes. *Endocrinol Metab Clin North Am* (2010) 39(3):481–97. doi: 10.1016/j.ecl.2010.05.011
6. You WP, Henneberg M. Type 1 Diabetes Prevalence Increasing Globally and Regionally: The Role of Natural Selection and Life Expectancy at Birth. *BMJ Open Diabetes Res Care* (2016) 4(1):e000161. doi: 10.1136/bmjdrc-2015-000161
7. Buzhor E, Leshansky L, Blumenthal J, Barash H, Warshawsky D, Mazor Y, et al. Cell-Based Therapy Approaches: The Hope for Incurable Diseases. *Regener Med* (2014) 9(5):649–72. doi: 10.2217/rme.14.35
8. Primavera M, Giannini C, Chiarelli F. Prediction and Prevention of Type 1 Diabetes. *Front Endocrinol (Lausanne)* (2020) 11:248:248. doi: 10.3389/fendo.2020.00248
9. Rewers M, Ludvigsson J. Environmental Risk Factors for Type 1 Diabetes. *Lancet* (2016) 387(10035):2340–8. doi: 10.1016/S0140-6736(16)30507-4
10. Korf H, Gysemans C, Overbergh L, Mathieu C. J. I. D. M. Pathogenesis of Type 1 Diabetes Mellitus: Immunologic Pathways. *International Diabetes Monitor* (2010) 22:121–7.
11. Kaur N, Bhadada SK, Minz RW, Dayal D, Kochhar R. Interplay Between Type 1 Diabetes Mellitus and Celiac Disease: Implications in Treatment. *Dig Dis* (2018) 36(6):399–408. doi: 10.1159/000488670
12. Makharia GK. Celiac Disease Screening in Southern and East Asia. *Dig Dis* (2015) 33(2):167–74. doi: 10.1159/000369537
13. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Nageshwar Reddy D. Role of the Normal Gut Microbiota. *World J Gastroenterol* (2015) 21(29):8787–803. doi: 10.3748/wjg.v21.i29.8787
14. Han H, Li Y, Fang J, Liu G, Yin J, Li T, et al. Gut Microbiota and Type 1 Diabetes. *Int J Mol Sci* (2018) 19(4):995. doi: 10.3390/ijms19040995
15. Knip M, Honkanen J. Modulation of Type 1 Diabetes Risk by the Intestinal Microbiome. *Curr Diabetes Rep* (2017) 17(11):105. doi: 10.1007/s11892-017-0933-9
16. Drexhage HA, Dik WA, Leenen PJ, Versnel MA. The Immune Pathogenesis of Type 1 Diabetes: Not Only Thinking Outside the Cell But Also Outside the Islet and Out of the Box. *Diabetes* (2016) 65(8):2130–3. doi: 10.2337/dbi16-0030
17. Mishra SP, Wang S, Nagpal R, Miller B, Singh R, Taraphder S, et al. Probiotics and Prebiotics for the Amelioration of Type 1 Diabetes: Present and Future Perspectives. *Microorganisms* (2019) 7(3):67. doi: 10.3390/microorganisms7030067
18. de Oliveira GLV, Leite AZ, Higuchi BS, Gonzaga MI, Mariano VS. Intestinal Dysbiosis and Probiotic Applications in Autoimmune Diseases. *Immunology* (2017) 152(1):1–12. doi: 10.1111/imm.12765
19. Shi G, Sun C, Gu W, Yang M, Zhang X, Zhai N, et al. Free Fatty Acid Receptor 2, a Candidate Target for Type 1 Diabetes, Induces Cell Apoptosis Through ERK Signaling. *J Mol Endocrinol* (2014) 53(3):367–80. doi: 10.1530/JME-14-0065
20. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. Expert Consensus Document. The International Scientific Association for Probiotics and Prebiotics Consensus Statement on the Scope and Appropriate Use of the Term Probiotic. *Nat Rev Gastroenterol Hepatol* (2014) 11(8):506–14. doi: 10.1038/nrgastro.2014.66
21. Hsieh PS, Ho HH, Hsieh SH, Kuo YW, Tseng HY, Kao HF, et al. Lactobacillus Salivarius AP-32 and Lactobacillus Reuteri GL-104 Decrease Glycemic Levels and Attenuate Diabetes-Mediated Liver and Kidney Injury in Db/Db Mice. *BMJ Open Diabetes Res Care* (2020) 8(1):e001028. doi: 10.1136/bmjdrc-2019-001028
22. Hsieh PS, Tsai YC, Chen YC, Teh SF, Ou CM, King VA. Eradication of Helicobacter Pylori Infection by the Probiotic Strains Lactobacillus Johnsonii MH-68 and L. Salivarius Ssp. Salicinius AP-32. *Helicobacter* (2012) 17(6):466–77. doi: 10.1111/j.1523-5378.2012.00992.x
23. A. American Diabetes. Standards of Medical Care in Diabetes–2010. *Diabetes Care* (2010) 33 Suppl 1:S11–61. doi: 10.2337/dc10-S011
24. Kawasaki E, Maruyama T, Imagawa A, Awata T, Ikegami H, Uchigata Y, et al. Diagnostic Criteria for Acute-Onset Type 1 Diabetes Mellitus (2012): Report of the Committee of Japan Diabetes Society on the Research of Fulminant and Acute-Onset Type 1 Diabetes Mellitus. *J Diabetes Investig* (2014) 5(1):115–8. doi: 10.1111/jdi.12119

25. 8 Pharmacologic Approaches to Glycemic Treatment. Standards of Medical Care in Diabetes—2018. *Diabetes Care* (2018) 41(Supplement 1):S73–85. doi: 10.2337/dc18-S008%JDiabetesCare
26. American Diabetes Association. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2021. *Diabetes Care* (2021) 44(Suppl 1):S73–84. doi: 10.2337/dc21-S006
27. Ho J, Reimer RA, Doulla M, Huang C. Effect of Prebiotic Intake on Gut Microbiota, Intestinal Permeability and Glycemic Control in Children With Type 1 Diabetes: Study Protocol for a Randomized Controlled Trial. *Trials* (2016) 17(1):347. doi: 10.1186/s13063-016-1486-y
28. UK Prospective Diabetes Study (UKPDS) Group. Intensive Blood-Glucose Control With Sulphonylureas or Insulin Compared With Conventional Treatment and Risk of Complications in Patients With Type 2 Diabetes (UKPDS 33). *Lancet* (1998) 352(9131):837–53. doi: 10.1016/S0140-6736(98)07019-6
29. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of Glycaemia With Macrovascular and Microvascular Complications of Type 2 Diabetes (UKPDS 35): Prospective Observational Study. *BMJ* (2000) 321(7258):405–12. doi: 10.1136/bmj.321.7258.405
30. Diabetes Canada Clinical Practice Guidelines Expert C, McGibbon A, Adams L, Ingersoll K, Kader T, Tugwell B. Glycemic Management in Adults With Type 1 Diabetes. *Can J Diabetes* (2018) 42(Suppl 1):S80–7. doi: 10.1016/j.jcjd.2017.10.012
31. Atkinson MA, Eisenbarth GS. Type 1 Diabetes: New Perspectives on Disease Pathogenesis and Treatment. *Lancet* (2001) 358(9277):221–9. doi: 10.1016/S0140-6736(01)05415-0
32. Ruan Y, Sun J, He J, Chen F, Chen R, Chen H. Effect of Probiotics on Glycemic Control: A Systematic Review and Meta-Analysis of Randomized, Controlled Trials. *PloS One* (2015) 10(7):e0132121. doi: 10.1371/journal.pone.0132121
33. Asemi Z, Zare Z, Shakeri H, Sabihi SS, Esmaillzadeh A. Effect of Multispecies Probiotic Supplements on Metabolic Profiles, Hs-CRP, and Oxidative Stress in Patients With Type 2 Diabetes. *Ann Nutr Metab* (2013) 63(1-2):1–9. doi: 10.1159/000349922
34. Ejtahed HS, Mohtadi-Nia J, Homayouni-Rad A, Niafar M, Asghari-Jafarabadi M, Mofid V. Probiotic Yogurt Improves Antioxidant Status in Type 2 Diabetic Patients. *Nutrition* (2012) 28(5):539–43. doi: 10.1016/j.nut.2011.08.013
35. Shakeri H, Hadaegh H, Abedi F, Tajabadi-Ebrahimi M, Mazroii N, Ghandi Y, et al. Consumption of Synbiotic Bread Decreases Triacylglycerol and VLDL Levels While Increasing HDL Levels in Serum From Patients With Type-2 Diabetes. *Lipids* (2014) 49(7):695–701. doi: 10.1007/s11745-014-3901-z
36. Ito C, Mito K, Hara H. Review of Criteria for Diagnosis of Diabetes Mellitus Based on Results of Follow-Up Study. *Diabetes* (1983) 32(4):343–51. doi: 10.2337/diab.32.4.343
37. World Health Organization. *Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia: Report of a WHO/IDF Consultation*. Geneva: WHO Press (2006).
38. Shih CT, Yeh YT, Lin CC, Yang LY, Chiang CP. Akkermansia Muciniphila is Negatively Correlated With Hemoglobin A1c in Refractory Diabetes. *Microorganisms* (2020) 8(9):1360. doi: 10.3390/microorganisms8091360
39. Karvonen AM, Sordillo JE, Gold DR, Bacharier LB, O'Connor GT, Zeiger RS, et al. Gut Microbiota and Overweight in 3-Year Old Children. *Int J Obes (Lond)* (2019) 43(4):713–23. doi: 10.1038/s41366-018-0290-z
40. Xu Y, Wang N, Tan HY, Li S, Zhang C, Feng Y. Function of Akkermansia Muciniphila in Obesity: Interactions With Lipid Metabolism, Immune Response and Gut Systems. *Front Microbiol* (2020) 11:219. doi: 10.3389/fmicb.2020.00219
41. Biagi E, Franceschi C, Rampelli S, Severgnini M, Ostan R, Turroni S, et al. Gut Microbiota and Extreme Longevity. *Curr Biol* (2016) 26(11):1480–5. doi: 10.1016/j.cub.2016.04.016
42. Groele L, Szajewska H, Szypowska A. Effects of Lactobacillus Rhamnosus GG and Bifidobacterium Lactis Bb12 on Beta-Cell Function in Children With Newly Diagnosed Type 1 Diabetes: Protocol of a Randomised Controlled Trial. *BMJ Open* (2017) 7(10):e017178. doi: 10.1136/bmjopen-2017-017178
43. Marcial GE, Ford AL, Haller MJ, Gezan SA, Harrison NA, Cai D, et al. Lactobacillus Johnsonii N6.2 Modulates the Host Immune Responses: A Double-Blind, Randomized Trial in Healthy Adults. *Front Immunol* (2017) 8:655. doi: 10.3389/fimmu.2017.00655
44. Zheng P, Li Z, Zhou Z. Gut Microbiome in Type 1 Diabetes: A Comprehensive Review. *Diabetes Metab Res Rev* (2018) 34(7):e3043. doi: 10.1002/dmrr.3043
45. Ahmadi Z, Arababadi MK, Hassanshahi G. CXCL10 Activities, Biological Structure, and Source Along With its Significant Role Played in Pathophysiology of Type I Diabetes Mellitus. *Inflammation* (2013) 36(2):364–71. doi: 10.1007/s10753-012-9555-1
46. Wuyts A, Proost P, Lenaerts JP, Ben-Baruch A, Van Damme J, Wang JM. Differential Usage of the CXC Chemokine Receptors 1 and 2 by Interleukin-8, Granulocyte Chemotactic Protein-2 and Epithelial-Cell-Derived Neutrophil Attractant-78. *Eur J Biochem* (1998) 255(1):67–73. doi: 10.1046/j.1432-1327.1998.2550067.x
47. Schall TJ. Biology of the RANTES/SIS Cytokine Family. *Cytokine* (1991) 3(3):165–83. doi: 10.1016/1043-4666(91)90013-4
48. Meagher C, Arreaza G, Peters A, Strathdee CA, Gilbert PA, Mi QS, et al. CCL4 Protects From Type 1 Diabetes by Altering Islet Beta-Cell-Targeted Inflammatory Responses. *Diabetes* (2007) 56(3):809–17. doi: 10.2337/db06-0619
49. Chatzigeorgiou A, Harokopos V, Mylona-Karagianni C, Tsouvalas E, Aidinis V, Kamper EF. The Pattern of Inflammatory/Anti-Inflammatory Cytokines and Chemokines in Type 1 Diabetic Patients Over Time. *Ann Med* (2010) 42(6):426–38. doi: 10.3109/07853890.2010.495951
50. Rouvier E, Luciani MF, Mattei MG, Denizot F, Golstein P. CTLA-8, Cloned From an Activated T Cell, Bearing AU-Rich Messenger RNA Instability Sequences, and Homologous to a Herpesvirus Saimiri Gene. *J Immunol* (1993) 150(12):5445–56.
51. Wan YY, Flavell RA. The Roles for Cytokines in the Generation and Maintenance of Regulatory T Cells. *Immunol Rev* (2006) 212:114–30. doi: 10.1111/j.0105-2896.2006.00407.x
52. You S, Thieblemont N, Alyanakian MA, Bach JF, Chatenoud L. Transforming Growth Factor-Beta and T-Cell-Mediated Immunoregulation in the Control of Autoimmune Diabetes. *Immunol Rev* (2006) 212:185–202. doi: 10.1111/j.0105-2896.2006.00410.x
53. Esposito K, Nappo F, Marfella R, Giugliano G, Giugliano F, Ciotola M, et al. Inflammatory Cytokine Concentrations are Acutely Increased by Hyperglycemia in Humans: Role of Oxidative Stress. *Circulation* (2002) 106(16):2067–72. doi: 10.1161/01.cir.0000034509.14906.ae
54. Palmer JP, Fleming GA, Greenbaum CJ, Herold KC, Jansa LD, Kolb H, et al. C-Peptide is the Appropriate Outcome Measure for Type 1 Diabetes Clinical Trials to Preserve Beta-Cell Function: Report of an ADA Workshop, 21-22 October 2001. *Diabetes* (2004) 53(1):250–64. doi: 10.2337/diabetes.53.1.250
55. Akirav EM, Lebastchi J, Galvan EM, Henegariu O, Akirav M, Ablamunits V, et al. Detection of Beta Cell Death in Diabetes Using Differentially Methylated Circulating DNA. *Proc Natl Acad Sci USA* (2011) 108(47):19018–23. doi: 10.1073/pnas.1111008108
56. Ballan R, Saad SMI. Characteristics of the Gut Microbiota and Potential Effects of Probiotic Supplements in Individuals With Type 2 Diabetes Mellitus. *Foods* (2021) 10(11):2528. doi: 10.3390/foods10112528

Conflict of Interest: Author H-HH, W-YL, Y-WK, Y-YH, and S-YT were employed by Bioflag Biotech Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fendo.2022.836023

Shared Genetic Basis and Causal Relationship Between Television Watching, Breakfast Skipping and Type 2 Diabetes: Evidence From a Comprehensive Genetic Analysis

Dongze Chen[1], *Hanyu Wu*[2], *Xinpei Wang*[1], *Tao Huang*[3,4*] *and Jinzhu Jia*[1,5*]

[1] Department of Biostatistics, School of Public Health, Peking University, Beijing, China, [2] Department of Bioinformatics, School of Life Science, Peking University, Beijing, China, [3] Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China, [4] Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing, China, [5] Center for Statistical Science, Peking University, Beijing, China

OPEN ACCESS

Edited by:
Jie Li,
Guangdong Provincial People's Hospital, China

Reviewed by:
Alphonse Thanaraj Thangavel, Dasman Diabetes Institute, Kuwait
Ghislain Rocheleau, Icahn School of Medicine at Mount Sinai, United States

***Correspondence:**
Tao Huang
huangtaotao@pku.edu.cn
Jinzhu Jia
jzjia@math.pku.edu.cn

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: *15 December 2021*
Accepted: *01 March 2022*
Published: *24 March 2022*

Citation:
Chen D, Wu H, Wang X, Huang T and Jia J (2022) Shared Genetic Basis and Causal Relationship Between Television Watching, Breakfast Skipping and Type 2 Diabetes: Evidence From a Comprehensive Genetic Analysis. Front. Endocrinol. 13:836023. doi: 10.3389/fendo.2022.836023

Background: Epidemiological investigations have established unhealthy lifestyles, such as excessive leisurely sedentary behavior (especially TV/television watching) and breakfast skipping, increase the risk of type 2 diabetes (T2D), but the causal relationship is unclear. We aimed to understand how single nucleotide variants contribute to the co-occurrence of unhealthy lifestyles and T2D, thereby providing meaningful insights into disease mechanisms.

Methods: Combining summary statistics from genome-wide association studies (GWAS) on TV watching (N = 422218), breakfast skipping (N = 193860) and T2D (N = 159208) in European pedigrees, we conducted comprehensive pairwise genetic analysis, including high-definition likelihood (HDL-method), cross-phenotype association studies (CPASSOC), GWAS-eQTL colocalization analysis and transcriptome-wide association studies (TWAS), to understand the genetic overlap between them. We also performed bidirectional two-sample Mendelian randomization (MR) analysis for causal inference using genetic instrumental variables, and two-step MR mediation analysis was used to assess any effects explained by body mass index, lipid traits and glycemic traits.

Results: HDL-method showed that T2D shared a strong genetic correlation with TV watching (r_g = 0.26; P = 1.63×10^{-29}) and skipping breakfast (r_g = 0.15; P =2.02×10^{-6}). CPASSOC identifies eight independent SNPs shared between T2D and TV watching, including one novel shared locus. TWAS and CPASSOC showed that shared genes were enriched in lung, esophageal, adipose, and thyroid tissues and highlighted potential shared regulatory pathways for lipoprotein metabolism, pancreatic β-cell function, cellular senescence and multi-mediator factors. MR showed TV watching had a causal effect on T2D (β_{IVW} = 0.629, P_{IVW} = 1.80×10^{-10}), but no significant results were observed between breakfast skipping and T2D. Mediation analysis provided evidence that body mass index,

fasting glucose, hemoglobin A1c and high-density lipoprotein are potential factors that mediate the causal relationship between TV and T2D.

Conclusions: Our findings provide strong evidence of shared genetics and causation between TV watching and T2D and facilitate our identification of common genetic architectures shared between them.

Keywords: TV watching, breakfast skipping, type 2 diabetes, Mendelian randomization, genome genetic correlation

HIGHLIGHTS

- The strongest positive genetic correlation was observed between TV watching and type 2 diabetes.
- Cross-trait meta-analysis identifies eight independent genomic loci shared between type 2 diabetes and television watching, one of which is novel.
- Implicated genes suggest potential treatment targets and signaling pathways for type 2 diabetes and television watching.
- Transcriptome-wide association studies and cross-trait meta-analysis support the role of lipoprotein metabolism, cellular senescence and multi-mediator factors may account for the shared metabolic pathway and causes between TV watching and T2D.
- Mendelian randomization study showed TV watching had strong causal effect on T2D (β_{IVW} = 0.629, P_{IVW} = 1.80×10^{-10}).

INTRODUCTION

Type 2 diabetes (T2D) is a global epidemic that affects more than 463 million people and is a leading cause of morbidity and mortality worldwide. Family-based studies have shown that T2D is highly heritable, with an estimated heritability range of 20%-80% (1, 2). Currently, worldwide prevalent unhealthy lifestyles (especially TV watching and breakfast skipping) are also considered to be the key contributors to T2D. However, whether such an unhealthy lifestyle is causally associated or shares a genetic basis with T2D remains largely unknown.

A growing body of evidence from observational studies suggests that the risk of T2D is positively associated with prolonged TV watching (3–6) and breakfast skipping (7–9). A prospective study showed that TV watching is always related to higher energy intake than expenditure and leads to higher BMI (10), which affects metabolism by releasing non-esterified fatty acids (NEFAs) (11). Increasing plasma NEFA levels then leads to inadequate insulin secretion and insulin resistance (low insulin sensitivity), together contributing to the development of T2D (11). The association between breakfast skipping and T2D is also reported to be partially mediated by body mass index (BMI) (9). Furthermore, breakfast skippers are more likely to have lower serum HDL cholesterol levels (12), which is widely confirmed to be associated with an increased risk of T2D in Mendelian randomization studies (13). Therefore, we hypothesized that a common genetic etiology and the mediating role of BMI or HDL may at least partially explain the association between T2D and TV watching and breakfast skipping.

Evidence from observational studies is limited for making causal inferences, as such associations may be due to (residual) confounding and/or reverse causality (14). Considering that genetics is unlikely to be influenced by these factors, it is informative to use genetic variants as instrumental variables to investigate the causal relationships behind these associations. To date, genome-wide association studies (GWAS) have been able to detect 145, 128 and 6 genome-wide significant independent SNP signals for T2D, TV watching and breakfast skipping, respectively. Many of the significant loci for TV watching are also susceptibility loci for T2D, suggesting a possible common genetic etiology between them (15–17). Meanwhile, a growing number of Mendelian randomization studies based on strong instrumental variables (IVs) have shown a causal relationship between TV watching and numerous adverse outcomes, such as cerebrovascular diseases (18), coronary artery disease (17), chronic kidney disease (19) and lung cancer (20). However, Mendelian randomization cannot deal with pleiotropy, where genetic variation is associated with multiple traits, since it will break the single pathway hypothesis of MR (21). Research suggests that cross-phenotypic (CP) associations can recognize genetic pleiotropy in human diseases and highlight shared biological pathways compared to single-trait analysis (22). However, little research has been done on CP association analysis between T2D with TV watching and breakfast skipping.

Therefore, to increase our understanding of potential causality and shared genetic architecture between TV watching, breakfast skipping and T2D, we conducted a comprehensive genetic analysis. We performed a bidirectional MR and mediation analysis using summary statistics from public external URL (https://data.mendeley.com/datasets/mxjj6czsrd/1), the Common Metabolic Diseases Knowledge Portal (CMDKP) website (for exposures) and the Diabetes Genetics Replication And Meta-analysis (DIAGRAMv3) Consortium (for type 2 diabetes). To further identify genomic loci shared between T2D and exposures,

Abbreviations: T2D, Type 2 diabetes; BMI, body mass index; SNV, Single nucleotide variant; SNP, Single nucleotide polymorphism; GWAS, genome-wide association study; HDL, High density lipoprotein; HDL-method, High-definition likelihood; LDSC, Linkage disequilibrium score regression; CPASSOC, Cross phenotype association study; eQTL, expression quantitative trait loci; MR, Mendelian randomization; TWAS, Transcriptome wide association study; GTEx, Genotype-tissue expression portal.

we used cross-phenotype association (CPASSOC) analysis and transcriptome-wide association (TWAS) studies to explore shared genetic components among these complex phenotypes.

MATERIALS AND METHODS

Data Source and Study Population

The study was conducted using publicly available GWAS summary data. Details on the study characteristics, participants, and ethics declarations for each dataset can be found in the original publications (16, 17, 23). The hitherto largest GWAS of self-reported TV watching was conducted based on the United Kingdom Biobank (UKB) population cohort (N = 422218) (17). A total of 45.7% of participants were male, with a mean age of 57.4 [standard deviation (SD) 8.0] years at the first assessment of the cohort, and the mean daily reported leisure TV watching was 2.8 h (SD 1.5). The most recent summary results for breakfast skipping were based on a proxy-phenotype (breakfast cereal skipping) GWAS obtained from the Common Metabolic Diseases Knowledge Portal website (16), which included 193860 participants with 24-hour retrospective dietary data from the UKB. We used the T2D GWAS summary statistics from the 2017 report of the DIAGRAMv3 Consortium, consisting of 26676 T2D cases and 132532 control individuals (23). All participants were of European ancestry and had no overlap between exposure (TV watching, breakfast skipping) and outcome (T2D) samples. The location of SNPs is based on the Genome Reference Consortium Human Build 37 (GRCh37).

Genetic Correlation Analysis

The more recent high-definition likelihood (HDL-method) (24) method and conventional cross-trait linkage disequilibrium score (LDSC) regression (25) were conducted to evaluate the genetic correlation ($\boldsymbol{r}_g$) between T2D and TV watching and breakfast skipping. HDL-method extends the LDSC method by modeling the relation between covariances among Z statistics for pairs of traits across multiple SNPs and a full matrix of cross-SNP LD scores. As the HDL-method yields more precise estimates of genetic correlations than LDSC, we chose the HDL-method as the primary result. The HDL-method uses the LD reference computed from 335265 genomic British individuals in the UKB.

Cross Trait Meta-Analysis

Genetic correlation depicts the genome-wide average sharing of genetic effects between traits. To identify genetic variants shared between traits, we applied cross-trait GWAS meta-analysis using the cross-phenotype association (CPASSOC) (26) method to combine the association evidence for TV watching and breakfast skipping with T2D based on the criteria of both $\boldsymbol{r}_g > 10\%$ and $P_{bonferroni} < 0.05$ from HDL-method. CPASSOC combines effect estimates and standard error of GWAS summary statistics to test the hypothesis of association between a SNP and two traits and assumes that effects may exist only within a subset of traits (27). We used the heterogonous version of cross-phenotype association (SHet), which is based on a sample size-weighted, fixed-effect model and is more powerful when there is a heterogonous effect present between studies (26).

We applied PLINK1.9 clumping function (parameters: –clump-p1 2.5e-8 –clump-p2 1e-5 –clump-r2 0.4 –clump-kb 500) to determine index loci that are independent of each other, i.e., variants with P value less than 1×10^{-5} have an r^2 greater than 0.4 and less than 500 kb away from the peak will be assigned to that peak's clump. We identified all genes falling within each clump region. A P value of 2.5×10^{-8} ($5\times10^{-8}/2$) was used as genome-wide significance level for cross-trait meta-analysis to account for 2 meta-analyses. SNPs with a meta-analysis P value less than 2.5×10^{-8} and trait-specific P value less than 1×10^{-5} were selected for downstream analysis.

GWAS-eQTL Coloclization Analysis

To investigate whether the shared index SNPs from CPASSOC and their expression quantitative trait loci (eQTLs) co-localized with candidate causal variants, we performed colocalization analysis, COLOC, which uses Bayesian posterior probability to assess colocalization (28). We extracted cis-eQTL data from the Genotype-Tissue Expression (GTEx) Portal v7 for 48 single tissues (29). The SNP-associated locus was defined as within a 1-Mb window for each of the shared SNPs. The posterior probability H4 hypothesis was calculated to determine whether shared SNPs are associated with two traits. In our study, loci with posterior probability H4 > 0.9 were considered to be co-localized.

Transcriptome-Wide Association Studies

For TV watching, skipping breakfast and T2D, we used transcriptome-wide association studies (TWAS) to identify genes whose cis-regulated gene expression was associated with the corresponding traits. Then, we further evaluated shared tissue-gene pairs between different traits. We performed TWAS analysis using FUSION software and its precomputed transcript expression reference weights, as well as eQTL data from GTEx v.7 (30). Bonferroni correction was applied to determine significant association results after multiple comparisons for all tissue-gene pairs tested for each trait ($\boldsymbol{P}_{Bonferroni} < 0.05$). To increase the significance of the TWAS results, we used the most recent and authoritative summary data for T2D obtained from DIAGRAM. This study was performed in 2018 by Mahajan et al., who mined additional novel T2D susceptibility SNP loci by combining data from 898130 (including UKB sample) individuals of European descent (31).

Mendelian Randomization Analysis

Finally, we implemented a bidirectional MR using TwoSample MR package to test the causal relationship between T2D and unhealthy lifestyles, where the associations for IV-exposure and IV-outcome came from two nonoverlapping groups of participants. Since different MR methods have different degrees of explanation and contexts of application and differ in statistical efficiency, we adopt many MR methods to estimate causal effects. The causal effect estimates from the multiplicative random effects inverse variance weighted (IVW) model were used as the primary result. We conducted a range of sensitivity analyses

using multiplicative random effects inverse variance weighted heterogeneity test, weighted median, MR–Egger regression, MR-Steiger, MR-Robust Adjusted Profile Scores (MR-RAPS), MR-Pleiotropy Residual Sum and Outlier (MR-PRESSO) analysis and leave-one-out cross-validation analysis. The weighted median approach provides consistent and robust estimates even if more than 50% of the IVs are invalid (32). The intercept of MR–Egger regression can be used to evaluate the directional pleiotropy of IVs (33). We applied MR-Steiger to assure that the causal direction between the hypothesized exposure and outcome was correctly assigned (34). Considering the measurement error in SNP exposure effects, MR-RAPS is unbiased when there are many weak instruments and is robust to systematic and idiosyncratic pleiotropy (35). MR-PRESSO and leave-one-out cross-validation analysis are mainly used to detect anomalous IVs (36, 37).

Furthermore, the effect allele frequency reported in the corresponding GWAS was used to detect and exclude all palindromic SNPs to determine the corresponding strand between two GWAS in harmonization section. For trait pairs with significant causal relationships, we searched the GWAS catalog (https://www.ebi.ac.uk/gwas/) to exclude IVs with genome-wide significance for potential confounding traits (e.g., educational attainment, cognitive performance, smoking behavior, alcohol consumption, hypertension, BMI, waist-to-hip ratio, body fat percentage, cardiovascular disease, etc.) and reran the MR to obtain more robust MR estimates. For TV watching, breakfast skipping and T2D, independent genetic instruments were selected at GWAS p value < 5×10^{-8} and LD r^2 < 0.001 based on the 1000 Genomes European phase 3 reference panel. Given the multiple comparisons, in this study, we considered a P threshold < 0.05 as suggestive significance, while Bonferroni-corrected P threshold was used as statistically significant (P<0.05/6 = 0.008).

To further assess the direct effects of TV watching on T2D, we performed two-step MR mediation analysis. We selected body mass index (BMI), 4 lipid traits [including high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol, triglyceride (TG), total cholesterol (TC)], and 6 glycemic traits [including fasting glucose (FG), fasting insulin (FI), 2-h postprandial glucose (2hGlu), hemoglobin A1c (HbA1c), homeostatic model assessment of beta cell function (HOMA-β), homeostatic model assessment of insulin resistance (HOMA-IR)] as potential mediators of liability to TV watching in T2D. Two-step MR is based on the coefficient product method to calculate indirect (or mediator) effects (**Figure 1**). This process involves calculating two MR estimates, one for the causal effect of exposure on the mediator and the other for the causal effect of the mediator on the outcome. These two estimates are then multiplied together to estimate the indirect effect (38). GWAS summary statistics for BMI, lipid traits, and glycemic traits were obtained from the Genetics of ANthropometric Traits (GIANT) Consortium, the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC), and the Global Lipid Genetics Consortium (GLGC), respectively. The source literature corresponding to the three mediated traits can be found here (39–42). There was no sample size overlap between exposures and mediators and little overlap between mediators and outcomes in the selected GWAS data. Bonferroni-corrected P threshold (P<0.05/11) was used as statistical significance accounting for the 11 mediation analyses.

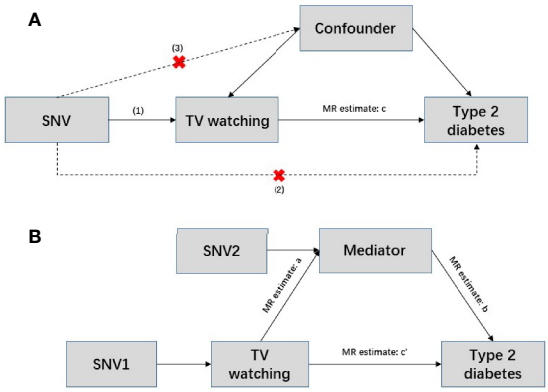


FIGURE 1 | Conceptual diagram of Mendelian randomization and mediation analysis. **(A)** Mendelian randomization is based on the following three assumptions. (1) Genetic variants are strongly associated with exposure ($p<5\times10^{-8}$); (2) instrumental variables can only act on the outcome through exposure, and there is no direct association with the outcome; and (3) instrumental variables are independent of any confounding factors. In this situation, c represents the total effect, SNV: single nucleotide variant. **(B)** Two-step Mendelian randomization, where a represents the effect of the exposure on the mediator; b represents the effect of mediator on the outcome; c' represents the direct effect; and a and b are estimated separately using separate genetic instrumental variables for both the exposure and mediator. These estimates are then multiplied together to estimate the indirect effect of the mediator (a * b), and the direct effect c' = c – a*b.

RESULTS

Genetic Correlations

T2D showed a strong positive genetic association with TV watching (r_g = 0.26; P = 1.63×10^{-29}) and skipping breakfast (r_g = 0.15; P =2.02×10^{-6}). The results suggested a potential common genetic basis and thus warranted further investigation of the underlying mechanisms using cross trait meta-analysis and instrumental variable analysis (**Table 1**).

Cross Trait Meta-Analysis

We identified eight index loci shared between T2D and TV watching (P_{meta} < 2.5×10^{-8} and single-trait P < 1×10^{-5}). However, we did not find any shared loci between T2D and breakfast skipping. GWAS-eQTL colocalization analysis had no significant results, but it identified a specific region at 12q14.3 that might be an expression quantitative trait locus between T2D and TV watching (tissue: lung, mapped gene: *HMGA2*, $P_{nominal}$ = 1.79×10^{-4}, H4 = 1.29×10^{-3}). Two of our CPASSOC index SNPs are located at the 12q14.3 region mapping to *HMGA2* gene. *HMGA2* encodes a protein belonging to the non-histone

TABLE 1 | Genetic correlation of type 2 diabetes with TV watching and breakfast skipping, estimated by high-definition likelihood method (HDL-method) and linkage disequilibrium score regression (LDSC).

Method	Trait	r_g	SE	r_g, 95%CI	pvalue	h^2(SE)
HDL-method	TV watching	0.26	0.023	0.21 to 0.31	1.63E-29	0.13(0.004)
	breakfast skipping	0.15	0.032	0.09 to 0.21	2.02E-6	0.05(0.002)
LDSC	TV watching	0.28	0.030	0.22 to 0.34	1.28E-21	0.13(0.004)
	breakfast skipping	0.14	0.043	0.06 to 0.22	1.30E-3	0.05(0.003)

Summary statistics for each trait were merged with Hapmap3 SNPs excluding the HLA region to estimate r_g; p value < 0.05/2; h^2 indicates the heritability of the corresponding phenotype.

chromosomal high-mobility group (HMG) protein family, and the protein contains structural DNA-binding domains and may act as a transcriptional regulating factor. Significantly higher expression of *HMGA2* mRNA in white adipose tissue has been reported in patients with T2D (43).

More importantly, we identified one novel locus shared between T2D and TV watching (11q13.1, index SNP: rs78028320, mapped gene: *CFL1*, $P_{meta} = 2.68\times10^{-9}$). *CFL1* is a typical protein-coding gene that encodes cofilin-1, an intracellular actin regulatory protein that plays an important role in regulating the organization of the actin cytoskeleton. Phosphorylated (inactive) cofilin-1 is upregulated in diabetic glomeruli, suggesting alterations in actin dynamics (44). In addition, podocytes in glomeruli are the key structure for maintaining the selective filtration barrier of the kidney. Its loss and structural abnormalities contribute to the progression of diabetic nephropathy (45). It has also been reported that mice deleted of *CFL1* in podocytes developed increased albuminuria and developed renal dysfunction, as indicated by a rise in creatinine (46).

The most significant locus overall was index SNP rs4420638 (mapped gene: *APOC1*, $P_{meta} = 2.42\times10^{-14}$). The mapped gene *APOC1* (apolipoprotein C1) is a protein-coding gene engaged in the inhibition of cholesteryl ester transfer protein (CETP). A study showed that *APOC1* was highly expressed in clear cell renal cell carcinoma (47), and a variant of *APOC1* called T45S led to elevated rates of T2D (48). The second strongest SNP was rs4565329 (mapped gene: *CENPW*, $P_{meta} = 7.64\times10^{-14}$). *CENPW* encodes a centromere protein that plays a central role in the assembly of kinetochore proteins, mitotic progression and chromosome segregation. The association between *CENPW* and T2D has been reported in previous genome-wide meta-analysis (49). SNP rs74333814 was also shared between TV watching and T2D (mapped gene: *ARAP1*, $P_{meta} = 3.84\times10^{-13}$). *ARAP1* encodes protein that is thought to regulate the cell-specific trafficking of a receptor protein involved in apoptosis. Findings suggest that *ARAP1* engages in islet insulin content and secretion and is thus likely to mediate the effects on diabetes susceptibility (50). Significantly, previous studies also showed that *APOC1* (51) and *ARAP1* (52) had a significant effect on BMI.

Transcriptome-Wide Association Studies

We next delved into the genetic level and examined shared TWAS genes between TV watching, breakfast skipping and T2D. After Bonferroni correction, a total of 10127 gene-tissue pairs were found to be significantly associated with T2D in 48 GTEx tissues, in addition to 7540 and 143 gene-tissue pairs associated with TV watching and breakfast skipping, respectively. We found 365 TWAS-significant genes shared between T2D and TV watching, with significant system-wide overlap, especially in the endocrine system, cardiovascular system, digestive system and nervous system (**Figure 2**). Intriguingly, 6 of the 365 shared TWAS-significant genes were also identified in CPASSOC, including *CENPW*, *ARAP1*, *CFL1*, *HMGA2*, *ABO* and *ATG16L2*. The functions of the first four genes have been described in detail in the CPASSOC section, and here, we focus on the two genes *ABO* and *ATG16L2*. The *ABO* (9q34.2) gene encodes the blood group ABO systemic transferase and is ubiquitously expressed in many tissues and cell types (53). Genetic variation at the ABO locus and ABO blood group have been found to be associated with the risk of venous thromboembolism (54) and type 2 diabetes (55). *ATG16L2* (11q13.4) is a protein-coding gene whose function is not fully understood, and it has been shown to play a unique function in autophagy. Analysis of transcriptomic data shows that autophagy plays a major role in the molecular pathology of T2D and AD (56).

However, for T2D and breakfast skipping, we observed only 12 shared TWAS-significant genes, mainly enriched in the endocrine system (**Figure 2**). Notably, we found that *EIF2S2P3* was the most enriched and significant among the 12 shared genes. *EIF2S2P3* is located at 10p23.33 and is a pseudogene. It has been reported to be associated with T2D (56), but its function remains unclear.

Mendelian Randomization Analysis

In our MR study, for T2D, TV watching, and breakfast skipping, we selected 35, 127 and 5 SNPs as IVs, respectively. The detailed characteristics of the IVs are shown in **Tables S1**-**S4**, and the screening flow of IVs is shown in **Figure 3**. F statistics provide an indication of the strength of the instrument and can be calculated using formula $F = \frac{n-k-1}{k} \cdot \frac{r^2}{1-r^2}$ (n is sample size, k is the number of IVs, and r^2 refers to how much variation in the trait can be explained by the set of genetic instruments used) (57). Given that r^2 is not generally provided in GWAS summary data, we used the formula $r^2 = \sum \left[\frac{\beta^2 \cdot 2 \cdot f \cdot (1-f)}{\beta^2 \cdot 2 \cdot f \cdot (1-f) + se^2 \cdot 2 \cdot n \cdot f \cdot (1-f)}\right]$ (f is effect allele frequency, n is sample size, β is effect estimate for each SNP and *se* is standard error for each SNP) (58) to obtain r^2 estimates. The F statistics for T2D, TV watching and breakfast skipping IVs are 69.86, 142.42 and 49.69, respectively (F >10 demonstrates that the analysis is unlikely to be affected by weak instrumental bias) (59).

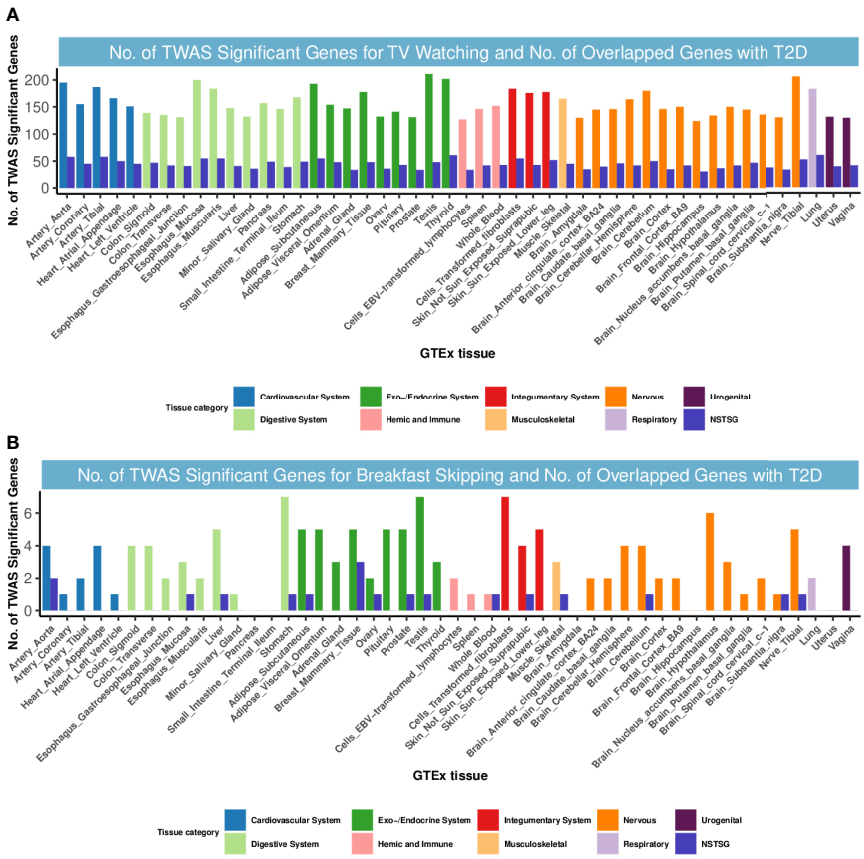


FIGURE 2 | Numbers of significant genes related to TV watching and breakfast skipping and the number of shared genes with T2D. Significant genes were identified by $P_{Bonferroni} < 0.05$. GTEx, genotype-tissue expression project; GWAS, genome-wide association studies; TWAS, transcriptome-wide association study; NSTSG, Number of shared TWAS significant genes between traits; T2D: type 2 diabetes. **(A)** No. of TWAS Significant Genes for TV watching and No. of Overlapped Genes with T2D. **(B)** No. of TWAS Significant Genes for breakfast skipping and No. of Overlapped Genes with T2D.

As shown in **Table 2**, TV watching was positively associated with the risk of type 2 diabetes [OR (95% CI)$_{IVW}$ = 1.86 (1.54, 2.26), $P = 1.80\times10^{-10}$; OR_{WM} = 1.82 (1.43, 2.32), $P = 1.12\times10^{-6}$; $OR_{MR\text{-}RAPS}$ = 1.78 (1.50, 2.11), $P = 3.13\times10^{-11}$; $OR_{MR\text{-}PRESSO\ Outlier\text{-}corrected}$ = 1.84 (1.56, 2.16), $P = 1.22\times10^{-11}$], with all P values reaching the Bonferroni-corrected threshold and without any evidence of pleiotropy ($P_{MR\text{-}Egger\text{-}intercept}$ = 0.41). This causal effect became more significant in the sensitivity analysis

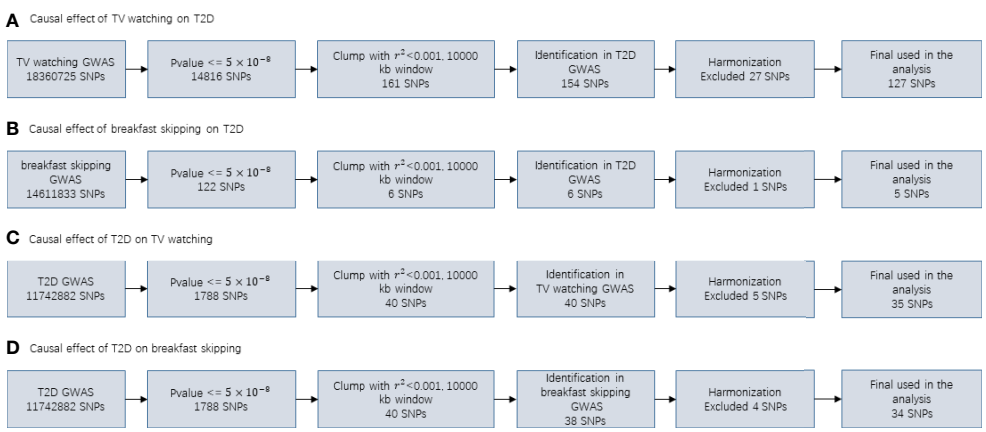


FIGURE 3 | Flowcharts visualizing the process for instrument definition, extraction and harmonization for the two-sample MR analyses conducted in the present study.

TABLE 2 | Causal relationships between TV watching, skipping breakfast and T2D (findings adjusted for multiple comparisons).

Exposure	Outcome	N_snp	Method	beta	OR	95%CI[#]	SE	p_value	Heterogeneity_P_value	Intercept_P_value	Steiger_P_value
TV watching	T2D	127	IVW	0.629	1.86	(1.54,2.26)	0.098	**1.80E-10**	1.66E-05	NA	1.12E-168
			WM	0.599	1.82	(1.44,2.3)	0.12	**6.36E-07**	NA	NA	
			MR–Egger	0.253	1.29	(0.52,3.17)	0.46	5.83E-01	1.60E-05	0.41	
			MR-RAPS	0.577	1.78	(1.5,2.11)	0.087	**3.13E-11**	NA	NA	
			MR-PRESSO: raw	0.569	1.77	(1.49,2.09)	0.086	**6.07E-10**	NA	NA	
			MR-PRESSO : Outlier-corrected	0.609	1.84	(1.56,2.16)	0.083	**1.22E-11**	NA	NA	
skipping breakfast	T2D	5	IVW	0.232	1.26	(0.51,3.14)	0.465	6.18E-01	0.11	NA	3.67E-16
			WM	0.752	2.12	(0.89,5.07)	0.444	8.99E-02	NA	NA	
			MR–Egger	2.111	8.25	(0.31,219.57)	1.674	2.97E-01	0.16	0.33	
			MR-RAPS	0.255	1.29	(0.63,2.67)	0.37	4.90E-01	NA	NA	
			MR-PRESSO: raw	0.239	1.27	(0.6,2.69)	0.383	5.61E-01	NA	NA	
			MR-PRESSO: Outlier-corrected	NA	NA	NA	NA	NA	NA	NA	
T2D	TV watching	35	IVW	-0.003	NA	(-0.017 0.011)	0.007	6.16E-01	3.04E-09	NA	2.87E-290
			WM	0.001	NA	(-0.011,0.013)	0.006	8.79E-01	NA	NA	
			MR–Egger	0.012	NA	(-0.021,0.045)	0.017	4.82E-01	4.78E-09	0.34	
			MR-RAPS	-0.002	NA	(-0.016,0.012)	0.007	7.55E-01	NA	NA	
			MR-PRESSO: raw	-0.002	NA	(-0.014, 0.010)	0.006	8.07E-01	NA	NA	
			MR-PRESSO: Outlier-corrected	-0.001	NA	(-0.011, 0.009)	0.005	8.85E-01	NA	NA	
T2D	skipping breakfast	34	IVW	-0.002	NA	(-0.016, 0.012)	0.007	7.72E-01	1.29E-03	NA	1.28E-203
			WM	-0.001	NA	(-0.017, 0.015)	0.008	9.25E-01	NA	NA	
			MR–Egger	0.009	NA	(-0.024, 0.042)	0.017	5.99E-01	1.15E-03	0.49	
			MR-RAPS	0.004	NA	(-0.010, 0.018)	0.007	5.24E-01	NA	NA	
			MR-PRESSO: raw	0.002	NA	(-0.010,0.014)	0.006	7.58E-01	NA	NA	
			MR-PRESSO : Outlier-corrected	-0.001	NA	(-0.013,0.011)	0.006	9.08E-01	NA	NA	

T2D, type 2 diabetes; CI, confidence interval; IVW, inverse variance weighted; MR, Mendelian randomization; NA, not applicable; N_snp: number of instrumental variables; OR, odds ratio; SE, standard error; SNP, single nucleotide polymorphism; WM, weighted median. When T2D is used as the outcome, there is an OR value.

[#]: 95% CIs of ORs are presented for the analysis of T2D as outcome, while 95% CIs of β values are presented for the analysis of the other outcomes.

p_value in bold refers to achieving statistical significance (p_value < 0.05/6).

excluding 16 SNPs associated with potential confounders (**Table 3**) [OR (95% CI)$_{IVW}$ = 1.94 (1.60, 2.36), $P = 3.74\times10^{-11}$; OR_{WM} = 1.82 (1.41, 2.35), $P = 3.27\times10^{-6}$; $OR_{MR\text{-}RAPS}$ = 1.78 (1.50, 2.11), $P = 3.13\times10^{-11}$; OR $_{MR\text{-}PRESSO\ :\ Outlier\text{-}corrected}$ = 1.84 (1.56, 2.16), $P = 1.22\times10^{-11}$]. The confounding traits associated with the 16 SNPs can be found in **Table S5**. However, there was no significant causal effect estimate from breakfast skipping to T2D. Due to shared biological pathways, T2D may further influence unhealthy lifestyles. To explore whether there is reverse causality, we performed an inverse MR analysis. We did not observe any significant association between genetic predisposition to T2D with TV watching and breakfast

TABLE 3 | The association between TV watching and risk of type 2 diabetes after remove 16 SNPs associated with confounding traits.

Exposure	Outcome	N_snp	Method	beta	OR	CI	SE	p_value	Heterogeneity_P_value	Intercept_P_value	Steiger_P_value
TV watching	**T2D**	111	IVW	0.66	1.94	(1.6,2.36)	0.1	3.74E-11	1.66E-05	NA	1.1E-168
		111	WM	0.59	1.82	(1.41,2.35)	0.129	3.27E-06	NA	NA	1.1E-168
		111	MR Egger	0.60	1.83	(0.71,4.69)	0.481	0.21	1.60E-05	0.41	1.1E-168
		111	MR-RAPS	0.58	1.78	(1.5,2.11)	0.087	3.13E-11	NA	NA	1.1E-168
		111	MR-PRESSO:raw	0.57	1.77	(1.49,2.09)	0.086	6.07E-10	NA	NA	1.1E-168
		111	MR-PRESSO: Outlier-corrected	0.61	1.84	(1.56,2.16)	0.083	1.22E-11	NA	NA	1.1E-168

T2D, type 2 diabetes; CI, confidence interval; IVW, inverse variance weighted; MR, Mendelian randomization; NA, not applicable; N_snp, number of instrumental variables; OR, odds ratio; SE, standard error; SNP, single nucleotide polymorphism; WM, weighted median. When T2D is used as the outcome, there is an OR value.

skipping (**Table 2** all $P > 0.05$). The leave-one-out cross-validation analysis showed that the overall estimates were not overdriven by any particular SNP (**Figures S1**-**S4**). The MR Steiger results showed that all causal estimates were in the intended direction (all $\boldsymbol{P}_{\text{MR Steiger}} \ll 0.05$, **Table 2**). The nearly symmetric funnel plots indicate no evidence of pleiotropy in the analysis (**Figures S5**-**S8**). In summary, instrumental variable analysis suggests a potential causal effect of increased TV watching time on an increased risk of T2D.

Epidemiological studies have shown that prolonged TV watching leads to increased BMI (60), lower HDL cholesterol (61), and higher fasting glucose concentrations (62) and that BMI and blood glycolipid traits are known risk factors for T2D (63), suggesting a potential mediating role for these traits in the association between TV watching and T2D. We performed a two-step MR mediation analysis to explain the mediation proportion for BMI, 4 lipid traits, and 6 glycemic traits. As shown in **Table 4**, the results revealed that four potential mediators produced a significant mediating effect. After adjusting for HbA1c, FG, and HDL, the estimates of causal effects produced moderate attenuation (OR: 1.78 adjusted for HbA1c, 1.71 adjusted for FG and 1.75 adjusted for HDL). In contrast, the association between TV watching and the risk of T2D was much more attenuated after adjusting for BMI (OR: 1.55 adjusted for BMI). Mediation analysis showed that the causal association between TV watching and T2D risk was partially mediated by BMI (mediation percentage = 29.10%), FG (mediation percentage = 13.51%), HDL (mediation percentage = 9.86%) or HbA1c (mediation percentage = 7.31%). Adjusting for these four factors simultaneously and adjusting for each factor separately produced results that were in the same direction as the results without adjustment, although the effect size was attenuated. In addition, we did not observe significant mediating effects for the other 7 glycemic-lipid traits.

Finally, we calculated the statistical power of this study using the mRnd website (64) (https://shiny.cnsgenomics.com/mRnd/). With the current sample size of T2D and the phenotypic variance of TV watching explained by IVs (4.1%, **Table S3**), at an alpha level of 0.05, we had 99% power to determine that each standard deviation increase in TV watching time increased the overall risk of T2D by 86% (i.e., an OR_{IVW} of 1.86, **Table 2**).

DISCUSSION

In the present study, we conducted a comprehensive genetic analysis to explore causal relationships and genetic overlap between T2D and TV watching and breakfast skipping by using summary statistics from GWAS. In the first instance, we showed that there was a strong positive genetic correlation between T2D and both exposures. Second, shared genetic structure at the locus level was identified between T2D and TV watching in cross-trait association analysis. Third, in the TWAS study between T2D and TV watching, we identified TWAS-significant genes, especially in tissues from the endocrine system, cardiovascular system, digestive system and nervous system. Finally, and most importantly, bidirectional MR showed that TV watching was positively associated with the risk of T2D. Mediation analysis identified four different traits as potential mediating factors between TV watching and T2D. Our results in the present study highlighted that TV watching plays an important role in the risk of T2D. The genetic overlaps elucidate potential shared biological pathways, thus providing new ideas and opportunities for T2D treatment and drug design.

The results of genetic correlation analysis are highly consistent with observational studies showing that breakfast skipping (8) and TV watching are significantly associated with an increased risk of T2D (4). These findings do not necessarily imply that TV watching per se causes T2D; rather, we believe that prolonged TV watching and breakfast skipping significantly affect the risk of developing diabetes in the future. There are two possible explanations for the observed positive association between TV watching and the risk of T2D. First, prolonged TV watching may result in lower energy expenditure and higher caloric intake, which are directly associated with obesity and weight gain (65, 66). Second, individuals who spend more time watching TV tend to eat more processed meats, snacks, and sweets and fewer vegetables and fruits, and such a diet may inversely affect diabetes risk (67). The average time spent watching TV is significantly associated with elevated levels of leptin and LDL cholesterol and lower levels of HDL cholesterol and apolipoprotein, which are important plasma biomarkers of T2D (68). Similarly, skipping breakfast may also trigger hyperglycemia and high glycated hemoglobin after lunch and dinner, further leading to impaired insulin response and thus increasing the risk of T2D (69). For these possible mechanistic pathways, we made

TABLE 4 | Two-step Mendelian randomization mediation analysis of the association between TV watching (exposure) and type 2 diabetes (outcome).

Mediator	Exposure → Mediator			Mediator → Outcome			Indirect causal effect by coefficient product	Direct causal effect	Adjust OR	Proportion of mediation
	IVW causal effect	IVW p value	MR Egger Intercept p value	IVW causal effect	IVW p value	MR Egger Intercept p value				
Adjust for BMI	0.315	**2.76E-06**	0.195	0.581	**5.14E-04**	0.563	0.183	0.439	1.55	29.10%
Adjust for TC	0.112	1.08E-01	0.119	-0.1	4.21E-02	0.307	NA	NA	NA	NA
Adjust for TG	0.24	**3.18E-06**	0.207	0.106	1.61E-01	0.028	NA	NA	NA	NA
Adjust for HDL	-0.289	**1.22E-05**	0.002	-0.213	**7.15E-04**	0.008	0.062	0.561	1.75	9.86%
Adjust for LDL	0.171	6.58E-03	0.197	-0.033	4.96E-01	0.344	NA	NA	NA	NA
Adjust for FG	0.053	**9.45E-04**	0.589	1.602	**4.03E-08**	0.015	0.085	0.537	1.71	13.51%
Adjust for FI	0.088	**1.09E-06**	0.898	1.318	6.19E-02	0.253	NA	NA	NA	NA
Adjust for HOMA-β	0.074	**4.01E-03**	0.862	-2.595	1.76E-01	0.221	NA	NA	NA	NA
Adjust for HOMA-IR	0.176	**2.13E-08**	0.596	0.346	2.03E-01	0.405	NA	NA	NA	NA
Adjust for 2hGlu	0.063	3.16E-01	0.506	0.921	1.78E-02	0.823	NA	NA	NA	NA
Adjust for HbA1c	0.038	**5.92E-04**	0.944	1.223	**3.08E-03**	0.183	0.046	0.576	1.78	7.31%
Adjust for ALL	NA	NA	NA	NA	NA	NA	0.376	0.253	1.29	59.78%

*BMI, body mass index; TC, total cholesterol; TG, triglyceride; HDL, high density lipoprotein; LDL, low density lipoprotein; FG, fasting glucose; FI, fasting insulin; HOMA-β, homeostatic model assessment of beta cell function; HOMA-IR, homeostatic model assessment of insulin resistance; 2hGlu, 2-h postprandial glucose; HbA1c, hemoglobin; NA, not applicable; The IVW causal effect size was the beta coefficient estimated by IVW models for corresponding outcome; Direct causal effect: this value is obtained by subtracting the indirect effect from 0.629 as show in **Table 5**; IVW p values < 0.05/11 indicate statistical significance and are marked in bold font, and mediation analysis is significant only if both MR steps reach statistical significance; Proportion of mediation = Indirect causal effect by coefficient product/0.629.*

TABLE 5 | Cross-trait meta-analysis results between type 2 diabetes and television watching ($P_{meta} < 2.5\times10^{-8}$ and single-trait $P < 1\times10^{-5}$).

Index.SNP	CHR	Genome position	EA	NEA	EAF	T2D		TV watching		P_{meta}	Genes	variant annotation
						BETA	P	BETA	P			
rs4420638	19	19q13.32	A	G	0.84	0.110	1.50E-09	0.014	3.60E-07	2.42E-14	[APOC1,APOE,PVRL2,TOMM40]	downstream
rs4565329	6	6q22.32	T	C	0.48	0.073	4.40E-09	0.010	1.50E-06	7.64E-14	[CENPW]	intron
rs74333814	11	11q13.4	T	C	0.86	-0.095	5.80E-09	-0.014	3.50E-06	3.84E-13	[ARAP1,ATG16L2,FCHSD2,MIR4692, STARD10]	intron
rs243024	2	2p16.1	A	G	0.55	0.066	3.90E-08	0.011	1.00E-06	4.39E-12	[AC007381.3]*	upstream
rs2258238	12	12q14.3	A	T	0.88	-0.110	1.60E-07	-0.016	7.40E-06	1.23E-11	[HMGA2,RPSAP52]	intron
rs10400419	12	12q14.3	T	C	0.57	-0.067	1.70E-07	-0.010	9.60E-06	1.34E-10	[HMGA2]*	intergenic
rs550057	9	9q34.2	T	C	0.76	0.065	3.40E-06	0.012	3.00E-06	1.81E-09	[ABO]	intron
rs78028320	11	11q13.1	A	G	0.82	0.069	5.80E-06	0.013	3.90E-06	2.68E-09	[CFL1]*	intergenic

*EA, effect allele; NEA, noneffect allele; P_{meta} is the cross-trait meta-analysis P value. CHR, chromosome; T2D, type 2 diabetes; genes in * are the nearest genes to this locus.*

presumptions and validated them in the subsequent shared genetic structure analysis and MR-mediated analysis.

CPASSOC and TWAS showed that the shared genes between TV watching and T2D were mostly enriched in the endocrine system and cardiovascular system, suggesting an underlying correlation between the biological pathway and these tissues. Study shows that the *CFL1* gene, which controls cell proliferation and cell death, is overexpressed in the subcutaneous adipose

tissue of subjects who have gained weight, suggesting that the *CFL1* gene affects the risk of T2D through a mediating pathway of BMI (70). Reports have demonstrated that elevated *APOC1* gene expression is significantly associated with the risk of T2D and TG levels; also, apoC1 glycosylation has been observed in patients with T2D, which impairs the ability of *APOC1* to inhibit plasma cholesteryl ester transporter protein activity, suggesting that elevated apoC1 expression may increase the risk of T2D through lipoprotein metabolic pathways (71, 72). *APOC1* has also been reported to activate lecithin-cholesterol acyltransferase (LCAT), which in turn promotes HDL cholesterol esterification and increases HDL levels (73). Furthermore, increased *HMGA2* expression can be expected to lead to increased expression of $p14^{Arf}$, an inducer of cellular senescence, and the accumulation of senescent cells triggers inflammation associated with insulin resistance, driving the development of T2D, predicting that TV watching induces a signaling pathway linked to cellular senescence to increase the risk of T2D (43). Of additional interest to us is the fact that individuals who watch television for long periods of time consume more food and energy, increasing the burden on the digestive system (74). Additionally, patients with T2D often experience gastrointestinal disturbances, suggesting that gastrointestinal disturbances play a collider role in the association between TV watching and T2D (75). The exact mechanism of the digestive system in this association needs to be further elaborated. Moreover, previous research shows that *APAR1* affects the function of pancreatic β-cells and that the proinsulin-raising allele of *ARAP1* is related to a decreasing risk of T2D (76). The opposite conclusion was also reported: T2D pathogenic activity is mediated by *STARD10* expression instead of *ARAP1* (77), but both genes are located in a specific region, 11q13.4, which was identified in our cross-trait analysis, implying that pancreatic β-cell and proinsulin processing may be located in the biological pathway between TV watching and T2D. Our study suggests that multisystem, multitissue, polygenic effects may have a synergistic effect on the risk of T2D, but this needs more experimental evidence for further clarification.

Overall, using the MR study design, we found strong causal relationship between TV watching time and an increased risk of T2D. The observed causal effect was greatly attenuated when the mediating role of BMI, glycemia, and lipids was taken into account, suggesting that BMI, glycemia, and lipids play a key role in the association. Our finding is consistent with most previous observational studies and meta-analyses showing that prolonged TV watching is associated with an increased risk of T2D. A recent systematic review and dose–response meta-analysis based on 11 prospective studies published from 2001-2016 showed a linear association between TV watching and T2D (78), which was again validated in a recent meta-analysis (79). Our results are also supported by previous epidemiological studies that used Cox proportional hazards regression, controlling for multiple time-independent (i.e., constant across all cycles) and time-related (i.e., varying from cycle to cycle) covariates, to clarify that watching more than 4 hours of television and video per day at age 16 increases the risk of developing T2D (80). Moreover, this association was also verified in a multivariate logistic regression study based on an East Asian population that took into account gender differences (6). In addition, cross-sectional and longitudinal studies assessing the association between TV watching time and cardiometabolic biomarkers among multiple ethnic groups corroborated the plausibility of our choice of mediating variables and provided some potential mechanistic pathways that act through these mediators (62, 68, 81). However, a recent MR analysis of sedentary behavior with T2D and glycemic traits contradicts our results, finding no causal relationship between sedentary behavior and T2D. Two reasons may explain this discrepancy, one of which is that sedentary behavior is assessed by accelerometers, which is not conducive to measuring posture and sedentariness and estimating energy expenditure (82). In addition, the presence of the Hawthorne effect makes it possible for subjects to change their habituation (83). Second, although they also used data from UKB, the sample size was so small ($N = 91084$) that they could not select enough IVs to improve the statistical power (number of IVs = 6 in their study) (84). We also acknowledge the discrepancy between the results of breakfast skipping and T2D, and the findings of traditional epidemiological investigations may be partly due to fewer IVs for breakfast skipping.

In contrast to traditional observational studies and randomized controlled trials, the highlight of this study is the MR approach, which allows estimation of the causal effect of unhealthy lifestyles on T2D with a large sample size and high precision, controlling for potential reverse causality and confounders to the maximum extent possible. In addition, this study used various methods for sensitivity analysis, especially excluding SNPs related to potential confounders, to enhance the strength of instrumental variables and improve the robustness of estimation. Two-step MR mediation analysis was used in our study. When the results are binary variables (e.g., T2D), the estimation accuracy obtained by this method is higher than that obtained by multivariate Mendelian randomization (MVMR) (85). However, several potential shortcomings need to be acknowledged. First, in TWAS and GWAS-eQTL analysis, small eQTL samples are not sufficient to detect relatively weak signals, reducing the efficacy of the method. Second, our study is limited to individuals of European ancestry and cannot be generalized to other ethnicities. Third, no sex-specific MR analysis was conducted for the association between TV watching and T2D in our study. In addition, the analysis of breakfast skipping was limited to a few IVs and could not produce results with high power and reliability. Finally, further exploration of unhealthy lifestyle and T2D association mechanisms in the future, such as larger replication studies, sex-specific studies based on individual data, and more studies of mediating factors (hypertension, physical activity, education attainment, diet, leptin level, etc.), would greatly benefit our findings.

Our comprehensive genetic analysis identified shared genetic similarities between TV watching and T2D, suggesting a strong intrinsic genetic link between this trait pair. We further used MR to find convincing evidence supporting a putative causal role between TV watching and T2D, but mediation analyses suggest that this effect is largely mediated by BMI, HbA1c, FG, and HDL. As obesity, hyperglycemia, and hyperlipidemia are recognized as established risk factors for T2D, our findings underscore the importance of actionable prevention strategies for T2D. However, to date, the complex interactions between TV watching and T2D do not appear to be fully understood, and further studies are needed to deepen our

understanding of the biological pathways by which TV watching influences T2D.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors. Custom ***R*** scripts used to generate results in this study can be made available upon request.

ETHICS STATEMENT

The study was conducted using publicly available summary-level genetic data, and no ethical approval was requested.

AUTHOR CONTRIBUTIONS

JJ and TH conceived and designed the study. DC and HW performed the data preparation and statistical analysis. DC and HW wrote the manuscript. DC and HW contributed equally to this article. All authors helped interpret the data, reviewed and edited the final paper and approved the submission.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.836023/full#supplementary-material

REFERENCES

1. Meigs JB, Cupples LA, Wilson PW. Parental Transmission of Type 2 Diabetes: The Framingham Offspring Study. *Diabetes* (2000) 49:2201–7. doi: 10.2337/diabetes.49.12.2201
2. Ali O. Genetics of Type 2 Diabetes. *World J Diabetes* (2013) 4:114–23. doi: 10.4239/wjd.v4.i4.114
3. Hu FB, Li TY, Colditz GA, Willett WC, Manson JE. Television Watching and Other Sedentary Behaviors in Relation to Risk of Obesity and Type 2 Diabetes Mellitus in Women. *JAMA* (2003) 289:1785–91. doi: 10.1001/jama.289.14.1785
4. Grontved A, Hu FB. Television Viewing and Risk of Type 2 Diabetes, Cardiovascular Disease, and All-Cause Mortality: A Meta-Analysis. *JAMA* (2011) 305:2448–55. doi: 10.1001/jama.2011.812
5. Bennett DA, Du H, Bragg F, Guo Y, Wright N, Yang L, et al. Physical Activity, Sedentary Leisure-Time and Risk of Incident Type 2 Diabetes: A Prospective Study of 512 000 Chinese Adults. *BMJ Open Diabetes Res Care* (2019) 7:e000835. doi: 10.1136/bmjdrc-2019-000835
6. Ikehara S, Iso H, Maruyama K, Ukawa S, Tamakoshi A. Television Viewing Time, Walking Time, and Risk of Type 2 Diabetes in Japanese Men and Women: The Japan Collaborative Cohort Study. *Prev Med* (2019) 118:220–5. doi: 10.1016/j.ypmed.2018.11.006
7. Mekary RA, Giovannucci E, Willett WC, Van Dam RM, Hu FB. Eating Patterns and Type 2 Diabetes Risk in Men: Breakfast Omission, Eating Frequency, and Snacking. *Am J Clin Nutr* (2012) 95:1182–9. doi: 10.3945/ajcn.111.028209
8. Bi H, Gan Y, Yang C, Chen Y, Tong X, Lu Z. Breakfast Skipping and the Risk of Type 2 Diabetes: A Meta-Analysis of Observational Studies. *Public Health Nutr* (2015) 18:3013–9. doi: 10.1017/S1368980015000257
9. Ballon A, Neuenschwander M, Schlesinger S. Breakfast Skipping Is Associated With Increased Risk of Type 2 Diabetes Among Adults: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. *J Nutr* (2019) 149:106–13. doi: 10.1093/jn/nxy194
10. Krishnan S, Rosenberg L, Palmer JR. Physical Activity and Television Watching in Relation to Risk of Type 2 Diabetes: The Black Women's Health Study. *Am J Epidemiol* (2009) 169:428–34. doi: 10.1093/aje/kwn344
11. Al-Goblan AS, Al-Alfi MA, Khan MZ. Mechanism Linking Diabetes Mellitus and Obesity. *Diabetes Metab Syndr Obes* (2014) 7:587–91. doi: 10.2147/DMSO.S67400
12. Deshmukh-Taskar P, Nicklas TA, Radcliffe JD, O'neil CE, Liu Y. The Relationship of Breakfast Skipping and Type of Breakfast Consumed With Overweight/Obesity, Abdominal Obesity, Other Cardiometabolic Risk Factors and the Metabolic Syndrome in Young Adults. The National Health and Nutrition Examination Survey (NHANES): 1999-2006. *Public Health Nutr* (2013) 16:2073–82. doi: 10.1017/S1368980012004296
13. Haase CL, Tybjærg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. HDL Cholesterol and Risk of Type 2 Diabetes: A Mendelian Randomization Study. *Diabetes* (2015) 64:3328–33. doi: 10.2337/db14-1603
14. Verduijn M, Siegerink B, Jager KJ, Zoccali C, Dekker FW. Mendelian Randomization: Use of Genetics to Enable Causal Inference in Observational Studies. *Nephrol Dialysis Transplant* (2010) 25:1394–8. doi: 10.1093/ndt/gfq098
15. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the Accuracy of Validated Target Identification Through Coding Variant Fine-Mapping in Type 2 Diabetes. *Nat Genet* (2018) 50:559–71. doi: 10.1038/s41588-018-0084-1
16. Dashti HS, Merino J, Lane JM, Song Y, Smith CE, Tanaka T, et al. Genome-Wide Association Study of Breakfast Skipping Links Clock Regulation With Food Timing. *Am J Clin Nutr* (2019) 110:473–84. doi: 10.1093/ajcn/nqz076
17. Van De Vegte YJ, Said MA, Rienstra M, van der Harst P, Verweij N. Genome-Wide Association Studies and Mendelian Randomization Analyses for Leisure Sedentary Behaviours. *Nat Commun* (2020) 11:1770. doi: 10.1038/s41467-020-15553-w
18. Yang F, Chen S, Qu Z, Wang K, Xie X, Cui H. Genetic Liability to Sedentary Behavior in Relation to Stroke, Its Subtypes and Neurodegenerative Diseases: A Mendelian Randomization Study. *Front Aging Neurosci* (2021) 13:769. doi: 10.3389/fnagi.2021.757388
19. Park S, Lee S, Kim Y, Lee Y, Kang MW, Kim K, et al. Causal Effects of Physical Activity or Sedentary Behaviors on Kidney Function: An Integrated Population-Scale Observational Analysis and Mendelian Randomization Study. *Nephrol Dialysis Transplant* (2021). doi: 10.1093/ndt/gfab153
20. Gao Y, Mi J, Liu Z, Song Q. Leisure Sedentary Behavior and Risk of Lung Cancer: A Two-Sample Mendelian Randomization Study and Mediation Analysis. *Front Genet* (2021) 12. doi: 10.3389/fgene.2021.763626
21. Relton CL, Davey Smith G. Mendelian Randomization: Applications and Limitations in Epigenetic Studies. *Epigenomics* (2015) 7:1239–43. doi: 10.2217/epi.15.88
22. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in Complex Traits: Challenges and Strategies. *Nat Rev Genet* (2013) 14:483–95. doi: 10.1038/nrg3461
23. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* (2017) 66:2888–902. doi: 10.2337/db16-1253
24. Ning Z, Pawitan Y, Shen X. High-Definition Likelihood Inference of Genetic Correlations Across Human Complex Traits. *Nat Genet* (2020) 52:859–64. doi: 10.1038/s41588-020-0653-y
25. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score Regression Distinguishes Confounding From Polygenicity in Genome-Wide Association Studies. *Nat Genet* (2015) 47:291–5. doi: 10.1038/ng.3211
26. Zhu Z, Hasegawa K, Camargo CAJr., Liang L. Investigating Asthma Heterogeneity Through Shared and Distinct Genetics: Insights From Genome-Wide Cross-Trait Analysis. *J Allergy Clin Immunol* (2021) 147:796–807. doi: 10.1016/j.jaci.2020.07.004

27. Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, et al. Meta-Analysis of Correlated Traits *via* Summary Statistics From GWASs With an Application in Hypertension. *Am J Hum Genet* (2015) 96:21–36. doi: 10.1016/j.ajhg.2014.11.011
28. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian Test for Colocalisation Between Pairs of Genetic Association Studies Using Summary Statistics. *PloS Genet* (2014) 10:e1004383. doi: 10.1371/journal.pgen.1004383
29. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank* (2015) 13:311–9. doi: 10.1089/bio.2015.0032
30. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative Approaches for Large-Scale Transcriptome-Wide Association Studies. *Nat Genet* (2016) 48:245–52. doi: 10.1038/ng.3506
31. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps. *Nat Genet* (2018) 50:1505–13. doi: 10.1038/s41588-018-0241-6
32. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization With Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* (2016) 40:304–14. doi: 10.1002/gepi.21965
33. Bowden J, Davey Smith G, Burgess S. Mendelian Randomization With Invalid Instruments: Effect Estimation and Bias Detection Through Egger Regression. *Int J Epidemiol* (2015) 44:512–25. doi: 10.1093/ije/dyv080
34. Hemani G, Tilling K, Davey Smith G. Orienting the Causal Relationship Between Imprecisely Measured Traits Using GWAS Summary Data. *PloS Genet* (2017) 13:e1007081. doi: 10.1371/journal.pgen.1007081
35. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical Inference in Two-Sample Summary-Data Mendelian Randomization Using Robust Adjusted Profile Score. *Ann Stat* (2020) 48:1742–69. doi: 10.1214/19-AOS1866
36. Noyce AJ, Kia DA, Hemani G, Nicolas A, Price TR, De Pablo-Fernandez E, et al. Estimating the Causal Influence of Body Mass Index on Risk of Parkinson Disease: A Mendelian Randomisation Study. *PloS Med* (2017) 14:e1002314. doi: 10.1371/journal.pmed.1002314
37. Verbanck M, Chen CY, Neale B, Do R. Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred From Mendelian Randomization Between Complex Traits and Diseases. *Nat Genet* (2018) 50:693–8. doi: 10.1038/s41588-018-0099-7
38. Carter AR, Sanderson E, Hammerton G, Richmond RC, Davey Smith G, Heron J, et al. Mendelian Randomisation for Mediation Analysis: Current Methods and Challenges for Implementation. *Eur J Epidemiol* (2021) 36:465–78. doi: 10.1007/s10654-021-00757-1
39. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, et al. A Genome-Wide Approach Accounting for Body Mass Index Identifies Genetic Variants Influencing Fasting Glycemic Traits and Insulin Resistance. *Nat Genet* (2012) 44:659–69. doi: 10.1038/ng.2274
40. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and Refinement of Loci Associated With Lipid Levels. *Nat Genet* (2013) 45:1274–83. doi: 10.1038/ng.2797
41. Turcot V, Lu Y, Highland HM, Schurmann C, Justice AE, Fine RS, et al. Protein-Altering Variants Associated With Body Mass Index Implicate Pathways That Control Energy Intake and Expenditure in Obesity. *Nat Genet* (2018) 50:26–41. doi: 10.1038/s41588-019-0447-2
42. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The Trans-Ancestral Genomic Architecture of Glycemic Traits. *Nat Genet* (2021) 53:840–60. doi: 10.1038/s41588-021-00852-9
43. Markowski DN, Thies HW, Gottlieb A, Wenk H, Wischnewsky M, Bullerdiek J. HMGA2 Expression in White Adipose Tissue Linking Cellular Senescence With Diabetes. *Genes Nutr* (2013) 8:449–56. doi: 10.1007/s12263-013-0354-6
44. Wasik AA, Koskelainen S, Hyvönen ME, Musante L, Lehtonen E, Koskenniemi K, et al. Ezrin Is Down-Regulated in Diabetic Kidney Glomeruli and Regulates Actin Reorganization and Glucose Uptake *via* GLUT1 in Cultured Podocytes. *Am J Pathol* (2014) 184:1727–39. doi: 10.1016/j.ajpath.2014.03.002
45. Pagtalunan ME, Miller PL, Jumping-Eagle S, Nelson RG, Myers BD, Rennke HG, et al. Podocyte Loss and Progressive Glomerular Injury in Type II Diabetes. *J Clin Invest* (1997) 99:342–8. doi: 10.1172/JCI119163
46. Garg P, Verma R, Cook L, Soofi A, Venkatareddy M, George B, et al. Actin-Depolymerizing Factor Cofilin-1 Is Necessary in Maintaining Mature Podocyte Architecture. *J Biol Chem* (2010) 285:22676–88. doi: 10.1074/jbc.M110.122929
47. Xiao H, Xu Y. Overexpression of Apolipoprotein C1 (APOC1) in Clear Cell Renal Cell Carcinoma and Its Prognostic Significance. *Med Sci Monit* (2021) 27:e929347. doi: 10.12659/MSM.929347
48. Kasthuri RS, Mcmillan KR, Flood-Urdangarin C, Harvey SB, Wilson-Grady JT, Nelsestuen GL. Correlation of a T45S Variant of Apolipoprotein C1 With Elevated BMI in Persons of American Indian and Mexican Ancestries. *Int J Obes (Lond)* (2007) 31:1334–6. doi: 10.1038/sj.ijo.0803569
49. Zhao W, Rasheed A, Tikkanen E, Lee JJ, Butterworth AS, Howson JMM, et al. Identification of New Susceptibility Loci for Type 2 Diabetes and Shared Etiological Pathways With Coronary Heart Disease. *Nat Genet* (2017) 49:1450–7. doi: 10.1038/ng.3943
50. Carrat G, Meur G, Rutter G. Roles of the Type 2 Diabetes Associated Gene Products Arap1 and StarD10 in the Control of Insulin Secretion. *Diabetic Med* (2013) 30:31–1. doi: 10.1016/j.molmet.2020.101015
51. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, et al. The Influence of Age and Sex on Genetic Associations With Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PloS Genet* (2015) 11:e1005378. doi: 10.1371/journal.pgen.1006166
52. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic Studies of Body Mass Index Yield New Insights for Obesity Biology. *Nature* (2015) 518:197–206. doi: 10.1038/nature14177
53. Greenwell P. Blood Group Antigens: Molecules Seeking a Function? *Glycoconj J* (1997) 14:159–73. doi: 10.1023/A:1018581503164
54. Kim J, Kraft P, Hagan KA, Harrington LB, Lindstroem S, Kabrhel C. Interaction of a Genetic Risk Score With Physical Activity, Physical Inactivity, and Body Mass Index in Relation to Venous Thromboembolism Risk. *Genet Epidemiol* (2018) 42:354–65. doi: 10.1002/gepi.22118
55. Qi L, Cornelis MC, Kraft P, Jensen M, Van Dam RM, Sun Q, et al. Genetic Variants in ABO Blood Group Region, Plasma Soluble E-Selectin Levels and Risk of Type 2 Diabetes. *Hum Mol Genet* (2010) 19:1856–62. doi: 10.1093/hmg/ddq057
56. Caberlotto L, Nguyen TP, Lauria M, Priami C, Rimondini R, Maioli S, et al. Cross-Disease Analysis of Alzheimer's Disease and Type-2 Diabetes Highlights the Role of Autophagy in the Pathophysiology of Two Highly Comorbid Diseases. *Sci Rep* (2019) 9:3965. doi: 10.1038/s41598-019-39828-5
57. Burgess S, Thompson SG. Avoiding Bias From Weak Instruments in Mendelian Randomization Studies. *Int J Epidemiol* (2011) 40:755–64. doi: 10.1093/ije/dyr036
58. Shim H, Chasman DI, Smith JD, Mora S, Ridker PM, Nickerson DA, et al. A Multivariate Genome-Wide Association Analysis of 10 LDL Subfractions, and Their Response to Statin Treatment, in 1868 Caucasians. *PloS One* (2015) 10:e0120758. doi: 10.1371/journal.pone.0120758
59. Staiger D, Stock JH. Instrumental Variables Regression With Weak Instruments. *Econometrica* (1997) 65:557–86. doi: 10.2307/2171753
60. Al-Ghamdi SH. The Association Between Watching Television and Obesity in Children of School-Age in Saudi Arabia. *J Family Community Med* (2013) 20:83–9. doi: 10.4103/2230-8229.114767
61. Martinez-Gomez D, Rey-López JP, Chillón P, Gómez-Martínez S, Vicente-Rodríguez G, Martín-Matillas M, et al. Excessive TV Viewing and Cardiovascular Disease Risk Factors in Adolescents. The AVENA Cross-Sectional Study. *BMC Public Health* (2010) 10:274–4. doi: 10.1186/1471-2458-10-274
62. Dunstan DW, Salmon J, Healy GN, Shaw JE, Jolley D, Zimmet PZ, et al. Association of Television Viewing With Fasting and 2-H Postchallenge Plasma Glucose Levels in Adults Without Diagnosed Diabetes. *Diabetes Care* (2007) 30:516–22. doi: 10.2337/dc06-1996
63. Antwi J, Lavin R, Sullivan S, Bellavia M. Perception of and Risk Factors for Type 2 Diabetes Among Students Attending an Upstate New York College: A Pilot Study. *Diabetol Metab Syndr* (2020) 12:25. doi: 10.1186/s13098-020-00535-1

64. Brion MJ, Shakhbazov K, Visscher PM. Calculating Statistical Power in Mendelian Randomization Studies. *Int J Epidemiol* (2013) 42:1497–501. doi: 10.1093/ije/dyt179
65. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical Activity and Television Watching in Relation to Risk for Type 2 Diabetes Mellitus in Men. *Arch Intern Med* (2001) 161:1542–8. doi: 10.1001/archinte.161.12.1542
66. Vereecken CA, Todd J, Roberts C, Mulvihill C, Maes L. Television Viewing Behaviour and Associations With Food Habits in Different Countries. *Public Health Nutr* (2006) 9:244–50. doi: 10.1079/PHN2005847
67. Hu FB. Sedentary Lifestyle and Risk of Obesity and Type 2 Diabetes. *Lipids* (2003) 38:103–8. doi: 10.1007/s11745-003-1038-4
68. Nang EEK, Salim A, Wu Y, Tai ES, Lee J, Van Dam RM. Television Screen Time, But Not Computer Use and Reading Time, is Associated With Cardio-Metabolic Biomarkers in a Multiethnic Asian Population: A Cross-Sectional Study. *Int J Behav Nutr Phys Act* (2013) 10:70. doi: 10.1186/1479-5868-10-70
69. Jakubowicz D, Wainstein J, Ahren B, Landau Z, Bar-Dayan Y, Froy O. Fasting Until Noon Triggers Increased Postprandial Hyperglycemia and Impaired Insulin Response After Lunch and Dinner in Individuals With Type 2 Diabetes: A Randomized Clinical Trial. *Diabetes Care* (2015) 38:1820–6. doi: 10.2337/dc15-0761
70. Sales V, Patti M-E. The Ups and Downs of Insulin Resistance and Type 2 Diabetes: Lessons From Genomic Analyses in Humans. *Curr Cardiovasc Risk Rep* (2013) 7:46–59. doi: 10.1007/s12170-012-0283-8
71. Bouillet B, Gautier T, Blache D, Pais De Barros J-P, Duvillard L, Petit J-M, et al. Glycation of Apolipoprotein C1 Impairs Its CETP Inhibitory Property: Pathophysiological Relevance in Patients With Type 1 and Type 2 Diabetes. *Diabetes Care* (2014) 37:1148–56. doi: 10.2337/dc13-1467
72. Bouillet B, Gautier T, Aho LS, Duvillard L, Petit JM, Lagrost L, et al. Plasma Apolipoprotein C1 Concentration is Associated With Plasma Triglyceride Concentration, But Not Visceral Fat, in Patients With Type 2 Diabetes. *Diabetes Metab* (2016) 42:263–6. doi: 10.1016/j.diabet.2016.01.003
73. Soutar AK, Garner CW, Baker HN, Sparrow JT, Jackson RL, Gotto AM, et al. Effect of the Human Plasma Apolipoproteins and Phosphatidylcholine Acyl Donor on the Activity of Lecithin: Cholesterol Acyltransferase. *Biochemistry* (1975) 14:3057–64. doi: 10.1021/bi00685a003
74. Temple JL, Giacomelli AM, Kent KM, Roemmich JN, Epstein LH. Television Watching Increases Motivated Responding for Food and Energy Intake in Children. *Am J Clin Nutr* (2007) 85:355–61. doi: 10.1093/ajcn/85.2.355
75. Du YT, Rayner CK, Jones KL, Talley NJ, Horowitz M. Gastrointestinal Symptoms in Diabetes: Prevalence, Assessment, Pathogenesis, and Management. *Diabetes Care* (2018) 41:627–37. doi: 10.2337/dc17-1536
76. Kulzer JR, Stitzel ML, Morken MA, Huyghe JR, Fuchsberger C, Kuusisto J, et al. A Common Functional Regulatory Variant at a Type 2 Diabetes Locus Upregulates ARAP1 Expression in the Pancreatic Beta Cell. *Am J Hum Genet* (2014) 94:186–97. doi: 10.1016/j.ajhg.2013.12.011
77. Carrat GR, Hu M, Nguyen-Tu M-S, Chabosseau P, Gaulton KJ, Van De Bunt M, et al. Decreased STARD10 Expression Is Associated With Defective Insulin Secretion in Humans and Mice. *Am J Hum Genet* (2017) 100:238–56. doi: 10.1016/j.ajhg.2017.01.011
78. Patterson R, Mcnamara E, Tainio M, De Sá TH, Smith AD, Sharp SJ, et al. Sedentary Behaviour and Risk of All-Cause, Cardiovascular and Cancer Mortality, and Incident Type 2 Diabetes: A Systematic Review and Dose Response Meta-Analysis. *Eur J Epidemiol* (2018) 33:811–29. doi: 10.1007/s10654-018-0380-1
79. Guo C, Zhou Q, Zhang D, Qin P, Li Q, Tian G, et al. Association of Total Sedentary Behaviour and Television Viewing With Risk of Overweight/Obesity, Type 2 Diabetes and Hypertension: A Dose-Response Meta-Analysis. *Diabetes Obes Metab* (2020) 22:79–90. doi: 10.1111/dom.13867
80. Scandiffio JA, Janssen I. Do Adolescent Sedentary Behavior Levels Predict Type 2 Diabetes Risk in Adulthood? *BMC Public Health* (2021) 21:969. doi: 10.1186/s12889-021-10948-w
81. Li C, Beech B, Crume T, D'agostino RBJr., Dabelea D, Kaar JL, et al. Longitudinal Association Between Television Watching and Computer Use and Risk Markers in Diabetes in the SEARCH for Diabetes in Youth Study. *Pediatr Diabetes* (2015) 16:382–91. doi: 10.1111/pedi.12163
82. Lee IM, Shiroma EJ. Using Accelerometers to Measure Physical Activity in Large-Scale Epidemiological Studies: Issues and Challenges. *Br J Sports Med* (2014) 48:197–201. doi: 10.1136/bjsports-2013-093154
83. Pedišić Ž, Bauman A. Accelerometer-Based Measures in Physical Activity Surveillance: Current Practices and Issues. *Br J Sports Med* (2015) 49:219–23. doi: 10.1136/bjsports-2013-093407
84. Meisinger C, Linseisen J, Leitzmann M, Baurecht H, Baumeister SE. Association of Physical Activity and Sedentary Behavior With Type 2 Diabetes and Glycemic Traits: A Two-Sample Mendelian Randomization Study. *BMJ Open Diabetes Res Care* (2020) 8. doi: 10.1136/bmjdrc-2020-001896
85. Vanderweele TJ, Vansteelandt S. Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *Am J Epidemiol* (2010) 172:1339–48. doi: 10.1093/aje/kwq332

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Endocrinology

ORIGINAL RESEARCH
published: 22 July 2022
doi: 10.3389/fendo.2022.905367

The Global Burden of Type 2 Diabetes Attributable to Tobacco: A Secondary Analysis From the Global Burden of Disease Study 2019

Jianjun Bai[1], *Fang Shi*[1], *Yudiyang Ma*[1], *Donghui Yang*[1], *Chuanhua Yu*[1,2*] *and Jinhong Cao*[1*]

[1] Department of Epidemiology and Biostatistics, School of Public Health, Wuhan University, Wuhan, China,
[2] Global Health Institute, Wuhan University, Wuhan, China

OPEN ACCESS

Edited by:
Changwei Li,
Tulane University School of Public Health and Tropical Medicine, United States

Reviewed by:
Yan-Bo Zhang,
Huazhong University of Science and Technology, China
János Sándor,
University of Debrecen, Hungary

***Correspondence:**
Chuanhua Yu
yuchua@whu.edu.cn
Jinhong Cao
caojhky2007@163.com

Specialty section:
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

Received: 27 March 2022
Accepted: 20 June 2022
Published: 22 July 2022

Citation:
Bai J, Shi F, Ma Y, Yang D, Yu C and Cao J (2022) The Global Burden of Type 2 Diabetes Attributable to Tobacco: A Secondary Analysis From the Global Burden of Disease Study 2019. Front. Endocrinol. 13:905367. doi: 10.3389/fendo.2022.905367

Objectives: Growing epidemiological studies have reported the relationship between tobacco and health loss among patients with type 2 diabetes (T2D). This study aimed to explore the secular trend and spatial distribution of the T2D burden attributable to tobacco on a global scale to better understand regional disparities and judge the gap between current conditions and expectations.

Methods: As a secondary analysis, we extracted data of tobacco-attributable T2D burden from the 2019 Global Burden of Disease Study (GBD). Joinpoint regression was adopted to determine the secular trend of age-standardized rates (ASR), with average annual percentage change (AAPC). Gaussian process regression (GPR) was used to explore the average expected relationship between ASRs and the socio-demographic index (SDI). Spatial autocorrelation was used to indicate if there is clustering of age-standardized DALY rate (ASDR) with Moran's I value. Multi-scale geographically weighted regression (MGWR) was to investigate the spatial distribution and scales of influencing factors in ASDR attributable to tobacco, with the regression coefficients for each influencing factor among 204 countries.

Results: Tobacco posed a challenge to global T2D health, particularly for the elderly and men from lower SDI regions. For women, mortality attributable to secondhand smoke was higher than smoking. A downward trend in age-standardized mortality rate (ASMR) of T2D attributable to tobacco was observed (AAPCs= -0.24; 95% CI -0.30 to -0.18), while the ASDR increased globally since 1990 (AAPCs= 0.19; 0.11 to 0.27). Oceania, Southern Sub-Saharan Africa, and Southeast Asia had the highest ASMRs and ASDRs, exceeding expectations based on the SDI. Also, "high-high" clusters were mainly observed in South Africa and Southeast Asian countries, which means a high-ASDR country is surrounded by high-ASDR neighborhoods in the above areas. According to MGWR model, smoking prevalence was the most sensitive influencing factor, with regression coefficients from 0.15 to 1.80.

Conclusion: The tobacco-attributable burden of T2D should be considered as an important health issue, especially in low-middle and middle-SDI regions. Meanwhile, secondhand smoke posed a greater risk to women. Regional disparities existed, with hot spots mainly concentrated in South Africa and Southeast Asian countries.

Keywords: type 2 diabetes, tobacco, mortality, world regions, spatial autocorrelation, DALYs - disability-adjusted life years

INTRODUCTION

Diabetes mellitus, characterized by elevated blood glucose, have been one of the leading causes of disease burden worldwide. According to the IDF Diabetes Atlas 10th edition, 537 million individuals (20-79 years of age) are suffering from diabetes, 6.7 million deaths were attributed to diabetes in 2021, and by 2045, there would be 783 million people with diabetes. Type 2 diabetes (T2D), caused by insulin resistance and β cell dysfunction, accounts for around 90% of diabetes cases and leads to a heavy burden on individuals and health systems (1). A 15% increased risk of premature death and an approximately 20-year reduction in life expectancy were observed in patients with T2D (2). Since T2D is closely related to behavioral and metabolic risk factors, avoiding known determinants through early behavioral modification is the most cost-efficient strategy to reduce disease burden.

Growing epidemiological studies have reported that tobacco smoking is one of the most important modifiable risk factors for T2D (3). According to the GBD 2019, smoking is the third leading risk factor for T2D burden worldwide, with 9.9% of the T2D burden attributable to smoking (4). The biological mechanism may involve oxidative stress, inflammatory markers, and glucose metabolism irregularities (5, 6). The 2014 Surgeon General's Report, for the first time, inferred causal association between smoking and T2D, as well as the potential dose-response relationship (7). A meta-analysis of 22 prospective studies (16,383 patients with T2D) revealed a similar association between smoking and T2D, with a pooled RR of 1.38 for T2D in current smokers compared to never smokers (8). A study in Australia found that smoking and diabetes are associated with an increased risk of mortality and micro- and macrovascular complications, which is intensified when combined (9). Besides, secondhand smoking has been a global health problem and more likely to occur indoors at work or home with active smokers. Secondhand smoke is considered to be associated with an increased risk of T2D (RR=1.22) (3). A national study in China also showed the positive relationship between secondhand smoke and T2D risk (10). Besides, the ultimate purpose of glycemic control is to prevent microvascular and macrovascular complications. Studies have shown that tobacco use is positively associated with important diabetic complications, including cardiovascular disease (CVD), neuropathy, nephropathy, and retinopathy (11). Meanwhile, the association between smoking and diabetes and CVD is well established. Smoking and diabetes interacted with each other in relation to increased risk of CVD events (12, 13). Thus, optimal management of tobacco use and control of blood glucose levels are essential to prevent diabetic complications, while also contributing to a reduction in the burden of cardiovascular disease in the whole population (14).

Given the increasing diabetes burden worldwide, the United Nations adopted Sustainable Development Goals (SDGs) target 3.4 and a series of measures to achieve a 30% reduction of premature mortality from non-communicable diseases (NCDs) by 2030 globally, including diabetes (15). The World Health Organization, the American Diabetes Association Guidelines, and the Italian Diabetes Clinical Guidelines all consider smoking as a preventable risk factor for T2D and support smoking cessation as one of the most important steps in preventing diabetes complications (16–18). Many developing countries such as China, Brazil, and South Africa, have also made great efforts on tobacco control, although progress varied substantially (19, 20). In clinical practice, physicians usually adopted smoking cessation as a basic intervention for long-term care of patients with diabetes. Whereas, some studies showed no significant changes in smoking rates among the diabetic population, comparable to the non-diabetic population (9, 21, 22). Hence, to judge the gap between current conditions and SDGs, it is essential to explore the spatial distribution disparities of tobacco-attributable T2D burden and evaluate the secular trends over the recent period on a global scale, especially among countries or regions with different social-economic levels. To our knowledge, there is no study available giving similar trends on a global scale.

In this study, we investigated the burden of T2D attributable to tobacco on a global scale, examined the secular trends by Joinpoint regression analysis, and explored the spatial distribution disparities through geographic analysis.

Abbreviations: T2D, type 2 diabetes; GBD, Global Burden of Diseases; DALYs, disability-adjusted life years; ASMR, age-standardized mortality rate; ASDR, age-standardized DALY rate; SDI, socio-demographic index; ASSP, age-standardized smoking prevalence; DTI, diabetes treatment index; ASR, Age-standardized rate; APC, annual percentage change; AAPC, the average annual percentage change; GPR, Gaussian Process Regression; MGWR, Multi-scale geographically weighted regression.

MATERIALS AND METHODS

Data Sources

The GBD 2019 provided a systematic and comprehensive annual assessment of 369 diseases and injuries, 87 behavioral, environmental, occupational, and metabolic risk factors among

204 countries or territories from 1990 to 2019. The reliability of the GBD data have been confirmed previously (23–25). The diabetic count reported by the GBD 2019 (460 million) was similar to that of the International Diabetes Federation 2019 (463 million) (26). As a secondary analysis of the GBD, the data of T2D burden attributable to tobacco and related influencing factors were extracted from the GBD 2019, including deaths, DALYs, age-standardized mortality rate (ASMR), age-standardized DALY rate (ASDR), socio-demographic index (SDI), age-standardized smoking prevalence (ASSP) and diabetes treatment index (DTI) among 204 countries or territories and 21 GBD regions. Age-standardized rates (ASR) were calculated by the GBD 2019 global standard population to eliminate the impact of age structure and population differences.

The DTI was used to evaluate the access and quality of diabetes care for a given set of interventions or services, varying from 0 to 100 (27). The SDI was considered a good indicator to reflect the health-related socio-economic developments. It is a composite indicator of lag-distributed income per capita, mean education for those aged 15 and older, and total fertility rate under 25, ranging from 0 to 1 (**Supplementary Material**) (23, 28). Based on the SDI, the 204 countries or territories were divided into five quintiles: low (< 0.46), low-middle (0.46-0.60), middle (0.61-0.69), high-middle (0.70-0.81), and high (> 0.81) SDI regions.

Our study was based on the publicly available the GBD database (GHDx). All data were publicly accessible online at (http://ghdx.healthdata.org/gbd-results-tool). Therefore, ethical approval is not applicable to our study.

Case Definition

The detailed methodology of the GBD 2019 has been described previously (23, 25, 28, 29). Briefly, to estimate all-cause mortality, cause-specific mortality, and YLLs, the GBD studies utilized standardized data identification, extraction, and processing methods to address data incompleteness, discrepancies in coding practices, and inconsistent age group and sex reports (30). In the GBD 2019, overall diabetes mellitus mortality was estimated using deaths directly attributed to diabetes mellitus. T2D deaths were defined by codes E11-E11.1, E11.3-E11.9 based on the Tenth Revision of the International Classification of Diseases. The GBD 2019 used a Bayesian hierarchical Cause of Death Ensemble model (CODEm) platform to build the best-fitted model with appropriate country-level covariates and analyze 20,830 site-years of vital registration data, and 448 site-years of sample-based vital registration data to estimate T2D mortality. The CoDCorrect process was then conducted to ensure that the cause-specific mortality and all-cause mortality estimates were internally consistent. YLLs were calculated by multiplying deaths by the residual life expectancy at the age of death based on the GBD 2019 reference life table (23, 25).

Nonfatal estimates were generated using data from the systematic literature search, hospital discharge, claims systems, household surveys, cohort studies, and disease registries (23, 30). T2D was defined as "fasting plasma glucose (FPG) ≥ 126 mg/dL (7 mmol/L) or reporting to be on drug or insulin treatment for type 2 diabetes" in the GBD 2019, and the corresponding sequelae were described as well (generic uncomplicated disease, diabetic neuropathy, vision impairment, etc.). Meanwhile, to ensure comparability of data across data sources, the GBD 2019 used MR-BRT analysis for bias adjustment methods to allow a more direct comparison between alternative case definitions (like HbA1c, OGTT, claims data) and/or study designs (23). A compartmental meta-regression tool, DisMod-MR2.1, was then used to synthesize all available data sources to produce internally consistent prevalence estimates. The YLD for each sequela was obtained by multiplying the prevalence and sequela-specific disability weight. After comorbidity correction, the sum of the YLDs for each general sequel denoted the total YLDs for T2D. DALYs were the sum of YLLs and YLDs, referring to all healthy life years lost from onset to death. The GBD generated 95% uncertainty intervals (UI) for all reported data based on the 25th and 975th ordered values of 1,000 draws of the posterior distribution.

Meanwhile, according to the GBD 2019, the burden of T2D attributable to tobacco was divided into two parts: smoking and secondhand smoke. Smoking case definitions were former and current smoking of any tobacco product daily or occasionally. Secondhand smoke was defined as current exposure to secondhand smoke at home, work, or other public places (28). Based on the GBD comparative risk assessment framework, population attributable fractions (PAFs) were used to quantify what proportion of disease burden in a specific population would be reduced if the exposure of certain causal factors were reduced to the theoretical minimum risk exposure level (TMREL) (28, 31). PAFs of disease outcomes were estimated based on exposure data, relative risk of outcomes, and the TMREL. Population surveys were the primary source of exposure data on smoking and secondhand smoke (28). Relative risks were derived from meta-analyses of cohort and case-control studies. The TMRELs for smoking and secondhand smoke were defined as zero. T2D burden attributable to tobacco was calculated by multiplying relevant PAFs by the T2D overall burden for each age-sex-location-year (31).

Statistical Analysis

We adopted the Joinpoint regression model to determine the secular trend of age-standardized rates. The Joinpoint regression model mainly uses the Grid search method to analyze and establish all possible change points and selects the points with small mean squared errors (MSE) as the joinpoints, which divided the overall trend into several segments (32). The annual percentage change (APC) for each segment, the average annual percentage change (AAPC) for overall trend, and 95% confidence intervals (CIs) were estimated by the Joinpoint model:

$$In(ASR) = \alpha + \beta_i x + \epsilon$$

$$APC = \left(e^{\beta_i} - 1\right) \times 100$$

$$AAPC = \left(\exp\frac{\sum w_i\beta_i}{\sum w_i} - 1\right) \times 100$$

Where x represents the calendar year, β_i represent the segmental regression coefficients, and w_i is the interval span of the different segments.

The average relationship between ASRs and SDI was calculated using Gaussian process regression (GPR) model. Gaussian processes are the basic principle behind GPR. Instead of inferring a distribution over the parameters of a parametric function, Gaussian processes could infer a distribution over functions directly, which defines a prior over functions. After having observed some function values, it can be converted into a posterior over functions. The general form is as follow:

$$p(f|\mathrm{X}) = N(f|\mu, \mathrm{K})$$

where f=($f(x_1)$,…, ($f(xn)$), μ=($m(x_1)$,…,$m(x_n)$), K is the kernel function. m is the mean function and it is common to use $m(x)$ =0. Thus, kernel function is the important part of GPR. We chose the classical "Radial Basis kernel function" to conduct the GPR with 10-fold cross validation by the "Kernlab" package of R software. Observed values are the actual disease burden rates in each location-year, while expected values were determined by GPR on the range of rates observed for each level of the SDI (30). We used these estimates of expected ASRs that were predicted based on the full range of the SDI to determine whether observed health patterns deviated from trends associated with changes along the socio-economic development spectrum (33). The associations between expected ASRs and the SDI were explored using the Pearson correlation analysis (30, 34, 35).

In addition, we adopted spatial autocorrelation to explore the spatial distribution of ASDR. Global spatial autocorrelation is used to determine whether the ASDR has aggregation characteristics in overall space, with the Global Moran's I, ranging from -1 to 1. Global Moran's I > 0 indicates similar values cluster together in a map; I < 0 indicates dissimilar values cluster together in a map; I = 0 indicates no spatial correlation (36). The calculation formula is as follows:

$$Global\ \ Moran's\ \ I = \frac{k\sum_{i=1}^{k}\sum_{j=1}^{k} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\left(\sum_{i=1}^{k}\sum_{j=1}^{k} w_{ij}\right)\sum_{i=1}^{k}(x_i - \bar{x})^2}$$

where $i \neq j$, k is the number of spatial units involved in the analysis; x_i and x_j represent the observation values of a certain factor x in spatial units i and j, respectively; $\bar{x}$ represents the average value of the attribute value, and w_{ij} is the spatial weight matrix, calculated by the queen contiguity weight matrix in GeoDa software.

Also, local spatial autocorrelation can locate the extent of spatial hot spots using the LISA cluster map. In LISA cluster map, "high" represents the ASDR being higher than the global average level, "low" represents the ASDR being lower than the average level. "High-high" cluster means a high-ASDR country is surrounded by high-ASDR neighborhoods, while "low-low" cluster means a low-ASDR country is surrounded by low-ASDR neighborhoods.

Furthermore, different processes can operate at different spatial scales and the impact of tobacco on T2D burden may be determined not only by global socio-economic development but also by local specific conditions, such as smoking rates and T2D treatment levels (37). We adopted multi-scale geographically weighted regression (MGWR) to explore the spatial distribution and scales of influencing factors in ASDRs among 204 countries in 2019. The traditional GWR model applies a constant bandwidth to illustrate the impact scales of different factors, which ignores the diversity of impact scales and does not align with the facts (38). MGWR is an extension of the GWR model that allows for exploring the associations at varying spatial scales and achieves that by using a varying bandwidth rather than a single, and constant bandwidth for the entire area, so as to provide more credible estimation results and the diverse impact scales of each factor (39). The calculation formula is as follows:

$$yi = \sum_{j=1}^{k}\beta_{bwj}(u_i, v_i)x_{ij} + \epsilon_i$$

where (u_i,v_i) is the spatial location of the i-th country, yi is the response variable, x_{ij} is the j-th explanatory variable, $\beta_{bwj}(u_i,v_i)$ is the j-th coefficient, and bwj in β_{bwj} is the bandwidth used by the regression coefficient of the j-th variable.

In practice, MGWR is usually regarded as a generalized additive model (below), thus it is possible to calibrate the model using the back-fitting algorithm with the classical GWR being the initial estimator (37, 40).

$$y = \sum_{j=1}^{k} f_j + \epsilon \ \ \left(f_j = \beta_{bwj}x_j\right)$$

where f_j is a smoothing function applied to the j-th explanatory variable. Spatial kernel and bandwidth selection criteria for MGWP were following Adaptive Bisquare and AICc principles. After collinearity analysis, we took SDI, ASSP, and DTI into account for the final MGWR model. The selected variables have relatively low multi-collinearity (all VIFs < 2.0).

Joinpoint regression model was conducted by Joinpoint program (version 4.8.0.1); Pearson correlation and Gaussian Process Regression were performed using R software (version 4.0.2); GeoDa and MGWP were for geographic analysis; ArcGIS for mapping. Detailed information on the above methods can be found in the **Supplementary Materials**. Two-sided P < 0.05 was considered to be statistically significant.

RESULTS

Global T2D Burden Attributable to Tobacco Since 1990

In 2019, T2D caused 1.47 (95% UI 1.37 to 1.57) million deaths and 66.30 (55.48 to 79.01) million DALYs worldwide, accounting for 94.96% and 93.54% of those in diabetes mellitus. An estimated 235.43 (163.51 to 299.27) thousand deaths and 11.86 (7.93 to 16.01) million DALYs of T2D were contributed to tobacco.

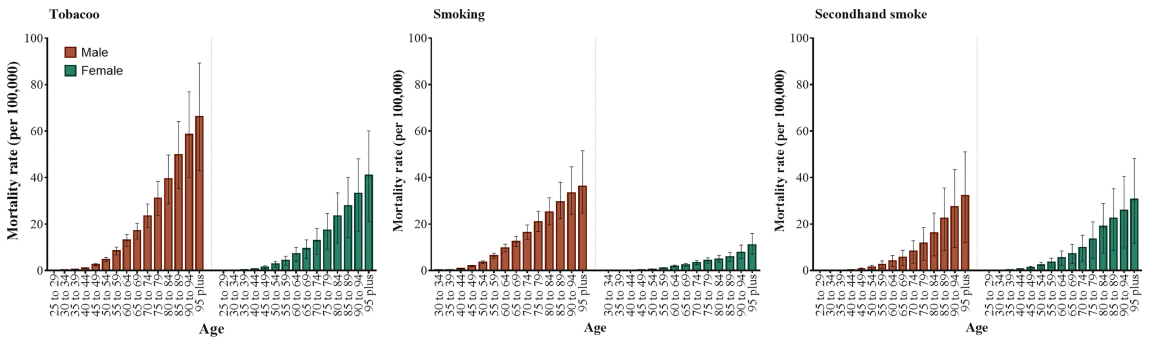


FIGURE 1 | Mortality of T2D attributable to tobacco, smoking, and secondhand smoke by age group and gender. Error bars represent the 95% UIs of mortality.

As shown in **Figure 1**, the mortality of T2D attributable to tobacco increased with age in both genders, and that of men was higher than women. The mortality trends were similar in smoking and secondhand smoke for men. For women, the mortality attributable to secondhand smoke was significantly higher than smoking.

Table 1 showed ASMR, ASDR, and AAPC of T2D attributed to tobacco from 1990 to 2019 by region. The gender disparities still existed, and the ASMR and ASDR (per 100,000) attributable to tobacco in men were higher than women (3.82 VS 2.14; 183.13 VS 104.47). We observed a declining trend of ASMR attributable to tobacco (AAPCs= -0.24; 95% CI -0.30 to -0.18), while the ASDR remained an increasing trend since 1990 globally (AAPCs= 0.19; 0.11 to 0.27). The ASMR and ASDR of T2D attributable to tobacco were 2.91 (2.01 to 3.70) and 142.00 (94.81 to 191.85) per 100,000 globally in 2019, respectively.

T2D Burden Attributable to Tobacco Since 1990 by SDI Regions

The ASRs of T2D attributable to tobacco remained the highest among low-middle and middle SDI regions over the study period. In 2019, the highest ASMR and ASDR (per 100,000) were observed in the low-middle SDI region, 4.47 (2.96 to 5.86) and 173.32 (110.43 to 237.06), respectively. The ASRs in the high SDI region remained the lowest. In 2019, the lowest ASMR and ASDR (per 100,000) were 1.31 (1.02 to 1.62) and 105.13 (73.18 to 142.87), respectively (**Table 1**).

Between 1990 and 2019, except for low-middle and middle SDI regions, the declining trend of ASMR attributable to tobacco were observed across the other SDI regions, especially after 2015 (**Figure 2**). High SDI regions showed the greatest decline in ASMR attributable to tobacco (AAPC= -1.81; -1.94 to -1.68). While in recent years, the decline in ASMR has stagnated among five SDI regions. Since 1990, low, low-middle, and middle SDI regions have shown rising trends in ASDR attributable to tobacco. Nevertheless, the declining trend was observed in the high SDI region from 1990 to 2019, with AAPC of -0.22 (-0.34 to -0.10). Also, all five SDI regions showed increasing trends of ASDR after 2017 (**Figure 2**). For smoking, similar trends were observed in ASRs. For secondhand smoke, all five SDI regions exhibited a consistent increase in ASDR, particularly in the low-middle SDI regions (**Supplementary Figure 1**).

T2D Burden Attributable to Tobacco Since 1990 by GBD Regions

Geographic differences existed among 21 GBD regions. The highest ASMR and ASDR (per 100,000) attributable to tobacco occurred in Oceania (24.86 and 823.21), followed by Southern Sub-Saharan Africa (9.72 and 290.68) and Southeast Asia (7.49 and 263.72). Besides, Australasia, high-income Asia Pacific and Eastern Europe were regions with the lowest ASMR (0.85, 0.65 and 0.94) and ASDR (58.55, 70.99 and 75.04). Central Asia showed the most noticeable growth in ASMR and ASDR over the past decades, with AAPCs of 2.96 (2.61 to 3.32) and 2.44 (2.19 to 2.70), respectively. The most significant decrease in ASMR was observed in high-income Asia Pacific (AAPCs= -3.17; -3.44 to -2.90), the most significant reduction in ASDR was in tropical Latin America (AAPCs= -1.94; -2.19 to -1.68), followed by high-income Asia Pacific. The three-segment trends of ASRs among 21 GBD regions are shown in **Supplementary Table 1**.

The estimated relationship between the SDI and the expected ASRs of T2D attributable to tobacco was highly positive when the SDI was < 0.40 and highly negative when the SDI was > 0.40 (**Figure 3**). The expected ASRs based on GPR are shown as the blue line in **Figure 3**. The regions above the blue line for ASRs represent a lag behind expected improvements in T2D burden attributable to tobacco. Over the study period, the regions with much higher-than-expected ASRs included Oceania, Southern Sub-Saharan Africa, Southeast Asia, and the Caribbean. While among most other regions, improvements in T2D burden outpaced what would have been expected based on SDI improvements alone, especially in East Asia, Eastern Europe, Andean Latin America, and Australasia.

Spatial Autocorrelation and MGWR by Countries

Spatial autocorrelation and MGWR were conducted to explore the spatial distribution patterns and influencing factors of the T2D burden attributable to tobacco. Global Moran's I values demonstrated that positive autocorrelations of tobacco-attributable T2D ASDRs existed (all P < 0.05; **Supplementary**

Table 2). Meanwhile, according to local Moran's I index and LISA clustering map, "high-high" clusters were mainly observed in South Africa, the Kingdom of Saudi Arabia, and Southeast Asian countries, which were the main lower SDI countries (**Supplementary Figure 2**).

The R^2 (0.846) and Adj-R^2 (0.819) of MGWR was higher, and the AICc value was lower compared with GWR model (**Supplementary Table 3**), indicating the MGWR result was more reliable. MGWR results suggested that the influencing scale of different factors varied greatly. ASDRs attributable to tobacco were sensitive to the ASSP factor and spatial heterogeneity existed. The influencing scale of ASSP was 43, accounting for 21.08% of the 204 countries or territories, which was close to the sub-continent scale. Once beyond this scale, the coefficient would change. The influencing scale of the SDI was 195, a global scale, suggesting the influence of the SDI on space is relatively stable.

Figure 4 shows the mapping of coefficients of MGWR and corresponding P-value for the selected factors. The SDI had a positive effect on ASDR of T2D, but the effect was relatively weak,

TABLE 1 | ASMR, ASDR, and AAPC in type 2 diabetes attributed to tobacco from 1990 to 2019 by regions.

Location	ASMR per 100,000 (95% UI)			ASDR per 100,000 (95% UI)		
	1990	2019	AAPC (95% CI)	1990	2019	AAPC (95% CI) 1990-2019
Global	3.10 (2.25 to 3.83)	2.91 (2.01 to 3.70)	-0.24* (-0.30 to -0.18)	132.69 (93.52 to 172.97)	142.00 (94.81 to 191.85)	0.19* (0.11 to 0.27)
Sex						
Male	3.96 (3.20 to 4.74)	3.82 (2.98 to 4.62)	-0.14* (-0.20 to -0.07)	167.52 (129.34 to 209.19)	183.13 (135.76 to 236.02)	0.25* (0.16 to 0.34)
Female	2.45 (1.47 to 3.29)	2.14 (1.12 to 3.03)	-0.48* (-0.55 to -0.42)	102.44 (58.55 to 143.68)	104.47 (54.42 to 153.68)	0.03 (-0.09 to 0.15)
SDI regions						
Low	3.72 (2.42 to 4.91)	3.66 (2.36 to 4.90)	-0.07* (-0.15 to 0.00)	123.83 (81.45 to 164.99)	135.63 (85.21 to 184.83)	0.27* (0.21 to 0.34)
Low-middle	4.28 (2.93 to 5.53)	4.47 (2.96 to 5.86)	0.19 (-0.04 to 0.42)	150.82 (103.56 to 197.01)	173.32 (110.43 to 237.06)	0.42* (0.30 to 0.54)
Middle	3.98 (2.70 to 5.05)	3.97 (2.64 to 5.11)	0.04 (-0.06 to 0.14)	156.36 (105.05 to 206.67)	168.40 (111.25 to 226.88)	0.22* (0.11 to 0.33)
High-middle	2.67 (1.93 to 3.27)	2.15 (1.50 to 2.72)	-0.74* (-0.83 to -0.66)	123.55 (86.86 to 163.25)	124.06 (84.42 to 169.33)	-0.07 (-0.22 to 0.08)
High	2.18 (1.72 to 2.65)	1.31 (1.02 to 1.62)	-1.81* (-1.94 to -1.68)	112.62 (83.24 to 145.62)	105.13 (73.18 to 142.87)	-0.22* (-0.34 to -0.10)
GBD regions						
Andean Latin America	2.11 (1.36 to 2.83)	1.96 (1.19 to 2.77)	-0.17 (-0.45 to 0.12)	77.89 (50.89 to 105.84)	80.1 (49.93 to 112.93)	0.08 (-0.04 to 0.20)
Australasia	1.71 (1.34 to 2.08)	0.85 (0.60 to 1.09)	-2.47* (-2.98 to -1.96)	71.42 (53.33 to 92.12)	58.55 (39.05 to 81.77)	-0.66* (-0.76 to -0.57)
Caribbean	5.67 (4.00 to 7.18)	4.07 (2.79 to 5.56)	-1.22* (-1.52 to -0.91)	230.82 (161.40 to 303.76)	201.24 (135.24 to 274.59)	-0.45* (-0.65 to -0.24)
Central Asia	1.66 (1.12 to 2.10)	3.85 (2.67 to 4.97)	2.96* (2.61 to 3.32)	90.05 (59.50 to 122.23)	179.57 (120.4 to 241.27)	2.44* (2.19 to 2.70)
Central Europe	2.57 (1.97 to 3.11)	2.21 (1.60 to 2.86)	-0.54* (-0.69 to -0.39)	147.20 (106.17 to 193.7)	163.66 (112.37 to 225.12)	0.32* (0.25 to 0.40)
Central Latin America	8.06 (5.49 to 10.22)	5.34 (3.39 to 7.26)	-1.42* (-1.72 to -1.12)	306.36 (207.39 to 402.06)	225.79 (141.05 to 315.8)	-1.03* (-1.25 to -0.81)
Central Sub-Saharan Africa	3.89 (2.59 to 5.19)	3.03 (1.96 to 4.25)	-0.90* (-1.00 to -0.80)	128.07 (88.24 to 171.39)	117.00 (76.36 to 162.99)	-0.32* (-0.39 to -0.24)
East Asia	2.12 (1.44 to 2.74)	2.06 (1.46 to 2.67)	-0.03 (-0.35 to 0.28)	112.39 (76.35 to 151.87)	115.87 (79.07 to 158.04)	-0.11 (-0.62 to 0.40)
Eastern Europe	0.68 (0.46 to 0.87)	0.94 (0.66 to 1.21)	1.61* (0.67 to 2.55)	55.67 (37.66 to 77.51)	75.04 (51.75 to 103.78)	1.11* (0.51 to 1.72)
Eastern Sub-Saharan Africa	4.34 (2.95 to 5.68)	3.43 (2.33 to 4.69)	-0.84* (-0.91 to -0.77)	126.44 (85.23 to 167.16)	108.83 (72.58 to 148.09)	-0.54* (-0.59 to -0.49)
High-income Asia Pacific	1.62 (1.27 to 1.94)	0.65 (0.51 to 0.80)	-3.17* (-3.44 to -2.90)	94.86 (69.93 to 123.77)	70.99 (48.23 to 99.14)	-1.10* (-1.28 to -0.91)
High-income North America	2.67 (2.06 to 3.32)	1.80 (1.40 to 2.24)	-1.41* (-1.61 to -1.21)	145.15 (105.76 to 190.91)	130.13 (92.51 to 173.29)	-0.48* (-0.68 to -0.27)
North Africa and Middle East	5.06 (3.37 to 6.58)	4.37 (2.93 to 5.72)	-0.52* (-0.64 to -0.40)	176.65 (119.64 to 232.91)	203.48 (132.64 to 278.18)	0.53* (0.43 to 0.63)
Oceania	19.85 (13.09 to 27.34)	24.86 (15.94 to 34.86)	0.72* (0.61 to 0.83)	629.55 (424.26 to 839.13)	823.21 (545.77 to 1124.19)	0.88* (0.79 to 0.97)

(Continued)

TABLE 1 | Continued

Location	ASMR per 100,000 (95% UI)			ASDR per 100,000 (95% UI)		
	1990	2019	AAPC (95% CI)	1990	2019	AAPC (95% CI) 1990-2019
South Asia	4.06 (2.58 to 5.47)	4.24 (2.60 to 5.73)	0.16 (-0.29 to 0.61)	139.01 (91.33 to 186.89)	164.77 (99.68 to 229.89)	0.50* (0.34 to 0.66)
Southeast Asia	7.13 (4.89 to 9.13)	7.49 (4.96 to 9.67)	0.15* (0.04 to 0.26)	229.12 (158.33 to 294.82)	263.72 (177.49 to 346.34)	0.47* (0.34 to 0.59)
Southern Latin America	4.10 (2.89 to 5.14)	3.07 (2.09 to 3.93)	-1.02* (-1.18 to -0.86)	144.12 (102.76 to 186.19)	148.36 (101.28 to 202.52)	0.11* (0.01 to 0.20)
Southern Sub-Saharan Africa	9.04 (6.35 to 11.65)	9.72 (6.16 to 13.03)	0.39 (-0.18 to 0.96)	276.14 (194.84 to 352.02)	290.68 (186.92 to 386.16)	0.29 (-0.15 to 0.74)
Tropical Latin America	7.15 (5.32 to 8.75)	3.85 (2.64 to 5.05)	-2.15* (-2.30 to -2.00)	270.15 (200.07 to 341.72)	156.24 (105.88 to 210.38)	-1.94* (-2.19 to -1.68)
Western Europe	2.20 (1.73 to 2.67)	1.10 (0.84 to 1.36)	-2.42* (-2.50 to -2.33)	101.04 (74.35 to 131.73)	95.45 (64.31 to 133.5)	-0.22* (-0.26 to -0.17)
Western Sub-Saharan Africa	2.62 (1.53 to 3.69)	2.81 (1.66 to 3.88)	0.27* (0.19 to 0.34)	78.52 (46.56 to 109.44)	88.58 (53.35 to 125.72)	0.43* (0.37 to 0.48)

ASMR, age-standardized mortality rate; ASDR, age-standardized DALY rate; SDI, socio-demographic index; AAPC, the average annual percentage change; UI, uncertainty interval; CI, confidence interval.

** means that AAPCs (95% CIs) with "*" represent significance at P < 0.05.*

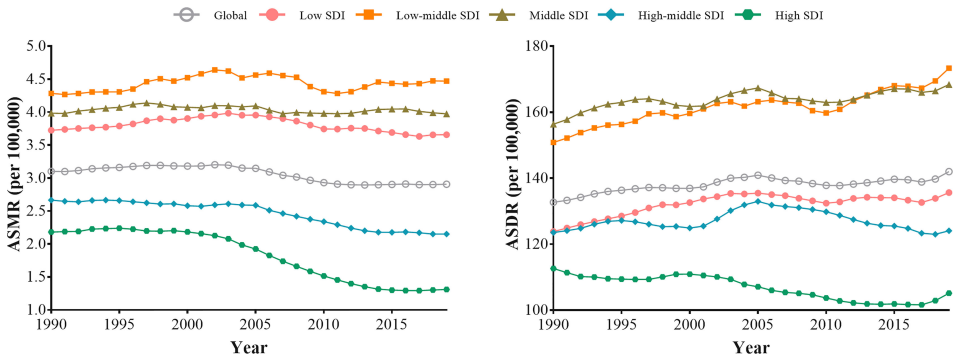


FIGURE 2 | The trends in the ASMR and ASDR for type 2 diabetes attributable to tobacco globally and among five SDI regions from 1990 to 2019.

with coefficients from 0.17 to 0.33. Furthermore, a clear directional feature of regression coefficients of the SDI was observed extending in the west-east direction. Also, negative regressions were observed between DTI and ASDRs, ranging from -0.87 to -0.13. DTI was a decisive factor in explaining the ASDR of T2D across North America, South America, and Australia. Additionally, positive regression coefficients were found for ASSP, varying from 0.15 to 1.80. Higher correlations were concentrated in countries from South America (Brazil, Ecuador, and Columbia). Meanwhile, the highest absolute value of coefficient suggested ASSP was the main influencing factor.

DISCUSSION

In the global analysis of the T2D burden attributable to tobacco, we found a declining trend in the global ASMR attributable to tobacco, whereas the ASDR remained an increasing trend since 1990 globally. The tobacco-attributable T2D burden posed great challenges to the elderly and men from low-middle and middle SDI regions. While secondhand smoke posed a greater risk for women. Meanwhile, "high-high" clusters were mainly observed in South Africa, the Kingdom of Saudi Arabia, and Southeast Asian countries.

Gender disparities were observed and the burden of T2D attributed to tobacco is more remarkable for men, partly because the smoking prevalence in men is higher and men are more likely to be exposed to secondhand smoke (41). Additionally, insulin resistance is more likely to occur in men (42), while estrogen can affect enzyme (CYP2A6) activity and promote nicotine metabolism for women (43). Meanwhile, smoking can also raise serum levels of heavy metals like lead, arsenic, and cadmium, which may affect glucose homeostasis and increase the risk of T2D (6, 44). Secondhand smoke, such as sidestream smoke, is mostly unfiltered and more likely to occur indoors at work or home, and its toxicity increases as it remains in the air for minutes or hours (45). Previous prospective studies have

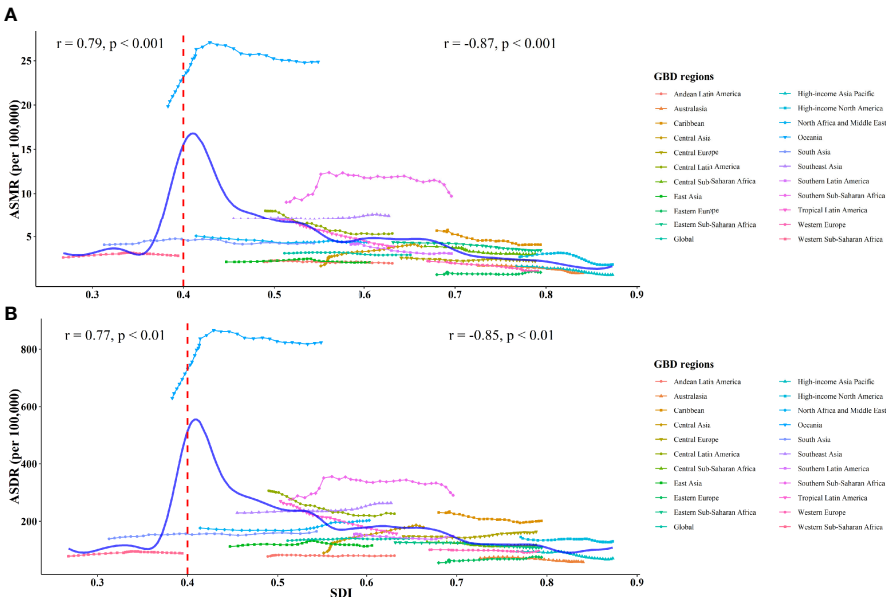


FIGURE 3 | ASMR **(A)** and ASDR **(B)** of type 2 diabetes attributed to tobacco globally and among 21 GBD regions from 1990 to 2019. The solid blue line represents the expected value based on the SDI. The Pearson correlation coefficients and P-values are displayed.

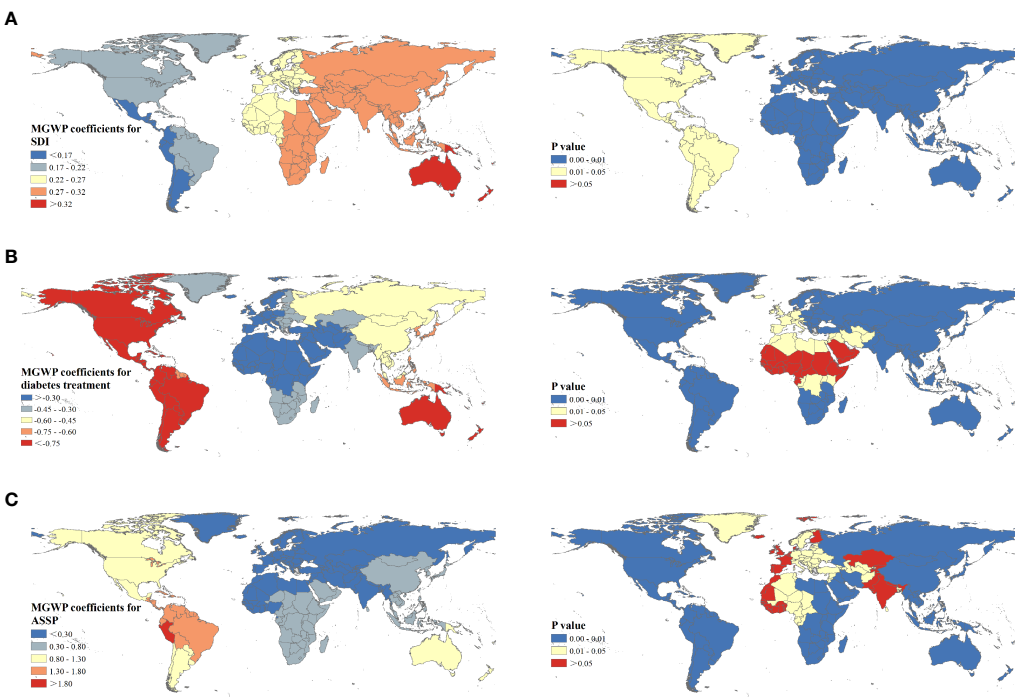


FIGURE 4 | Spatial patterns of regression coefficients and corresponding P values for socio-demographic index **(A)**, diabetes treatment index **(B)**, and age-standardized smoking prevalence **(C)** among 204 countries or territories.

suggested the association between secondhand smoke and T2D risk (46, 47). Also, compared to men, the T2D burden attributable to secondhand smoke in women was higher and secondhand smoke was the main contributor to womens' tobacco-attributable T2D burden. Probably because daily time, frequency, and total duration of secondhand smoke exposure were higher in women, although secondhand smoke exposure prevalence was lower in women (46). Moreover, a prospective cohort in the US found a 16% higher rate of T2D among female nonsmokers with secondhand smoke exposure than those without (48).

Geographical differences were also detected at the regional and national level. For the high SDI region, the T2D burden attributable to tobacco remained lowest and showed the most significant downward trend since 1990. Higher SDI usually means more government expenditure on T2D, earlier screening, more access to health services, adequate health infrastructure, and people being more health-conscious (35, 36). North America and the Caribbean (415 billion USD) alone accounted for 42.96% of global diabetes health expenditure (966 billion USD) in 2021 (49). Meanwhile, the negative relationships between the SDI and ASRs confirmed the above view when the SDI was>0.40. Low-middle and middle SDI regions showed the highest burden of T2D attributable to tobacco and continued increasing trends. Countries from these regions are mostly developing countries (China, Brazil, India, and Indonesia), which are the main tobacco producing and consuming countries worldwide (50). Despite declines in smoking prevalence observed worldwide, rapid urbanization, population growth, and aging may offset the potential gains of these decreases and drive increases in tobacco-attributable T2D burden among low-middle and middle SDI countries (41). Also, the tobacco-attributable T2D burden did not seem to follow the expected relationship with the SDI in low SDI regions, which was lower than that in low-middle and middle SDI regions. The absence of universal healthcare coverage, shortage of medical personnel, and inability to timely diagnose may cause missed diagnosis, under-reporting, and underestimates regarding the burden T2D attributable to tobacco use in low SDI regions (34, 51).

Additionally, "high-high" clusters were mainly observed in South Africa, the Kingdom of Saudi Arabia, and Southeast Asian countries; and the tobacco-attributable T2D burden in these regions was much higher than the expected value-based SDI. This may be explained by the relatively high tobacco use rate and poor T2D treatment. The Southeast Asia region has 26% of the world's population and was one of the largest tobacco producers and consumers, with 250 million tobacco smokers and a massive number of smokeless tobacco users while at the same time, this region accounted for just 1% of the global diabetes health expenditure (49, 52). Meanwhile, the tobacco use rate among adolescents (13-15 years old) remains high in Southeast Asia, with the highest rate being 30.3% in Bhutan (53). The epidemiologic study has shown that Asians, especially South Asians, are more genetically susceptible to T2D (54). South Africa has been experiencing rapid socio-economic growth and the transformation of disease burden to NCD in recent years. However, medical and health undertakings seem to lag behind socio-economic development and fail to meet the needs of NCD prevention and control (55). In 2019, 59.7% of people with diabetes were undiagnosed in Africa (26). Also, late diagnosis and poor glycemic control exacerbated the burden of diabetes and complications in South Africa (56).

Across the 204 countries or territories, smoking prevalence is the most sensitive and direct factor affecting tobacco-attributable T2D burden, followed by diabetes treatment. Over the past decade, many effective tobacco-control initiatives and interventions have been conducted to address the tobacco epidemic, including tobacco taxes, smoking bans in public places, smoking cessation interventions, and the WHO Framework Convention on Tobacco Control (57, 58). The overall smoking prevalence has decreased, but the number of smokers is still increasing due to population growth. Multinational studies showed that the proportion of women smoking was gradually increasing (59). Meanwhile, tobacco control policies may be blocked in developing countries, where tobacco revenues are essential to the national economy (60). In clinical guidelines, smoking cessation is recognized as an essential intervention of the long-term care of patients with diabetes. However, the physician may encounter obstacles in promoting smoking cessation to prevent diabetes (14). A national cohort study in Australia found similar prevalence of smoking among diabetics (13.5%) as general Australian population (13.8%), partially suggesting poor adherence to primary and/or secondary prevention recommendations for smoking cessation among the diabetic population (9). In the NHANES study of 24,649 participants, the age-adjusted smoking rate was 25.7% among diabetics and 24.1% among non-diabetics (21). In Africa, the prevalence of smoking among diabetics was 12.9%, also similar to the prevalence of smoking in the general African population (12%) as reported by the WHO during the same period (61). This may be partly because patient concerns about weight gain and withdrawal effects after cessation, and some studies have shown that short-term weight gain after quitting smoking may increase the risk of T2D (62, 63). Additionally, smokers may have insufficient understanding of diabetes, and even if they are aware of the harm of the disease, they may continue to smoke. Smoking, especially continued smoking after diagnosis of diabetes, was independently associated with diabetes complications (22, 64). Therefore, physicians should advocate for smoking cessation interventions in the early stages and combine other interventions such as diet, exercise, and reducing the cost of smoking cessation treatment for patients to maximize the benefits of smoking cessation for diabetes improvement. Besides, most interventions focus on high-risk groups, but recent studies have shown that strategies focusing on detecting and treating high-risk groups are not enough (65). Diabetes "prevention" is often just a "delay" for high-risk groups. More measures should focus on population-based primary care strategies, targeting preventable risk factors that are easily modifiable, particularly tobacco use. Meanwhile, primary diagnosis and surveillance of T2D should be strengthened to improve data reliability among low SDI countries.

The GBD studies came up with comprehensive quality estimates of global disease burden and fill a gap where actual data on disease burden are sparse or unavailable, yet several limitations should be acknowledged. Data on tobacco use were obtained through self-reporting, which may lead to underestimation for population groups with low social acceptance of smoking, especially among women in Asia and Africa (66). Additionally, the accuracy and robustness of the GBD estimate largely depend on the quality and quantity of data used in the modeling. Vital registration, verbal autopsy, and statistics systems are critical sources of vital statistics for mortality rates. However, the population coverage with these systems was disappointing among low-income regions, which may lead to underestimating the T2D burden, although the GBD has conducted many adjusted methods to reduce such bias. Since our study is based on the population level, ecological fallacy might emerge and the relationship between mortality, DALYs and the SDI, ASSP, and DTI, although explanatory, cannot be considered as a causality. Finally, as a secondary analysis of the GBD data, we have no additional detailed covariable data to control the bias, such as race, education and occupation.

CONCLUSIONS

Tobacco should be regarded as an essential and preventable risk factor for the burden of T2D, especially in low-middle and middle SDI regions. Great efforts have been made on tobacco control and a declining trend of ASMR of T2D attributable to tobacco was observed, while the ASDR of T2D increased globally. Gender and regional disparities existed. Tobacco-attributable T2D burden posed great challenges to the elderly and men while secondhand smoke posed a greater risk to women. Hot spots were concentrated in South Africa, the Kingdom of Saudi Arabia, and Southeast Asia and needed more attention with supportive policies to lessen the T2D burden. Also, low SDI regions should increase their health investment in NCDs and strengthen the capacity of diabetes diagnosis and surveillance.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

Our study was based on a publicly available GBD database (GHDx). No patients, the public or animals were involved in the design, or conduct, or reporting, or dissemination plans of our study. All data were publicly open access online at (http://ghdx.healthdata.org/gbd-results-tool). Therefore, ethical approval is not applicable for our study.

AUTHOR CONTRIBUTIONS

Study design: CY and JC. Data collection: JB. Data analyses: JB, and FS. Results visualisation: JB. Results interpretations: All authors. Manuscript writing: JB. Manuscript revising: JB, FS, YM, DY, CY, and JC. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded by the National Natural Science Foundation of China [grant number 82173626, 81773552], Wuhan Total Health Cost Accounting Project 2017-2020 [grant number WHWSZFY2021], Health commission of Hubei Province scientific research project [grant number WJ2019H304].

ACKNOWLEDGMENTS

The data set supporting the conclusions of this article is available in the GBD (Global Burden of Disease) Data Tool repository (http://ghdx.healthdata.org/gbd-results-tool). This study used publicly available deidentified data accessed from the Global Burden of Disease Study 2019 repository. Therefore, ethical approval, and statement from an ethics committee or institutional review board were not required.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.905367/full#supplementary-material

REFERENCES

1. *IDF Diabetes Atlas*. Available at: https://diabetesatlas.org/ (Accessed 2022-04-29).
2. Tancredi M, Rosengren A, Svensson AM, Kosiborod M, Pivodic A, Gudbjornsdottir S, et al. Excess Mortality Among Persons With Type 2 Diabetes. *N Engl J Med* (2015) 373:1720–32. doi: 10.1056/NEJMoa1504347
3. Pan A, Wang Y, Talaei M, Hu FB, Wu T. Relation of Active, Passive, and Quitting Smoking With Incident Type 2 Diabetes: A Systematic Review and Meta-Analysis. *Lancet Diabetes Endocrinol* (2015) 3:958–67. doi: 10.1016/S2213-8587(15)00316-2
4. Safiri S, Karamzad N, Kaufman JS, Bell AW, Nejadghaderi SA, Sullman MJM, et al. Prevalence, Deaths and Disability-Adjusted-Life-Years (DALYs) Due to Type 2 Diabetes and Its Attributable Risk Factors in 204 Countries and Territories, 1990-2019: Results From the Global Burden of Disease Study 2019. *Front Endocrinol (Lausanne)* (2022) 13:838027. doi: 10.3389/fendo.2022.838027
5. Chi Y, Wang X, Jia J, Huang T. Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic

Etiology and Causal Relationship. *Front Endocrinol (Lausanne)* (2022) 13:809445. doi: 10.3389/fendo.2022.809445
6. Maddatu J, Anderson-Baucum E, Evans-Molina C. Smoking and the Risk of Type 2 Diabetes. *Transl Res* (2017) 184:101–7. doi: 10.1016/j.trsl.2017.02.004
7. P. National Center for Chronic Disease, S. Health Promotion Office on, and Health, Reports of the Surgeon General. *The Health Consequences of Smoking-50 Years of Progress: A Report of the Surgeon General*. US), Atlanta (GA: Centers for Disease Control and Prevention (2014).
8. Akter S, Goto A, Mizoue T. Smoking and the Risk of Type 2 Diabetes in Japan: A Systematic Review and Meta-Analysis. *J Epidemiol* (2017) 27:553–61. doi: 10.1016/j.je.2016.12.017
9. Szwarcbard N, Villani M, Earnest A, Flack J, Andrikopoulos S, Wischer N, et al. The Association of Smoking Status With Glycemic Control, Metabolic Profile and Diabetic Complications- Results of the Australian National Diabetes Audit (ANDA). *J Diabetes Complicat* (2020) 34:107626. doi: 10.1016/j.jdiacomp.2020.107626
10. Zhang M, Yang Y, Liu F, Jia J, Xu Q, Wang L, et al. The Association of Husbands' Smoking With Wives' Dysglycemia Status: A Cross-Sectional Study Among Over 10 Million Chinese Women Aged 20-49. *J Diabetes* (2020) 12:354–64. doi: 10.1111/1753-0407.13009
11. Molla GJ, Ismail-Beigi F, Larijani B, Khaloo P, Moosaie F, Alemi H, et al. Smoking and Diabetes Control in Adults With Type 1 and Type 2 Diabetes: A Nationwide Study From the 2018 National Program for Prevention and Control of Diabetes of Iran. *Can J Diabetes* (2020) 44:246–52. doi: 10.1016/j.jcjd.2019.07.002
12. Roth GA, Johnson CO, Abate KH, Abd-Allah F, Ahmed M, Alam K, et al. The Burden of Cardiovascular Diseases Among US States, 1990-2016. *JAMA Cardiol* (2018) 3:375–89. doi: 10.1001/jamacardio.2018.0385
13. Yang Y, Peng N, Chen G, Wan Q, Yan L, Wang G, et al. Interaction Between Smoking and Diabetes in Relation to Subsequent Risk of Cardiovascular Events. *Cardiovasc Diabetol* (2022) 21:14. doi: 10.1186/s12933-022-01447-2
14. Yang YS, Sohn TS. Smoking as a Target for Prevention of Diabetes. *Diabetes Metab J* (2020) 44:402–4. doi: 10.4093/dmj.2020.0126
15. N.C.D.C. collaborators. NCD Countdown 2030: Pathways to Achieving Sustainable Development Goal Target 3.4. *Lancet* (2020) 396:918–34. doi: 10.1016/S0140-6736(20)31761-X
16. *Global Report on Diabetes: Executive Summary*. Available at: https://www.who.int/publications/i/item/who-nmh-nvi-16.3 (Accessed 2022-06-02).
17. Śliwińska-Mossoń M, Milnerowicz H. The Impact of Smoking on the Development of Diabetes and its Complications. *Diabetes Vasc Dis Res* (2017) 14:265–76. doi: 10.1177/1479164117701876
18. *Standard Di Cura Del Diabete Mellito AMD-SID*. Available at: https://aemmedi.it/wp-content/uploads/2018/06/AMD-Standard-unico-protetto.pdf (Accessed 2020-06-02).
19. Sanni S, Hongoro C, Ndinda C, Wisdom JP. Assessment of the Multi-Sectoral Approach to Tobacco Control Policies in South Africa and Togo. *BMC Public Health* (2018) 18:962. doi: 10.1186/s12889-018-5829-3
20. Wen H, Xie C, Wang F, Wu Y, Yu C. Trends in Disease Burden Attributable to Tobacco in China, 1990-2017: Findings From the Global Burden of Disease Study 2017. *Front Public Health* (2020) 8:237. doi: 10.3389/fpubh.2020.00237
21. Clair C, Meigs JB, Rigotti NA. Smoking Behavior Among US Adults With Diabetes or Impaired Fasting Glucose. *Am J Med* (2013) 126:541.e15–8. doi: 10.1016/j.amjmed.2012.11.029
22. Yeom H, Lee JH, Kim HC, Suh I. The Association Between Smoking Tobacco After a Diagnosis of Diabetes and the Prevalence of Diabetic Nephropathy in the Korean Male Population. *J Prev Med Public Health* (2016) 49:108–17. doi: 10.3961/jpmph.15.062
23. G.B.D. Diseases and C. Injuries. Global Burden of 369 Diseases and Injuries in 204 Countries and Territories, 1990-2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *Lancet* (2020) 396:1204–22. doi: 10.1016/S0140-6736(20)30925-9
24. G.B.D.S. Collaborators. Global, Regional, and National Burden of Stroke, 1990-2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* (2019) 18:439–58. doi: 10.1016/S1474-4422(19)30034-1
25. G.B.D. Disease, I. Injury and C. Prevalence. Global, Regional, and National Incidence, Prevalence, and Years Lived With Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990-2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* (2018) 392:1789–858. doi: 10.1016/S0140-6736(18)32279-7
26. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and Regional Diabetes Prevalence Estimates for 2019 and Projections for 2030 and 2045: Results From the International Diabetes Federation Diabetes Atlas, 9(Th) Edition. *Diabetes Res Clin Pract* (2019) 157:107843. doi: 10.1016/j.diabres.2019.107843
27. G.U.H.C. Collaborators. Measuring Universal Health Coverage Based on an Index of Effective Coverage of Health Services in 204 Countries and Territories, 1990-2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *Lancet* (2020) 396:1250–84. doi: 10.1016/S0140-6736(20)30750-9
28. G.B.D.R.F. Collaborators. Global Burden of 87 Risk Factors in 204 Countries and Territories, 1990-2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *Lancet* (2020) 396:1223–49. doi: 10.1016/S0140-6736(20)30752-2
29. G.B.D. Disease, I. Injury and C. Prevalence. Global, Regional, and National Incidence, Prevalence, and Years Lived With Disability for 310 Diseases and Injuries, 1990-2015: A Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet* (2016) 388:1545–602. doi: 10.1016/S0140-6736(16)31678-6
30. G.B.D. Child, C. Adolescent Health, Reiner RCJr., Olsen HE, Ikeda CT, Echko MM, et al. Diseases, Injuries, and Risk Factors in Child and Adolescent Health, 1990 to 2017: Findings From the Global Burden of Diseases, Injuries, and Risk Factors 2017 Study. *JAMA Pediatr* (2019) 173:e190337.
31. Liu X, Yu Y, Wang M, Mubarik S, Wang F, Wang Y, et al. The Mortality of Lung Cancer Attributable to Smoking Among Adults in China and the United States During 1990-2017. *Cancer Commun (Lond)* (2020) 40:611–9. doi: 10.1002/cac2.12099
32. Cui Y, Mubarik S, Li R, Nawsherwan, Yu C. Trend Dynamics of Thyroid Cancer Incidence Among China and the U.S. Adult Population From 1990 to 2017: A Joinpoint and Age-Period-Cohort Analysis. *BMC Public Health* (2021) 21:624. doi: 10.1186/s12889-021-10635-w
33. Haagsma JA, James SL, Castle CD, Dingels ZV, Fox JT, Hamilton EB, et al. Burden of Injury Along the Development Spectrum: Associations Between the Socio-Demographic Index and Disability-Adjusted Life Year Estimates From the Global Burden of Disease Study 2017. *Inj Prev 26* (2020), i12–26. doi: 10.1136/injuryprev-2019-043296
34. Liu C, Wang B, Liu S, Li S, Zhang K, Luo B, et al. Type 2 Diabetes Attributable to PM2.5: A Global Burden Study From 1990 to 2019. *Environ Int* (2021) 156:106725. doi: 10.1016/j.envint.2021.106725
35. Mehdipour P, Navidi I, Parsaeian M, Mohammadi Y, Moradi Lakeh M, Rezaei Darzi E, et al. Application of Gaussian Process Regression (GPR) in Estimating Under-Five Mortality Levels and Trends in Iran 1990 - 2013, Study Protocol. *Arch Iran Med* (2014) 17:189–92.
36. Luo L, Jiang J, Zhang G, Wang L, Wang Z, Yang J, et al. Stroke Mortality Attributable to Ambient Particulate Matter Pollution From 1990 to 2015 in China: An Age-Period-Cohort and Spatial Autocorrelation Analysis. *Int J Environ Res Public Health* (2017) 14. doi: 10.3390/ijerph14070772
37. Fotheringham AS, Yang W, Kang W. Multiscale Geographically Weighted Regression (MGWR). *Ann Am Assoc Geograph* (2017) 107:1247–65. doi: 10.1080/24694452.2017.1352480
38. Tran DX, Pearson D, Palmer A, Lowry J, Gray D, Dominati EJ. Quantifying Spatial Non-Stationarity in the Relationship Between Landscape Structure and the Provision of Ecosystem Services: An Example in the New Zealand Hill Country. *Sci Tot Environ* (2022) 808:152126. doi: 10.1016/j.scitotenv.2021.152126
39. Mollalo A, Vahedi B, Rivera KM. GIS-Based Spatial Modeling of COVID-19 Incidence Rate in the Continental United States. *Sci Tot Environ* (2020) 728:138884. doi: 10.1016/j.scitotenv.2020.138884
40. Oshan TM, Li Z, Kang W, Wolf LJ, Fotheringham AS. Mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS Int J Geo-Informat* (2019) 8:269. doi: 10.3390/ijgi8060269
41. G.B.D.T. Collaborators. Smoking Prevalence and Attributable Disease Burden in 195 Countries and Territories, 1990-2015: A Systematic Analysis From the Global Burden of Disease Study 2015. *Lancet* (2017) 389:1885–906. doi: 10.1016/S0140-6736(17)30819-X

42. Huebschmann AG, Huxley RR, Kohrt WM, Zeitler P, Regensteiner JG, Reusch JEB. Sex Differences in the Burden of Type 2 Diabetes and Cardiovascular Risk Across the Life Course. *Diabetologia* (2019) 62:1761–72. doi: 10.1007/s00125-019-4939-5
43. Benowitz NL, Lessov-Schlaggar CN, Swan GE, Jacob P. 3rd, Female Sex and Oral Contraceptive Use Accelerate Nicotine Metabolism. *Clin Pharmacol Ther* (2006) 79:480–8. doi: 10.1016/j.clpt.2006.01.008
44. Feseke SK, St-Laurent J, Anassour-Sidi E, Ayotte P, Bouchard M, Levallois P. Arsenic Exposure and Type 2 Diabetes: Results From the 2007-2009 Canadian Health Measures Survey. *Health Promot Chronic Dis Prev Can* (2015) 35:63–72. doi: 10.24095/hpcdp.35.4.01
45. Schick SF, Glantz SA. Sidestream Cigarette Smoke Toxicity Increases With Aging and Exposure Duration. *Tob Control* (2006) 15:424–9. doi: 10.1136/tc.2006.016162
46. Kim BJ, Kim JH, Kang JG, Kim BS, Kang JH. Association Between Secondhand Smoke Exposure and Diabetes Mellitus in 131 724 Korean Never Smokers Using Self-Reported Questionnaires and Cotinine Levels: Gender Differences. *J Diabetes* (2021) 13:43–53. doi: 10.1111/1753-0407.13082
47. Sun K, Liu D, Wang C, Ren M, Yang C, Yan L. Passive Smoke Exposure and Risk of Diabetes: A Meta-Analysis of Prospective Studies. *Endocrine* (2014) 47:421–7. doi: 10.1007/s12020-014-0194-1
48. Zhang L, Curhan GC, Hu FB, Rimm EB, Forman JP. Association Between Passive and Active Smoking and Incident Type 2 Diabetes in Women. *Diabetes Care* (2011) 34:892–7. doi: 10.2337/dc10-2087
49. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, Regional and Country-Level Diabetes Prevalence Estimates for 2021 and Projections for 2045. *Diabetes Res Clin Pract* (2021), 109119.
50. Burki TK. WHO Releases Latest Report on the Global Tobacco Epidemic. *Lancet Oncol* (2021) 22:1217. doi: 10.1016/S1470-2045(21)00464-2
51. Wilson ML, Fleming KA, Kuti MA, Looi LM, Lago N, Ru K. Access to Pathology and Laboratory Medicine Services: A Crucial Gap. *Lancet* (2018) 391:1927–38. doi: 10.1016/S0140-6736(18)30458-6
52. Rani M, Thamarangsi T, Agarwal N. Youth Tobacco Use in South-East Asia: Implications for Tobacco Epidemic and Options for Its Control in the Region. *Indian J Public Health* (2017) 61:S12–7. doi: 10.4103/ijph.IJPH_241_17
53. WHO. *The Fatal Link Between Tobacco and Cardiovascular Diseases in the WHO South-East Asia Region*. Available at: https://www.who.int/southeastasia/activities/the-fatal-link-between-tobacco-and-cardiovascular-diseases (Accessed 2021-10-18).
54. Shariff AI, Kumar N, Yancy WS Jr., Corsino L. Type 2 Diabetes and Atherosclerotic Cardiovascular Disease in South Asians: A Unique Population With a Growing Challenge. *Curr Diabetes Rep* (2020) 20:4. doi: 10.1007/s11892-020-1291-6
55. Mudie K, Jin MM, Tan, Kendall L, Addo J, Dos-Santos-Silva I, et al. Non-Communicable Diseases in Sub-Saharan Africa: A Scoping Review of Large Cohort Studies. *J Glob Health* (2019) 9:020409. doi: 10.7189/jogh.09.020409
56. Dalal S, Beunza JJ, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, et al. Non-Communicable Diseases in Sub-Saharan Africa: What We Know Now. *Int J Epidemiol* (2011) 40:885–901. doi: 10.1093/ije/dyr050
57. Puska P, Daube M. Impact Assessment of the WHO Framework Convention on Tobacco Control: Introduction, General Findings and Discussion. *Tob Control 28* (2019), s81–3. doi: 10.1136/tobaccocontrol-2018-054429
58. Khokhar MA, Siddiqi K. Smoke-Free Policies in the Global South. *Nicot Tob Res* (2020) 22:2125–6. doi: 10.1093/ntr/ntaa210
59. Warren CW, Lea V, Lee J, Jones NR, Asma S, McKenna M. Change in Tobacco Use Among 13-15 Year Olds Between 1999 and 2008: Findings From the Global Youth Tobacco Survey. *Glob Health Promot* (2009) 16:38–90. doi: 10.1177/1757975909342192
60. Otanez M, Glantz SA. Social Responsibility in Tobacco Production? Tobacco Companies' Use of Green Supply Chains to Obscure the Real Costs of Tobacco Farming. *Tob Control* (2011) 20:403–11. doi: 10.1136/tc.2010.039537
61. Noubiap JJ, Nansseu JR, Endomba FT, Ngouo A, Nkeck JR, Nyaga UF, et al. Active Smoking Among People With Diabetes Mellitus or Hypertension in Africa: A Systematic Review and Meta-Analysis. *Sci Rep* (2019) 9:588. doi: 10.1038/s41598-018-37858-z
62. Hu Y, Zong G, Liu G, Wang M, Rosner B, Pan A, et al. Smoking Cessation, Weight Change, Type 2 Diabetes, and Mortality. *N Engl J Med* (2018) 379:623–32. doi: 10.1056/NEJMoa1803626
63. Tian J, Venn A, Otahal P, Gall S. The Association Between Quitting Smoking and Weight Gain: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. *Obes Rev* (2015) 16:883–901. doi: 10.1111/obr.12304
64. Kar D, Gillies C, Nath M, Khunti K, Davies MJ, Seidu S. Association of Smoking and Cardiometabolic Parameters With Albuminuria in People With Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis. *Acta Diabetol* (2019) 56:839–50. doi: 10.1007/s00592-019-01293-x
65. Barry E, Roberts S, Oke J, Vijayaraghavan S, Normansell R, Greenhalgh T. Efficacy and Effectiveness of Screen and Treat Policies in Prevention of Type 2 Diabetes: Systematic Review and Meta-Analysis of Screening Tests and Interventions. *BMJ 356* (2017), i6538. doi: 10.1136/bmj.i6538
66. G.T. Collaborators. Spatial, Temporal, and Demographic Patterns in Prevalence of Smoking Tobacco Use and Attributable Disease Burden in 204 Countries and Territories, 1990-2019: A Systematic Analysis From the Global Burden of Disease Study 2019. *Lancet* (2021) 397:2337–60. doi: 10.1016/S0140-6736(21)01169-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fendo.2022.964132

OPEN ACCESS

EDITED BY
Changwei Li,
Tulane University School of Public Health and Tropical Medicine, United States

REVIEWED BY
Aqeel M. Alenazi,
Prince Sattam Bin Abdulaziz University, Saudi Arabia
Hao Peng,
Soochow University, China

*CORRESPONDENCE
Chen Chai
chaichen@whu.edu.cn

SPECIALTY SECTION
This article was submitted to Systems Endocrinology, a section of the journal Frontiers in Endocrinology

RECEIVED 06 July 2022
ACCEPTED 12 October 2022
PUBLISHED 03 November 2022

CITATION
Deng M-G, Cui H-T, Lan Y-B, Nie J-Q, Liang Y-H and Chai C (2022) Physical activity, sedentary behavior, and the risk of type 2 diabetes: A two-sample Mendelian Randomization analysis in the European population. *Front. Endocrinol.* 13:964132. doi: 10.3389/fendo.2022.964132

Physical activity, sedentary behavior, and the risk of type 2 diabetes: A two-sample Mendelian Randomization analysis in the European population

Ming-Gang Deng[1], Han-Tao Cui[1], Yong-Bing Lan[2], Jia-Qi Nie[1], Yue-Hui Liang[1] and Chen Chai[3]*

[1]School of Public Health, Wuhan University, Wuhan, China, [2]Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, China, [3]Emergency Center, Hubei Clinical Research Center for Emergency and Resuscitation, Zhongnan Hospital of Wuhan University, Wuhan, China

Physical activity (PA) and sedentary behaviors (SB) have been linked to the risk of type 2 diabetes (T2DM) in observational studies; however, it is unclear whether these associations are causative or confounded. This study intends to use summary genetic data from the UK Biobank and other consortiums in conjunction with the two-sample Mendelian Randomization (MR) approach to solve this problem. The inverse variance weighted (IVW) technique was utilized as the primary analysis, with sensitivity analyses using the MR-Egger, weighted-median, and MR-Pleiotropy RESidual Sum and Outlier (PRESSO) techniques. Inverse associations between self-reported moderate PA (OR: 0.3096, 95% CI: 0.1782-0.5380) and vigorous PA (OR: 0.2747, 95% CI: 0.1390-0.5428) with T2DM risk were found, respectively. However, accelerometer-based PA measurement (average acceleration) was not associated with T2DM risk (OR: 1.0284, 95% CI: 0.9831-1.0758). The time (hours/day) spent watching TV was associated with T2DM risk (OR: 2.3490, 95% CI: 1.9084-2.8915), while the time (hours/day) spent using the computer (OR: 0.8496, 95% CI: 0.7178-1.0056), and driving (OR: 3.0679, 95% CI: 0.8448-11.1415) were not associated with T2DM risk. The sensitivity analysis revealed relationships of a similar magnitude. Our study revealed that more PA and less TV viewing were related to a decreased T2DM risk, and provided genetic support for a causal relationship between PA, TV viewing, and T2DM risk.

KEYWORDS

physical activity, sedentary behaviors, type 2 diabetes, Mendelian Randomization, causal effect

Introduction

Type 2 diabetes mellitus (T2DM) is a common disease with increasing incidence globally. The incidence of T2DM in the European region has been continuously rising, from 190/100,000 in 1990 to 328/100,000 in 2019, according to the Global Burden of Disease-2019 (1). Prevention strategies targeting T2DM risk factors are essential to control the growing burden. Among various risk factors for T2DM, including ethnicity, obesity, an unhealthy diet, physical inactivity (PA), sedentary behaviors (SB), and so on (2), the effects of PA and SB have attracted special attention.

PA can provide an antidiabetic impact by improving blood glucose levels in people with T2DM (3), because it entails bodily movements and energy expenditure (4). Numerous prior observational studies have shown both moderate physical activity (MPA) (5, 6) and vigorous physical activity (VPA) (5, 6) were related to decreased risks of T2DM. Additionally, the incidence of T2DM decreased as the quantity of PA increased (7, 8). However, the link between PA and T2DM may not always be plausible, and it may be influenced by gender and the kind of PA. According to a study in Korean adults, working PA was not associated with T2DM risk, whereas transport PA was solely linked to T2DM in males (9).

SB is frequently characterized as sitting, watching television, or couch time, and is a potentially significant factor in health (10). SB was claimed to affect T2DM risk in many studies (11–15). However, no connection between the incidence of T2DM and objectively measured sedentary time was found by an earlier study (16). A population-based study found that watching TV was the main SB, accounting for roughly half of the total sedentary time, followed by computer viewing (17). The majority of studies used overall sedentary time or time spent watching television as the main indicator of SB (13, 15). Nevertheless, mentally active SB, like using a computer or driving (18), may have different impacts on T2DM risk than mentally passive SB, like watching TV, and to our knowledge, this issue has not yet received much attention.

Mendelian randomization (MR), is a well-established tool for causal inference by employing genetic variants as instrumental variables (IV) for exposures, e.g., PA and SB (19). Since the genetic variants are randomly assigned during meiosis, they are not susceptible to reverse causation bias and confounders, which are the general flaws of conventional epidemiological methods (20). Therefore, MR can yield stronger evidence by assessing whether the observed connection between risk variables and outcome is compatible with causal impact.

The objective of this study is to employ the MR technique to pinpoint the causal relationships among PA, SB, and T2DM in the European population. We hypothesis that more physical activity and less sedentary time were related to a lower risk of T2DM.

Methods

Study design

An overview of the study design is presented in Figure 1. Our study consisted of five components (1): identification of genetic variants to serve as instrumental variables for PA and SB (2). obtaining the instrumental SNPs for outcome summary data from genome-wide association studies (GWAS) of T2DM (3). harmonizing the exposure and outcome databases (4). performing MR analysis (5) evaluation of MR analysis assumptions and sensitivity analysis.

Three critical assumptions must be met for causal estimations derived from MR analysis to be valid (1): they associate with the risk factor of interest (the relevance assumption) (2); they share no common cause with the outcome (the independence assumption) (3); they do not affect

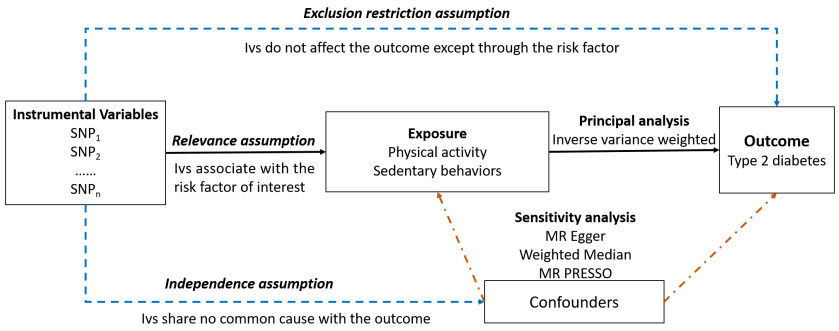


FIGURE 1
The overview of the study design.

the outcome except through the risk factor (the exclusion restriction assumption) (21).

GWAS summary data for outcome

Summary data on the associations of genetic variants with physician diagnosed T2DM were obtained from a recent GWAS meta-analysis of 62,892 T2DM patients and 596,424 controls of European ancestry with a total of 16 million gene variations (22). The study was composed of three contributing studies, including the full cohort release of the UK Biobank (UKB), Genetic Epidemiology Research on Aging (GERA), and Diabetes Genetics Replication and Meta-analysis (DIAGRAM).

GWAS summary data for exposures

Physical activities

Three types of PA were included in our study, e.g., self-reported moderate physical activity (MPA), self-reported vigorous physical activity (VPA), and accelerometer-based physical activity (Accelerometer-based PA). Self-reported levels of PA were measured *via* a touchscreen questionnaire, in a fashion similar to the International Physical Activity Questionnaire (23). For MPA, participants were directly asked "How many minutes did you usually spend doing moderate activities on a typical day, like carrying light loads, cycling at normal pace (Do not include walking)?". A total of 343,827 participants were included.

For VPA, participants were asked "In a typical WEEK, how many days did you do 10 minutes or more of vigorous physical activity? (These are activities that make you sweat or breathe hard such as fast cycling, aerobics, heavy lifting)", and those who indicated 1 or more such days were then asked "How many minutes did you usually spend doing vigorous activities on a typical DAY". These participants were then dichotomized into two groups (1): those who reported 0 days of VPA, and (2) those who reported both 3 or more days of VPA and a typical duration of VPA that is 25 minutes or greater. Individuals that did not fall into either of these two groups were excluded, and a total of 261,055 individuals were included, with 98,060 cases and 162,995 controls (24).

Accelerometer-based PA was assessed by an Axivity AX3 wrist-worn accelerometer, as previously described (25). Participants were informed in the invitation email and device mail-out letter that the accelerometer should be worn continuously and that they should carry on with their normal activities. PA information (overall acceleration average) was extracted from 100Hz raw triaxial acceleration data after calibration, removal of gravity and sensor noise, and identification of wear/non-wear episodes. Individuals with less than 3 days (72 hours) of data, or those not having data in each one hour of the 24-hour cycle, and outliers with values more than 4 standard deviations above the mean were excluded. 72 hours of wear was determined to be needed to be within 10% of a complete severe-day measure, which was based on missing data simulations by Doherty et al (25). A total of 91,084 participants were included, with the mean and standard deviation of average acceleration being 27.98 and 8.14, respectively.

Sedentary behaviors

Watching television, using the computer, and driving were identified as three types of SB in our study. For the ascertainment of sedentary time, during the first visit, participants were asked three questions, "On a typical day, how many hours do you spend watching TV?", "In a typical day, how many hours do you spend using the computer? (Do not include using a computer at work)" and "On a typical day, how many hours do you spend driving?". A total number of 437,887, 360,895, and 310,555 individuals were included for watching TV, using the computer, and driving, respectively. The hours/day of these sedentary behaviors were treated as exposure measurements.

Selection of instrumental variables

The genome-wide significance level was defined at $p < 5*10^{-8}$ to fulfill the relevance assumption, making instrumental variables robustly associated with the outcome. The *F*-statistic was calculated to estimate the strength of each SNP, utilizing the formula given previously (26). To fulfill the independence assumption, we employed the PLINK clumping method (27) to clump the SNPs and make them independent of each other ($r^2 < 0.001$, region size = 10000kb). Pleiotropy tests were also carried out to ensure that the exclusion restriction assumption was met (21).

SNPs that were unavailable in the outcome datasets were replaced by suitable proxy SNPs with minimum linkage disequilibrium $R^2 = 0.8$ and minor allele frequency threshold = 0.3, where available. Finally, 5, 7, 6, 89, 73, and 6 SNPs associated with MPA, VPA, Accelerometer-based PA, watching television, using the computer, and driving were identified, respectively. Summary statistics are presented in Supplementary Table S1.

Statistical analysis

The SNP-exposure and SNP-outcome coefficients were combined in a random-effects meta-analysis using the inverse-variance weighted (IVW) method to primarily provide an overall estimate of a causal impact.

Several sensitivity analyses were carried out to check and correct the presence of pleiotropy in the causal estimates. Cochran's *Q* was calculated to check the heterogeneity of the individual causal effect, with a *P*-value < 0.05, indicating the presence of pleiotropy. Consequently, weights were penalized to improve the robustness of the IVW method (28). MR-Egger intercept term was used to assess the horizontal pleiotropy, where deviation from zero indicates the directional pleiotropy. Moreover, the slope of the MR-Egger regression gives valid MR estimates when horizontal pleiotropy exists (29, 30). The complementary weighted-median method was used which can give valid MR estimates by assuming that at least 50% of IVs are effective and ordering the MR estimates of each IV weighted for the inverse of their variance (31). MR pleiotropy residual sum and outlier (MR-PRESSO) global test method was conducted to detect horizontal pleiotropy and the MR-PRESSO outlier test was conducted to correct the horizontal pleiotropy *via* outlier removal (32). The leave-one-out analysis was performed to assess the influence of a single SNP on the MR estimates.

The TwoSampleMR (version 0.5.6), MendelianRandomization (version 0.5.1), and MRPRESSO (version 1.0) packages were used for statistical analyses in R software (version 4.1.0, R Foundation for Statistical Computing). All statistical tests were two-tailed, and significance was considered at a Bonferroni corrected *p-value* below 0.0083 (correcting for 6 exposures and 1 outcome).

Results

MR estimates of causal effects of physical activity on type 2 diabetes

MR estimates between PA and T2DM risk are illustrated in Table 1. Increment in genetically predicted duration of MPA was associated with lower risk of T2DM (odds ratio [OR]: 0.3096, 95% confidence interval [95% CI]: 0.1782-0.5380, *P*-value < 0.0001). An inverse relationship was also found between VPA and T2DM risk, compared to those who reported 0 days of VPA per week, people who reported both 3 or more days of VPA per week had about 63% lower risk of T2DM (OR: 0.2747, 95% CI: 0.1390-0.5428, *P*-value = 0.0002). However, no significant association between Accelerometer-based PA and T2DM risk was found (OR: 1.0284, 95% CI: 0.9831-1.0758, *P*-value = 0.2290). There was some evidence of heterogeneity based on Cochran's *Q* (*Q*-value = 22.4315, *P*-value = 0.0004) for the accelerometer-based PA analysis. Consequently, weights were penalized for the IVW method.

The scatter plots of PA and T2DM risk are depicted in Figure 2. MR estimates for each SNP associated with PA in relation to T2DM risk are presented in Supplementary Figure S1, and the funnel plots of PA and T2DM risk association are presented in Supplementary Figure S2.

MR estimates of causal effects of sedentary behaviors on type 2 diabetes

MR estimates between SB and T2DM risk are illustrated in Table 2. A positive association between the duration of watching TV and T2DM risk was found (OR: 2.3490, 95% CI: 1.9084-2.8915, *P*-value < 0.0001). However, we found no association between the duration of using the computer or driving and T2DM risk. There was some evidence of heterogeneity based on Cochran's Q for the three types of SB; consequently, for these IVW models, weights were penalized to improve the robustness.

Scatter plots of SB and T2DM risk are presented in Figure 3. MR estimates for each SNP associated with SB in relation to T2DM risk are presented in Supplementary Figure S3, and the funnel plots of SB and T2DM risk association are presented in Supplementary Figure S4.

Evaluation of assumptions and sensitivity analyses

The strength of the genetic instruments denoted by the *F*-statistic was ≥ 10 for all the exposures, see Supplementary Table S1. No evidence of directional pleiotropy was found for all exposures (all MR-Egger intercept *P*-values > 0.05), see Tables 1 and 2. The estimates from the weighted-median method for the exposures were consistent with those of IVW methods, with an exception for driving (Tables 1, 2). The MR-PRESSO approach identified outlier SNPs for watching TV, using the computer, and driving, but similar magnitude associations were found after these outliers were excluded from the analysis (Table 1, 2). Moreover, the estimates from the leave-one-out analysis did not reveal any influential SNPs driving the overall association (Supplementary Figures S5, S6).

Discussion

The present study performed a Mendelian Randomization in the European population to investigate the genetic association of physical activity and sedentary behaviors with the risk of type 2 diabetes. Our study validated that self-reported physical activity, either moderate or vigorous intensity, is genetically connected with lower risks of developing T2DM, whereas acceleration-based physical activity is not. Watching TV is correlated with an elevated risk of developing T2DM, while using a computer, and driving are not.

Our findings that MPA and VPA are associated with a lower risk of T2DM are in line with those of previous studies. According to a previous study, those who regularly engage in MPA had a 30% lower risk of T2DM compared to sedentary

TABLE 1 MR estimates between physical activity and T2DM risk.

Methods	OR	95% CI	*P*-value	*Q*-value	*P*-value for heterogeneity ‡ or pleiotropy§
MPA					
IVW	0.3096	0.1782-0.5380	<0.0001	1.9273	0.7491
MR-Egger	1.4053	0.0720-27.4441	0.8369		0.3847
Weighted-Median	0.3167	0.1529-0.6560	0.0019		
MR-PRESSO	0.3096	0.2110-0.4544	0.0039		
VPA					
IVW	0.2747	0.1390-0.5428	0.0002	3.1608	0.7884
MR-Egger	0.3445	0.0016-74.8385	0.7138		0.9370
Weighted-Median	0.2569	0.1070-0.6168	0.0024		
MR-PRESSO	0.2747	0.1676-0.4503	0.0021		
Accelerometer-based PA					
IVW †	1.0284	0.9831-1.0758	0.2290	22.4315	0.0004
MR-Egger	1.1969	0.8054-1.7786	0.4241		0.4222
Weighted-Median	1.0228	0.9803-1.0671	0.2981		
MR-PRESSO	1.0015	0.9414-1.0654	0.9632		

MPA: moderate physical activity; VPA: vigorous physical activity; OR: odds ratio; CI: confidence intervals; Q-value: Cochran's Q statistic; IVW: inverse variance weighted
† Weights were penalized due to the presence of heterogeneity based on Cochran's Q statistic
‡ P-value for heterogeneity based on Cochran's Q statistic
§ P-value or pleiotropy based on MR-Egger interceptTable 2 MR estimates between sedentary behavior and T2DM risk

individuals (33).. Weekly VPA may decrease the risk of T2DM among those who are normal weight or overweight (34).. However, the link between PA and T2DM may not always be plausible, and it may be influenced by gender and the kind of PA. According to a study in Korean adults, wsorking PA was not associated with T2DM risk, whereas transport PA was solely linked to T2DM in males (9). One previous MR study reported that objectively measured average or vigorous physical activity and SB are not associated with the risk of T2DM (16). This outcome is in line with our findings that there is no correlation between acceleration-based PA and T2DM risk. Acceleration-based PA and self-reported PA were demonstrated to have a poor agreement in several studies (35, 36). In general adults, the self-reported time spent on MPA and VPA exceeded the time measured with the accelerometer (36). The disparity between self-reported and Accelerometer-measured MVPA increased with higher activity and intensity levels (35). The PA acquired from the accelerometer and the self-report is not conceptually

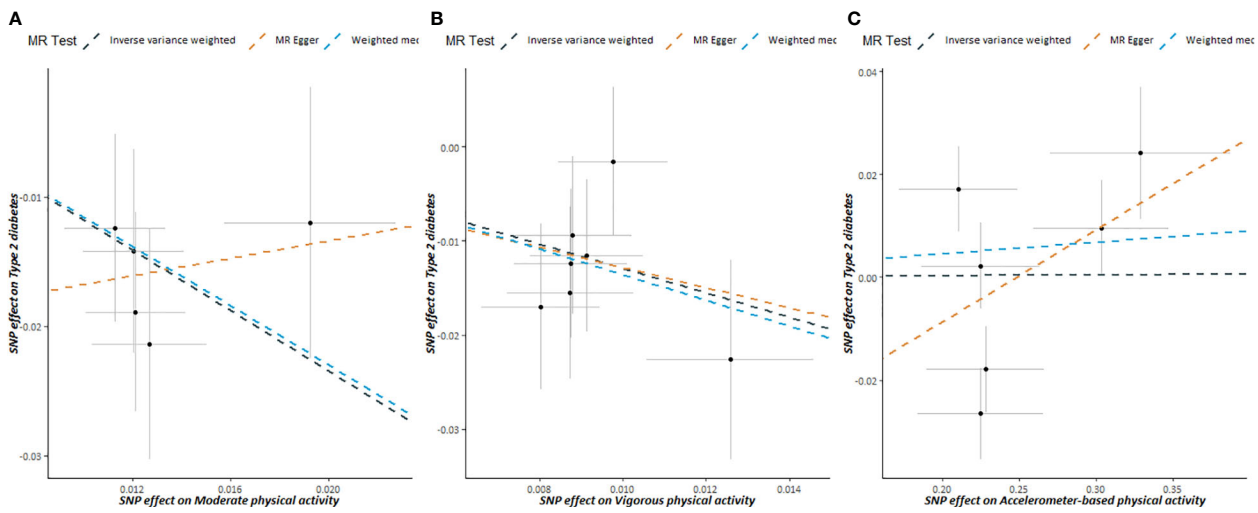


FIGURE 2
Scatter plots showing the correlation of genetic associations of physical activity with genetic associations with type 2 diabetes, **(A)** moderate physical activity; **(B)** vigorous physical activity; **(C)** accelerometer-based physical activity.

10.3389/fendo.2022.964132

TABLE 2 MR estimates between sedentary behavior and T2DM risk.

Methods	OR	95% CI	*P*-value	*Q*-value	*P*-value for heterogeneity ‡ or pleiotropy §
Watching TV					
IVW †	2.3490	1.9084-2.8915	<0.0001	283.8163	<0.0001
MR-Egger	4.5997	1.0169-20.8051	0.0470		0.4226
Weighted-Median	2.5143	1.9260-3.2824	<0.0001		
MR-PRESSO ¶	2.6092	2.0432-3.3320	<0.0001		
Using the computer					
IVW †	0.8496	0.7178-1.0056	0.0580	111.9547	0.0018
MR-Egger	1.6372	0.5794-4.6265	0.3520		0.3083
Weighted-Median	0.8834	0.6982-1.1176	0.3030		
MR-PRESSO ¶	0.9120	0.7555-1.1010	0.3411		
Driving					
IVW †	3.0679	0.8448-11.1415	0.0880	35.14185	<0.0001
MR-Egger	0.0533	0.0000-73665.88	0.7052		0.5988
Weighted-Median	3.4418	1.5140-7.8243	0.0030		
MR-PRESSO ¶	2.9869	0.6258-14.2560	0.2636		

OR, odds ratio; CI, confidence intervals; Q-value, Cochran's Q statistic; MR, Mendelian randomization; IVW, inverse variance weighted
† Weights were penalized due to the presence of heterogeneity based on Cochran's Q statistic
‡ P-value for heterogeneity based on Cochran's Q statistic
§ P-value or pleiotropy based on MR-Egger intercept
¶ Results after removal of outliers

equivalent, which explains why the Acceleration-based PA findings in our research vary from the MPA and VPA (37).

Numerous observational studies reported that watching TV is associated with a higher incidence of T2DM (38–40), but the finding was subject to unobserved confounders, or reverse causation. For the first time, the present study confirms this association *via* Mendelian Randomization, which avoids residual confounding by taking advantage of the instrument variable of genetic variants. A total of 89 genetic SNPs were identified to be associated with watching TV, which enables a robust estimation of the association between watching TV and the risk of T2DM. On average, watching TV increases T2DM more than one-fold.

The present study did not find computer usage was associated with the risk of T2DM. The finding concurs with

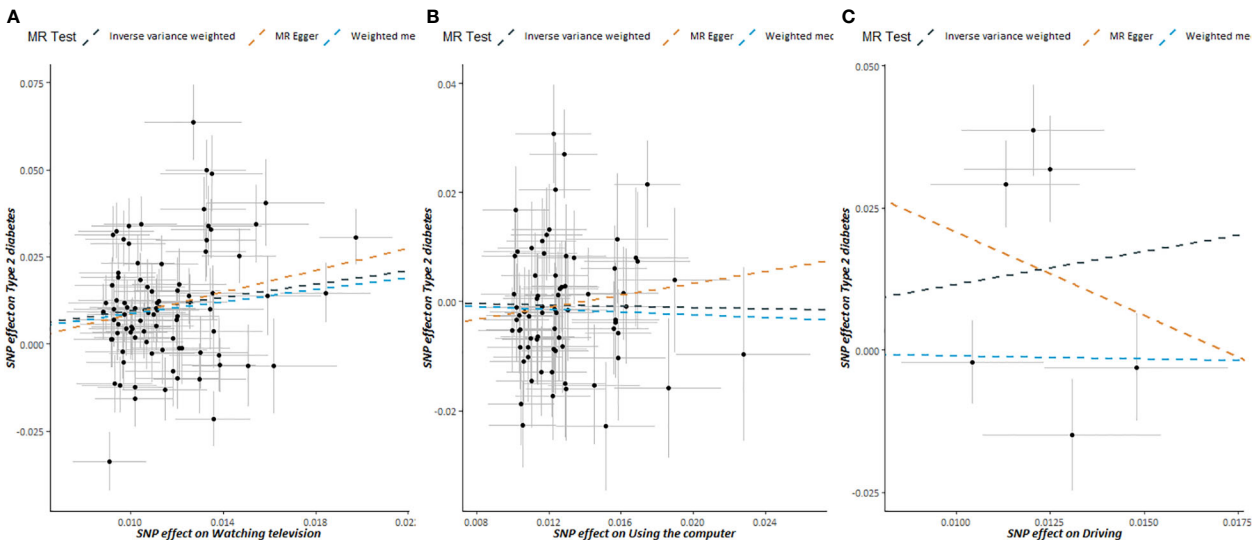


FIGURE 3
Scatter plots showing the correlation of genetic associations of sedentary behavior with genetic associations with type 2 diabetes, **(A)** watching television; **(B)** using the computer; **(C)** driving.

previous studies. One study showed that computer use was unrelated to HbA1c and blood lipids (41). Computer use also had no significant association with the risk of T2DM in Taiwanese older adults (42). It differs from watching television in that individuals may be more physically and cognitively engaged while using a computer than when watching television (43). Another possible explanation is that the computer time used in this article does not include computer time at work, which may be biased.

The present study did not find driving was associated with the risk of T2DM. Most recent research on driving and T2DM, to our knowledge, have focused on occupational drivers (44–46), who had a higher prevalence of T2DM. However, the average daily driving duration of participants in our study was roughly 1.18 hours, indicating that the driving employment may unlikely impact our findings. Furthermore, it is commonly understood that driving necessitates drivers to be cognitively engaged, which is not the case while watching television (18).

Several mechanisms could explain the links. First, the disparate impacts of TV viewing and physical activity may be mediated by diet and BMI. PA can reduce the risk of T2DM by reducing obesity, which is a greater risk of T2DM (47, 48). Individuals who engage in more vigorous physical activity may adhere to a healthier diet more closely, consume fewer snacks, and spend less time watching television (49). Second, according to a prospective study, PA played an important role in glycemic control (50), and the skeletal muscle, which is a primary tissue that determines blood glucose, by increasing insulin sensitivity (51, 52). Nevertheless, SB is not conducive to glycemic control (53). Additionally, PA's antioxidant and anti-inflammatory properties may help suppress the progress of T2DM (54).

There are certain strengths of our study. We apply the MR method to avoid confounding biases and reverse causation in observational studies. The genetic link between PA, SB, and T2DM risk was validated by our study. To our knowledge, we are the first to explore the genetic relationship between SB and T2DM from the angles of watching TV, using a computer, and driving.

Several limitations must be addressed. To begin, our research included only European individuals, which means that the findings cannot be properly extrapolated to other ethnic groups. Moreover, self-reported PA and SB may introduce measurement bias and result in estimations that differ from objective measures. Furthermore, since our research was limited to summary data, populations were not categorized according to sociodemographic factors (e.g., age, sex, or employment) when examining casual connections. Additionally, but certainly not least, light physical activity and other sedentary activities were not considered in the study. Finally, this study was a univariate MR study, without multi-adjustments for BMI, obesity, and other covariates, which were confounders for T2DM risks.

Conclusion

Our study revealed that more PA and less TV viewing were related to a decreased T2DM risk, and provided genetic support for a causal relationship between PA, TV viewing, and T2DM risk.

Data availability statement

The data that support the finding of this work was available in the UK Biobank at https://www.ukbiobank.ac.uk/ and the recently published article at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6195860/. These data were derived from the IEU OpenGWAS Project at https://gwas.mrcieu.ac.uk/, last accessed on 21 February 2022.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.964132/full#supplementary-material

References

1. Collaborators GDaI. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: A systematic analysis for the global burden of disease study 2019. *Lancet (London England)* (2020) 396(10258):1204–22. doi: 10.1016/S0140-6736(20)30925-9

2. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, 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 Practice* (2019) 157:107843. doi: 10.1016/j.diabres.2019.107843

3. Wake AD. Antidiabetic effects of physical activity: How it helps to control type 2 diabetes. *Diabetes Metab syndrome obesity: Targets Ther* (2020) 13:2909–23. doi: 10.2147/DMSO.S262289

4. Ainsworth B, Cahalin L, Buman M, Ross R. The current state of physical activity assessment tools. *Prog Cardiovasc Diseases* (2015) 57(4):387–95. doi: 10.1016/j.pcad.2014.10.005

5. Honda T, Kuwahara K, Nakagawa T, Yamamoto S, Hayashi T, Mizoue T. Leisure-time, occupational, and commuting physical activity and risk of type 2 diabetes in Japanese workers: A cohort study. *BMC Public Health* (2015) 15:1004. doi: 10.1186/s12889-015-2362-5

6. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ. Physical activity and the risk of type 2 diabetes: A systematic review and dose-response meta-analysis. *Eur J Epidemiol* (2015) 30(7):529–42. doi: 10.1007/s10654-015-0056-z

7. Wahid A, Manek N, Nichols M, Kelly P, Foster C, Webster P, et al. Quantifying the association between physical activity and cardiovascular disease and diabetes: A systematic review and meta-analysis. *J Am Heart Assoc* (2016) 5(9):e002495. doi: 10.1161/JAHA.115.002495

8. Hu FB, Sigal RJ, Rich-Edwards JW, Colditz GA, Solomon CG, Willett WC, et al. Walking compared with vigorous physical activity and risk of type 2 diabetes in women: A prospective study. *Jama* (1999) 282(15):1433–9. doi: 10.1001/jama.282.15.1433

9. Lee EB, Hong S, Min J, Park DH, Cho W, Suh SH, et al. Association between domain-specific physical activity and diabetes in Korean adults. *Sci Rep* (2021) 11(1):13066. doi: 10.1038/s41598-021-92560-x

10. Thyfault JP, Du M, Kraus WE, Levine JA, Booth FW. Physiology of sedentary behavior and its relationship to health outcomes. *Med Sci Sports Exercise* (2015) 47(6):1301–5. doi: 10.1249/MSS.0000000000000518

11. Bertoglia MP, Gormaz JG, Libuy M, Sanhueza D, Gajardo A, Srur A, et al. The population impact of obesity, sedentary lifestyle, and tobacco and alcohol consumption on the prevalence of type 2 diabetes: Analysis of a health population survey in Chile, 2010. *PloS One* (2017) 12(5):e0178092. doi: 10.1371/journal.pone.0178092

12. Hu FB, Li TY, Colditz GA, Willett WC, Manson JE. Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women. *Jama* (2003) 289(14):1785–91. doi: 10.1001/jama.289.14.1785

13. Patterson R, McNamara E, Tainio M, de Sá TH, Smith AD, Sharp SJ, et al. Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: A systematic review and dose response meta-analysis. *Eur J Epidemiol* (2018) 33(9):811–29. doi: 10.1007/s10654-018-0380-1

14. Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, Tennant B, et al. Sedentary behavior and health: Update from the 2018 physical activity guidelines advisory committee. *Med Sci Sports Exerc* (2019) 51(6):1227–41. doi: 10.1249/MSS.0000000000001935

15. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: A systematic review and meta-analysis. *Ann Internal Med* (2015) 162(2):123–32. doi: 10.7326/M14-1651

16. Meisinger C, Linseisen J, Leitzmann M, Baurecht H, Baumeister SE. Association of physical activity and sedentary behavior with type 2 diabetes and glycemic traits: A two-sample mendelian randomization study. *BMJ Open Diabetes Res Care* (2020) 8(2):e001896. doi: 10.1136/bmjdrc-2020-001896

17. Andrade-Gomez E, Garcia-Esquinas E, Ortola R, Martinez-Gomez D, Rodriguez-Artalejo F. Watching TV has a distinct sociodemographic and lifestyle profile compared with other sedentary behaviors: A nationwide population-based study. *PloS One* (2017) 12(12):e0188836. doi: 10.1371/journal.pone.0188836

18. Hallgren M, Nguyen TT, Owen N, Stubbs B, Vancampfort D, Lundin A, et al. Cross-sectional and prospective relationships of passive and mentally active sedentary behaviours and physical activity with depression. *Br J Psychiatry J Ment science* (2020) 217(2):413–9. doi: 10.1192/bjp.2019.60

19. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* (2014) 23(R1):R89–98. doi: 10.1093/hmg/ddu328

20. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *Jama* (2017) 318(19):1925–6. doi: 10.1001/jama.2017.17219

21. Davies NM, Holmes MV, Davey Smith G. Reading mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ (Clinical Res ed)* (2018) 362:k601. doi: 10.1136/bmj.k601

22. Xue A, Wu Y, Zhu Z, Zhang F, Kemper KE, Zheng Z, et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat Commun* (2018) 9(1):2941. doi: 10.1038/s41467-018-04951-w

23. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* (2003) 35(8):1381–95. doi: 10.1249/01.MSS.0000078924.61453.FB

24. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, Mandarino LJ, et al. Genome-wide association study of habitual physical activity in over 377,000 UK biobank participants identifies multiple variants including CADM2 and APOE. *Int J Obes (2005)* (2018) 42(6):1161–76. doi: 10.1038/s41366-018-0120-3

25. Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, et al. Large Scale population assessment of physical activity using wrist worn accelerometers: The UK biobank study. *PloS One* (2017) 12(2):e0169649. doi: 10.1371/journal.pone.0169649

26. Papadimitriou N, Dimou N, Tsilidis KK, Banbury B, Martin RM, Lewis SJ, et al. Physical activity and risks of breast and colorectal cancer: A mendelian randomisation analysis. *Nat Commun* (2020) 11(1):597. doi: 10.1038/s41467-020-14389-8

27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* (2007) 81(3):559–75. doi: 10.1086/519795

28. Burgess S, Bowden J, Dudbridge F, Thompson SG. Robust instrumental variable methods using multiple candidate instruments with application to mendelian randomization. *arXiv: Method* (2016). doi: 10.48550/arXiv.1606.03729

29. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *Int J Epidemiol* (2015) 44(2):512–25. doi: 10.1093/ije/dyv080

30. Burgess S, Thompson SG. Interpreting findings from mendelian randomization using the MR-egger method. *Eur J Epidemiol* (2017) 32(5):377–89. doi: 10.1007/s10654-017-0255-x

31. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* (2016) 40(4):304–14. doi: 10.1002/gepi.21965

32. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nat Genet* (2018) 50(5):693–8. doi: 10.1038/s41588-018-0099-7

33. Jeon CY, Lokken RP, Hu FB, van Dam RM. Physical activity of moderate intensity and risk of type 2 diabetes: A systematic review. *Diabetes Care* (2007) 30(3):744–52. doi: 10.2337/dc06-1842

34. Siegel LC, Sesso HD, Bowman TS, Lee IM, Manson JE, Gaziano JM. Physical activity, body mass index, and diabetes risk in men: A prospective study. *Am J Med* (2009) 122(12):1115–21. doi: 10.1016/j.amjmed.2009.02.008

35. Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of self-reported versus accelerometer-measured physical activity. *Med Sci sports exercise* (2014) 46(1):99–106. doi: 10.1249/MSS.0b013e3182a0595f

36. Domingos C, Correia Santos N, Pêgo JM. Association between self-reported and accelerometer-based estimates of physical activity in Portuguese older adults. *Sensors (Basel Switzerland)* (2021) 21(7):2258. doi: 10.3390/s21072258

37. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med* (2014) 48(13):1019–23. doi: 10.1136/bjsports-2014-093546

38. Grontved A, Hu FB. Television viewing and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: A meta-analysis. *JAMA* (2011) 305(23):2448–55. doi: 10.1001/jama.2011.812

39. Krishnan S, Rosenberg L, Palmer JR. Physical activity and television watching in relation to risk of type 2 diabetes: the black women's health study. *Am J Epidemiol* (2009) 169(4):428–34. doi: 10.1001/archinte.161.12.1542

40. Shiue I. Modeling indoor TV/screen viewing and adult physical and mental health: Health survey for England, 2012. *Environ Sci Pollut Res Int* (2016) 23 (12):11708–15. doi: 10.1007/s11356-016-6354-5

41. Li C, Beech B, Crume T, D'Agostino RBJr., Dabelea D, Kaar JL, et al. Longitudinal association between television watching and computer use and risk markers in diabetes in the SEARCH for diabetes in youth study. *Pediatr Diabetes* (2015) 16(5):382–91. doi: 10.1111/pedi.12163

42. Hsueh MC, Liao Y, Chang SH. Associations of total and domain-specific sedentary time with type 2 diabetes in Taiwanese older adults. *J Epidemiol* (2016) 26(7):348–54. doi: 10.2188/jea.JE20150095

43. Yu B, Gu Y, Bao X, Meng G, Wu H, Zhang Q, et al. Distinct associations of computer/mobile devices use and TV watching with depressive symptoms in adults: A large population study in China. *Depress Anxiety* (2019) 36(9):879–86. doi: 10.1002/da.22932

44. Martin WP, Sharif F, Flaherty G. Lifestyle risk factors for cardiovascular disease and diabetic risk in a sedentary occupational group: the Galway taxi driver study. *Irish J Med Sci* (2016) 185(2):403–12. doi: 10.1007/s11845-016-1442-6

45. Adedokun AO, Ter Goon D, Owolabi EO, Adeniyi OV, Ajayi AI. Prevalence, awareness, and determinants of type 2 diabetes mellitus among commercial taxi drivers in buffalo city metropolitan municipality south Africa: A cross-sectional survey. *Medicine* (2019) 98(9):e14652. doi: 10.1097/MD.0000000000014652

46. Kurosaka K, Daida H, Muto T, Watanabe Y, Kawai S, Yamaguchi H. Characteristics of coronary heart disease in Japanese taxi drivers as determined by coronary angiographic analyses. *Ind Health* (2000) 38(1):15–23. doi: 10.2486/indhealth.38.15

47. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and long-term weight gain in women and men. *N Engl J Med* (2011) 364 (25):2392–404. doi: 10.1056/NEJMoa1014296

48. Rana JS, Li TY, Manson JE, Hu FB. Adiposity compared with physical inactivity and risk of type 2 diabetes in women. *Diabetes Care* (2007) 30(1):53–8. doi: 10.2337/dc06-1456

49. Llavero-Valero M, Escalada-San Martín J, Martínez-González MA, Basterra-Gortari FJ, Gea A, Bes-Rastrollo M. Physical activity intensity and type 2 diabetes: Isotemporal substitution models in the "Seguimiento universidad de navarra" (SUN) cohort. *J Clin Med* (2021) 10(13):2744. doi: 10.3390/jcm10132744

50. Mikus CR, Oberlin DJ, Libla JL, Taylor AM, Booth FW, Thyfault JP. Lowering physical activity impairs glycemic control in healthy volunteers. *Med Sci Sports Exerc* (2012) 44(2):225–31. doi: 10.1249/MSS.0b013e31822ac0c0

51. Yang D, Yang Y, Li Y, Han R. Physical exercise as therapy for type 2 diabetes mellitus: From mechanism to orientation. *Ann Nutr Metab* (2019) 74(4):313–21. doi: 10.1159/000500110

52. Sylow L, Tokarz VL, Richter EA, Klip A. The many actions of insulin in skeletal muscle, the paramount tissue determining glycemia. *Cell Metab* (2021) 33 (4):758–80. doi: 10.1016/j.cmet.2021.03.020

53. Sardinha LB, Magalhaes JP, Santos DA, Judice PB. Sedentary patterns, physical activity, and cardiorespiratory fitness in association to glycemic control in type 2 diabetes patients. *Front Physiol* (2017) 8:262. doi: 10.3389/fphys.2017.00262

54. de Lemos ET, Oliveira J, Pinheiro JP, Reis F. Regular physical exercise as a strategy to improve antioxidant and anti-inflammatory status: benefits in type 2 diabetes mellitus. *Oxid Med Cell Longev* (2012) 2012:741545. doi: 10.1155/2012/741545

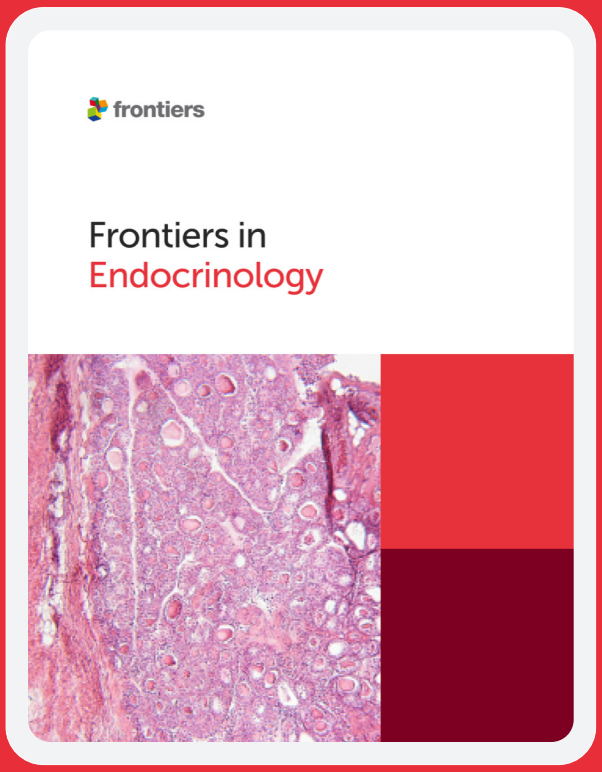
Frontiers in Endocrinology

Explores the endocrine system to find new therapies for key health issues

The second most-cited endocrinology and metabolism journal, which advances our understanding of the endocrine system. It uncovers new therapies for prevalent health issues such as obesity, diabetes, reproduction, and aging.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact