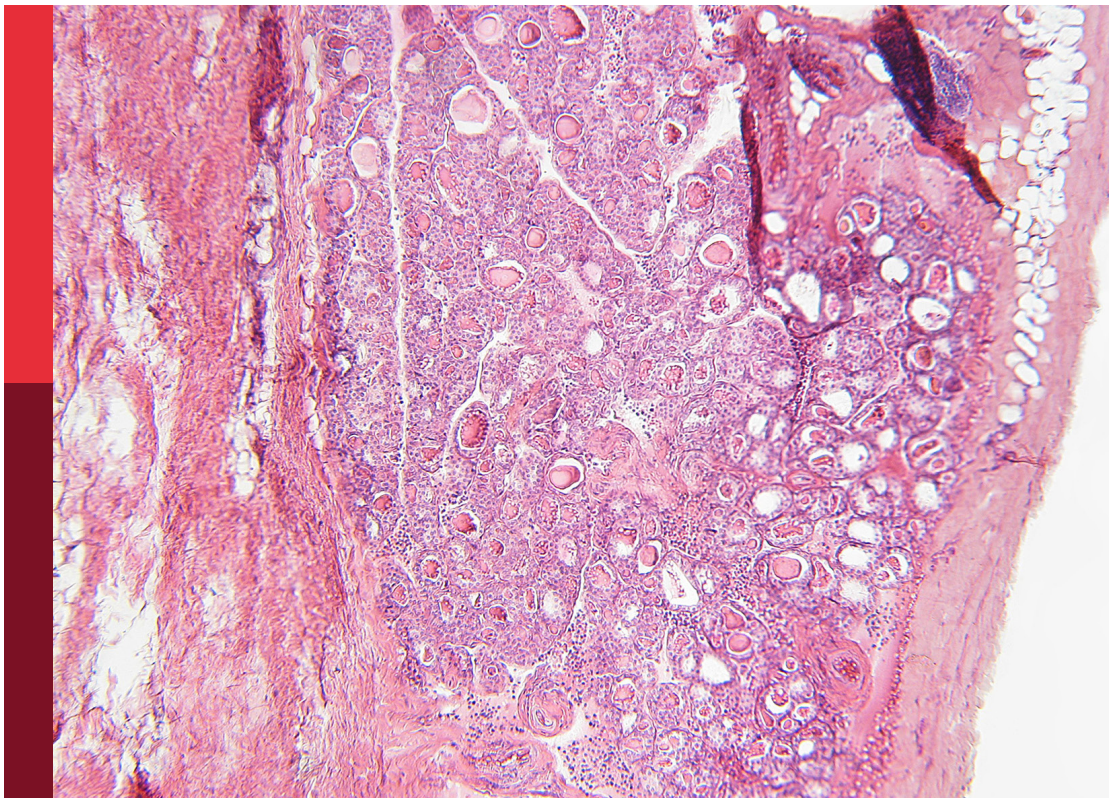
frontiers | Research Topics

# Exploring causal risk factors for metabolic and endocrine disorders

**Edited by**
Youxin Wang, Xiao Wang, Haifeng Hou
and Di Liu

**Published in**
Frontiers in Endocrinology
Frontiers in Immunology



ISSN 1664-8714
ISBN 978-2-8325-4465-5
DOI 10.3389/978-2-8325-4465-5

## 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

# Exploring causal risk factors for metabolic and endocrine disorders

**Topic editors**

Youxin Wang — North China University of Science and Technology, China
Xiao Wang — Lund University, Sweden
Haifeng Hou — Shandong First Medical University, China
Di Liu — Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences (CAS), China

**Citation**

Wang, Y., Wang, X., Hou, H., Liu, D., eds. (2024). *Exploring causal risk factors for metabolic and endocrine disorders*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-4465-5

# Table of contents

05 Editorial: Exploring causal risk factors for metabolic and endocrine disorders
Di Liu, Haifeng Hou, Xiao Wang and Youxin Wang

08 Hypoglycemia as a potential risk for patients taking clopidogrel: A systematic review and meta-analysis
Shi Chen, Jiaqi Qiang, Yuelun Zhang, Bin Zhao, Ran Tian, Tao Yuan, Ming Li, Mei Li, Yuxiu Li, Huijuan Zhu and Hui Pan

17 Causal relationship between polycystic ovary syndrome and chronic kidney disease: A Mendelian randomization study
Yufei Du, Fengao Li, Shiwei Li, Li Ding and Ming Liu

24 Identification of shared genetic architecture between non-alcoholic fatty liver disease and type 2 diabetes: A genome-wide analysis
Yajing Tan, Qian He and Kei Hang Katie Chan

34 The association between total bile acid and bone mineral density among patients with type 2 diabetes
Song Yang, Hongyun Li, Yuanyuan Gu, Qiang Wang, Li Dong, Chao Xu, Yuxin Fan, Ming Liu, Qingbo Guan and Lixing Ma

42 Remnant cholesterol, stronger than triglycerides, is associated with incident non-alcoholic fatty liver disease
Yiping Cheng, Qiang Zhang, Haizhen Li, Guangshuai Zhou, Ping Shi, Xu Zhang, Liying Guan, Fang Yan and Chao Xu

52 Metabolic drivers of dysglycemia in pregnancy: ethnic-specific GWAS of 146 metabolites and 1-sample Mendelian randomization analyses in a UK multi-ethnic birth cohort
Harriett Fuller, Mark M. Iles, J. Bernadette Moore and Michael A. Zulyniak

66 The genetically predicted causal relationship of inflammatory bowel disease with bone mineral density and osteoporosis: evidence from two-sample Mendelian randomization
Dengyong Xu, Yao Chen, Xing Gao, Weidong Xie, Ya Wang, Jiaying Shen, Guang Yang and Binbin Xie

74 Diagnostic value of serum cathepsin S in type 2 diabetic kidney disease
Xuejing Ren, Wanqing Wang, Huixia Cao and Fengmin Shao

85 The association between serum uric acid and hypertriglyceridemia: evidence from the national health and nutrition examination survey (2007–2018)
Mo-Yao Tan, Chao-Yue Mo, Fang Li and Qian Zhao

94 Assessing the usefulness of a newly proposed metabolic score for visceral fat in predicting future diabetes: results from the NAGALA cohort study
Ruijuan Yang, Maobin Kuang, Jiajun Qiu, Changhui Yu, Guotai Sheng and Yang Zou

106 Risk factors and drug discovery for cognitive impairment in type 2 diabetes mellitus using artificial intelligence interpretation and graph neural networks
Xin Zhang, Jiajia Xie, Xiong You and Houwu Gong

116 The causal association between polycystic ovary syndrome and susceptibility and severity of COVID-19: a bidirectional Mendelian randomization study using genetic data
Yu Si, Yuye Fei, Hua Ma, Yating Xu, Li Ning, Xiu Li and Qingling Ren

124 Association of serum 25-hydroxyvitamin D with urinary incontinence in elderly men: evidence based on NHANES 2007-2014
Li Liu, Mingming Xu, Hang Zhou, Xuexue Hao, Xiangyu Chen and Xiaoqiang Liu

131 The causal relationship between air pollution, obesity, and COVID-19 risk: a large-scale genetic correlation study
Jingwei Zhang, Jie Wen, Xin Wan and Peng Luo

143 Autoimmune polyglandular syndrome type 4: experience from a single reference center
Elisa Gatta, Valentina Anelli, Elena Cimino, Elena Di Lodovico, Elda Piovani, Irene Zammarchi, Giorgia Gozzoli, Virginia Maltese, Maria Cavadini, Barbara Agosti, Andrea Delbarba, Ilenia Pirola, Angela Girelli, Caterina Buoso, Francesca Bambini, Daniele Alfieri, Walter Bremi, Paolo Facondo, Roberto Lupo, Francesco Bezzi, Micaela Fredi, Anna Maria Mazzola, Elena Gandossi, Maura Saullo, Fiorella Marini, Massimo Licini, Letizia Chiara Pezzaioli, Laura Pini, Franco Franceschini, Chiara Ricci and Carlo Cappelli

153 Distinctive biochemistry profiles associated with hyperuricemia between Tibetans and Hans in China
Xue-Wen Ren, Kang Chen, Jue Wu, Zhang-Lin Yang, Tao Ji and Qing-Hong Zhang

172 The association of psychological stress with metabolic syndrome and its components: cross-sectional and bidirectional two-sample Mendelian randomization analyses
Cancan Li, Tianqi Tao, Yanyan Tang, Huimin Lu, Hongfeng Zhang, Huixin Li, Xiuhua Liu, Weiping Guan and Yixuan Niu

186 Mendelian randomization indicates causal effects of estradiol levels on kidney function in males
M. Kamal Nasr, Claudia Schurmann, Erwin P. Böttinger and Alexander Teumer

frontiers | Frontiers in Endocrinology

TYPE Editorial
PUBLISHED 01 February 2024
DOI 10.3389/fendo.2024.1357999



OPEN ACCESS

EDITED AND REVIEWED BY
Darko Stefanovski,
University of Pennsylvania, United States

*CORRESPONDENCE
Youxin Wang
sdwangyouxin@163.com;
wangyouxin@ncst.edu.cn

†These authors have contributed equally to this work

RECEIVED 19 December 2023
ACCEPTED 23 January 2024
PUBLISHED 01 February 2024

CITATION
Liu D, Hou H, Wang X and Wang Y (2024) Editorial: Exploring causal risk factors for metabolic and endocrine disorders. *Front. Endocrinol.* 15:1357999. doi: 10.3389/fendo.2024.1357999



# Editorial: Exploring causal risk factors for metabolic and endocrine disorders

Di Liu[1†], Haifeng Hou[2†], Xiao Wang[3†] and Youxin Wang[4*†]

[1]Center for Biomedical Information Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China, [2]School of Public Health, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, China, [3]Department of Clinical Sciences, Center for Primary Health Care Research, Lund University/Region Skåne, Malmö, Sweden, [4]School of Public Health, North China University of Science and Technology, Tangshan, China

KEYWORDS

**metabolic and endocrine disorders, risk factor, causal association, mendelian randomization, meta-analysis**

Editorial on the Research Topic
Exploring causal risk factors for metabolic and endocrine disorders

Metabolic and endocrine disorders remain a significant global health concern, with their prevalence steadily increasing in recent years. These disorders, including conditions but not limited to hypertension, diabetes and its complications, dyslipidemia, obesity, metabolic syndrome, hyperuricemia, non-alcoholic fatty liver disease, polycystic ovary syndrome, thyroid disorders, parathyroid disorders, pituitary disorders, and adrenal disorders, not only cause of morbidity and mortality and affect well-being but also pose substantial challenges to healthcare systems worldwide (1). The rising burden of metabolic and endocrine disorders has prompted extensive research to understand their causal risk factors. Identifying these factors is crucial for developing effective preventive strategies and improving public health outcomes. The main purpose of the editorial is to summarize the current state of research on causal risk factors for metabolic and endocrine disorders, highlighting the need for further investigation and collaboration in this critical area of study.

Etiological inference is the core of epidemiological research, which informs etiological modeling and prevention efforts. Large-scale epidemiological studies, genetic analyses, and clinical trials are the main study design for etiological inference, which have specific advantages and disadvantages that can complement each other to some extent (2, 3). In recent years, with the large-scale release of GWAS data, Mendelian randomization study has been widely applied in causal inference (4, 5), in which genetic data may partially address the limitations of confounding and reverse causality and provide more convincing evidence to explain the underlying causal associations. It remains challenges associated with identifying causal factors due to complex interactions between genetic, environmental, and lifestyle factors.

In this Research Topic, cross-sectional, cohort, meta-analysis, and Mendelian randomization study designs were used to explore the risk factors for metabolic and endocrine disorders. As metabolic and endocrine disorders could act as risk factors for diseases or adverse outcomes of risk factors, several studies investigated the bidirectional

causal relationship of metabolic and endocrine disorders as exposures or outcomes. In addition, there is a mutual relationship between metabolic disorders and endocrine dysregulation. When there is dysfunction in the endocrine system, it can lead to metabolic abnormalities and the development of metabolic disorders. On the other hand, metabolic disorders can also affect the functioning of the endocrine system.

The findings of Tan et al., determine the genetic structure shared between NAFLD and T2D, offering a new reference for the genetic pathogenesis and mechanism of NAFLD and T2D comorbidities.

Chen et al.'s systematic review and meta-analysis indicates that clopidogrel might be a modifiable and causal risk factor of hypoglycemia, especially in the Asian population.

Cheng et al. demonstrate that levels of triglycerides and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes independent of other risk factors, in the middle aged and elderly subset of the Chinese population, especially those who were women, non-CVD status, non-diabetes status and middle BMI status (24 to 28 kg/m$^2$).

In a Mendelian randomization study, Du et al. discover an important role of polycystic ovary syndrome in the development of chronic kidney disease, underlying the importance of regular follow-up of renal function in patients with polycystic ovary syndrome.

Xu et al. show that genetically predicted causal relationship of inflammatory bowel disease with bone mineral density and osteoporosis using evidence from two-sample Mendelian randomization.

Yang et al. demonstrate the potential role of bile acids in bone metabolism among T2DM patients in a north China population.

Using an ethnic-specific GWAS of 146 metabolites and 1-sample Mendelian randomization analyses in a UK multi-ethnic birth cohort, Fuller et al. confirm and demonstrate the presence of ethnic-specific causal relationships between metabolites and dysglycemia in mid-pregnancy in a UK population of SA and WE pregnant women.

Yang et al. assess the usefulness of a newly proposed metabolic score for visceral fat in predicting future diabetes, and finds that Metabolic Score for Visceral Fat is positively correlated with diabetes risk.

Ren et al. show that increased serum Cat-S is associated with the progression of albuminuria and decreased renal function in T2DM patients.

Nasr et al. conclude that serum estradiol levels may have a causal effect on kidney function, based on evidence from a MR analysis.

Gatta et al. show that 50% of patients developed autoimmune polyglandular syndrome type 4 within the first ten years, don't suggest any particular follow-up time and don't specify any particular disease.

Li et al. find that psychological stress is associated with hypertension both in cross-sectional and MR studies, suggesting targeting hypertension-related factors in interventions might improve mental and metabolic health.

Zhang et al. explore the risk factors for cognitive impairment (CI) in patients with type 2 diabetes mellitus (T2DM), screen potential therapeutic drugs for T2DM-CI, and provide evidence for preventing and treating T2DM-CI, using artificial intelligence interpretation and graph neural networks.

Based on data from the national health and nutrition examination survey (2007–2018), Tan et al. find an association between serum uric acid and hypertriglyceridemia.

Using data from the NHANES 2007-2014, Liu et al. show an association of serum 25-hydroxyvitamin D with urinary incontinence in elderly men.

Zhang et al. use clinical and genetic data from different public biological databases and perform two-sample and two-step Mendelian randomization analyses. The study suggests that exposure to heavy air pollutants causally increases risks for obesity.

Ren et al. investigate distinct risk factors for hyperuricemia in native Tibetan and immigrant Han populations in Tibet, China, and confirms the distinctive biochemistry between Tibetans and Hans.

Si et al. indicate a causality between polycystic ovary syndrome and susceptibility and severity of COVID-19 using a bidirectional Mendelian randomization study.

This unique insight will significantly contribute to the development of innovative prevention strategies for metabolic and endocrine disorders. This Research Topic aims to bring together important research findings from distinguished researchers and scientists worldwide. We strive to provide readers with ample opportunities to stay informed about the latest research and cutting-edge advances in this field through this collaboration. By fostering interdisciplinary collaboration and facilitating knowledge sharing, we firmly believe that we can pave the way for new avenues of research in the prevention of metabolic and endocrine disorders.

## Author contributions

YW: Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing. DL: Writing – original draft, Writing – review & editing. HH: Writing – original draft, Writing – review & editing. XW: Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by the National Key R&D Program of China (grant number 2017YFE0118800).

## 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



## References

1. G.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

2. Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol* (2016) 45:1776–86. doi: 10.1093/ije/dyv341

3. Liu D, Meng X, Tian Q, Cao W, Fan X, Wu L, et al. and multiple health outcomes: an umbrella review of observational studies, randomized controlled trials, and mendelian randomization studies. *Adv Nutr* (2022) 13:1044–62. doi: 10.1093/advances/nmab142

4. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Bmj* (2018) 362:k601. doi: 10.1136/bmj.k601

5. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. *Bmj* (2021) 375:n2233. doi: 10.1136/bmj.n2233

frontiers | Frontiers in Endocrinology

TYPE Systematic Review
PUBLISHED 28 February 2023
DOI 10.3389/fendo.2023.1091933



OPEN ACCESS

EDITED BY
Youxin Wang,
Capital Medical University, China

REVIEWED BY
Leila Warszawski,
Instituto Estadual de Diabetes e Endocrinologia Luiz Capriglione, Brazil
Zhelong Liu,
Huazhong University of Science and Technology, China
Yan Ying Guo,
People's Hospital of Xinjiang Uygur Autonomous Region, China

*CORRESPONDENCE
Hui Pan
panhui20111111@163.com

†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 07 November 2022
ACCEPTED 13 February 2023
PUBLISHED 28 February 2023

CITATION
Chen S, Qiang J, Zhang Y, Zhao B, Tian R, Yuan T, Li M, Li M, Li Y, Zhu H and Pan H (2023) Hypoglycemia as a potential risk for patients taking clopidogrel: A systematic review and meta-analysis.
*Front. Endocrinol.* 14:1091933.
doi: 10.3389/fendo.2023.1091933



# Hypoglycemia as a potential risk for patients taking clopidogrel: A systematic review and meta-analysis

Shi Chen[1†], Jiaqi Qiang[1,2†], Yuelun Zhang[3†], Bin Zhao[4], Ran Tian[5], Tao Yuan[1], Ming Li[1], Mei Li[1], Yuxiu Li[1], Huijuan Zhu[1] and Hui Pan[1,6*]

[1]Key Laboratory of Endocrinology of National Health Commission, Department of Endocrinology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [2]Eight-Year Program of Clinical Medicine, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [3]Medical Research Centre, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [4]Department of Pharmacy, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [5]Department of Cardiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [6]State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

**Background:** Clopidogrel is a cornerstone antiplatelet drug used in cardiovascular, cerebrovascular, and peripheral artery diseases. The sulfhydryl group of clopidogrel metabolite could induce insulin autoimmune syndrome (IAS) with hypoglycemia as the major symptom. Discontinuing clopidogrel and substituting it with ticagrelor has been revealed as an effective treatment in previous studies. Since hypoglycemia serves as a risk factor for cardiovascular and cerebrovascular events, we aimed to determine the association between hypoglycemia/IAS and clopidogrel and to investigate whether clopidogrel is a modifiable and causal risk factor of hypoglycemia/IAS.

**Methods:** MEDLINE, Embase, Cochrane databases, and clinical trial registries were searched for randomized controlled trials (RCTs) of clopidogrel from inception to 28 February 2022. RCTs comparing clopidogrel with placebo or other antiplatelet drugs were eligible if meeting the inclusion criteria: 1) clopidogrel was administrated 75 mg qd orally as a long-term antiplatelet prescription at least for months, and 2) hypoglycemia-inducible drugs were not used in the control arm. One investigator abstracted articles and performed a quality assessment. Uncertainties were resolved by discussions with two investigators independently. Odds ratio (OR) and risk difference (RD) were calculated and performed with subgroup analyses. The pre-specified protocol was registered in PROSPERO (CRD42022299622).

**Results:** Six trials with 61,399 participants in total fulfilled the criteria and were included in the meta-analysis. Clopidogrel might not be associated with higher hypoglycemia odds (OR 0.95, 95% CI 0.65 to 1.40). However, Asian participants ($p = 0.0437$) seemed more likely to develop clopidogrel-associated hypoglycemia. Clopidogrel-associated hypoglycemia occurred at the highest

rate of 0.03% (RD −0.00023, 95% CI −0.00077 to 0.00031), and this increased to 0.91% (RD 0.00210, 95% CI −0.00494 to 0.00914) in an aging population and to 0.18% (RD 0.00040, 95% CI −0.00096 to 0.00177) when Asian ratio of the population was elevated.

**Conclusions:** We raise the concern that clopidogrel might be a modifiable and causal risk factor of hypoglycemia. The Asian population might be more vulnerable and need additional care.

**Systematic review registration:** https://www.crd.york.ac.uk/prospero, identifier CRD42022299622.

KEYWORDS

clopidogrel, hypoglycemia, insulin autoimmune syndrome, meta-analysis, adverse event

# 1 Introduction

Hypoglycemia is a relatively uncommon condition with non-specific symptoms; however, its occurrence can be associated with daily performance disruption and physical injury involving the cardiovascular and central nervous systems (1, 2). Morbidities include myocardial ischemia, myocardial infarction, cardiac arrhythmia, transient ischemic attack, stroke, coma, seizure, cognitive impairment, and dementia (3, 4). Insulin autoimmune syndrome (IAS) is a cause of hypoglycemia characterized by hyperinsulinism and elevation in insulin autoimmune antibodies (IAAs), which might be triggered by drugs with sulfhydryl groups. The incidence is higher in populations with genetic background of susceptible human leukocyte antigen (HLA) alleles (1).

Clopidogrel is an anti-platelet drug prevalently prescribed for cardiovascular and cerebrovascular diseases, including acute coronary syndrome, myocardial infarction, stroke, and peripheral arterial disease. After activation in the hepatic P450 system, it transforms into a thiol derivative with a sulfhydryl group. As reported in previous cases (5), clopidogrel could induce IAS and trigger severe episodic hypoglycemia, making the drug threatening for the patients' primary diseases. The phenotype frequency of HLA-DRB1*0403–one of the most IAS-susceptible alleles—is 6.94% in Asians and 2.07% in Caucasians (6, 7) (calculation method is described in the Materials and Methods section). Given that the cardiovascular disease is a global burden and among the top leading causes of death (8, 9), it is reasonable to postulate there are several millions of individuals carrying susceptible HLA alleles and taking clopidogrel at the same time. Among this possible affected population, clopidogrel plays the role of a double-edged sword.

In this article, we were the first to raise the concern that if clopidogrel induced IAS in such a population, the harm could exacerbate the state of illness and even threaten their life. However, in previous studies comparing clopidogrel with other antiplatelet drugs, whether hypoglycemia was associated with clopidogrel remained undetermined. Whether hypoglycemia was more prevalent in populations with a higher frequency of susceptible HLA genotypes, such as Asians, also lacked evidence. Therefore, it was urgent to find out if clopidogrel was a risk factor for hypoglycemia through a systematic review and meta-analysis of randomized controlled trials (RCTs) of clopidogrel.

# 2 Materials and methods

## 2.1 Search strategy

The meta-analysis was conducted according to a pre-specified protocol (PROSPERO: CRD42022299622). RCTs comparing clopidogrel to other antiplatelet drugs or placebos were searched in clinical registries and published literature. Terms (clopidogrel OR Plavix) were searched on ClinicalTrials.gov, www.controlled-trials.com, and www.clinicaltrialsregister.eu from inception to 16 February 2022. The syntax (clopidogrel OR Plavix) was searched in free-text and MeSH terms on PubMed, Embase, and the Cochrane Library from inception to 28 February 2022. The search strategy in PubMed was as follows: ((“clopidogrel”[Title/Abstract] AND “clopidogrel”[MeSH Terms]) OR “plavix”[Title/Abstract]) AND (meta-analysis[Filter] OR meta analysis[Filter] OR systematic review [Filter]). Meta-analyses and systematic reviews were filtered out (see details in Supplementary Method). No restriction was set on language or publication time. For each search result, those meeting these criteria were included: a) systematic review or meta-analysis on RCTs of clopidogrel and b) clopidogrel was compared to placebo or other antiplatelet drugs. Those meeting these criteria were excluded: a) duplicated publications; b) clopidogrel was used in both groups; c) meta-analyses or systematic reviews on pharmacokinetics, pharmacodynamics, pharmacogenetics, or cost-effective studies; and d) *in vitro* or animal studies. Then, for each eligible meta-analysis and systematic review, abstracts and full texts were retrieved to screen and extract RCTs included in the search process.

## 2.2 Selection criteria

For each search result in the clinical registries and RCT extracted from the literature, we reviewed results displayed on the websites and related publications. Trial inclusion criteria were as follows: a) since trials with small sample size may have selection bias and overestimate the effect, we included trials recruiting 100 or more participants in total; b) hypoglycemia-inducible antiplatelet (or anticoagulant) drugs were not used in the control arm; c) since the optimal administrating dose of clopidogrel remained undetermined with variable and non-standard in children, we included trials enrolling adults (age ≥18 years); d) clopidogrel was administrated 75 mg qd orally as a long-term antiplatelet prescription at least for months; e) patients' combination drugs were balanced in the clopidogrel arm and the comparing arm; and f) the trial provided information on the occurrence of hypoglycemia. Drugs meeting these criteria were considered as hypoglycemia-inducible: a) chemical structures containing sulfhydryl groups and b) had been reported to induce hypoglycemia at therapeutic dose in literature (electronic searching strategy: generic name AND brand name AND (hypoglycemia OR hypoglycaemia OR (insulin autoimmune syndrome)).

## 2.3 Data extraction and quality assessment

A standard data extraction form was used to extract baseline characteristics, interventions, and adverse events amount of hypoglycemia. Other associated hypoglycemic symptoms, such as coma, seizure, and shock, were not counted. The quality of included trials was assessed using the Cochrane Collaboration risk-of-bias tool (10). Literature search, review, data extraction, and risk assessment were performed by JQ. Uncertainties were resolved by discussing with SC and consulting a third investigator (YZ) for arbitration.

## 2.4 Data synthesis and analysis

Publication bias was tested with a funnel plot and Egger's test (11). To quantify the potential risk of hypoglycemia in clopidogrel compared to control groups, odds ratios (ORs) and associated 95% confidence intervals (CIs) were first calculated using the random-effects model and the Mantel–Haenszel method. Statistical heterogeneity among trials was evaluated with $I^2$ statistics (12). Subgroup analyses of OR were performed by comparing clopidogrel with drugs in the control arm individually (ticagrelor, edoxaban, aspirin + extended-released dipyridamole, and placebo). In meta-regression analyses, the association between covariates (mean age, female ratio, Caucasian ratio, Asian ratio, and follow-up duration) and the OR was investigated using the DerSimonian–Laird method to estimate the between-study variance $Tau^2$ and the Hartung–Knapp method to adjust type I error. Sensitivity analyses were conducted by excluding trials with placebo, restricting to trials using aspirin as background treatment, using Peto's method to calculate the random-effects estimate of OR, and not using the Hartung–Knapp method adjustment in meta-regression.

To estimate the numbers needed to harm, a meta-analysis was next performed on risk difference (RD) by using the random-effects model and Mantel–Haenszel method. Subgroup analyses were performed using the median as the division boundaries, including mean age (>66.6 and ≤66.6 years), female sex ratio (>28.6% and ≤28.6%), Caucasian ratio (>81% and ≤81%), and Asian race ratio (>12% and ≤12%). Individual comparisons were also performed on clopidogrel versus drugs in the control arm (ticagrelor, placebo, aspirin + extended-released dipyridamole, and edoxaban).

Statistical analysis was completed in R (version 4.0.0) with the "meta" (version 4.16.2) (13) package. Confidence intervals not containing values of 1.00 and two-sided p-values of less than 0.05 were considered statistically significant.

## 2.5 HLA frequency calculation

HLA phenotype frequency is the percentage of individuals carrying the specific HLA allele in the population. We searched allele frequency (AF) and phenotype frequency (PF) of susceptible HLA alleles carried by IAS patients in the Allele Frequency Net Database (AFND) (6, 7). Asian race incorporated data of "Oriental" and "Asian" of ethnic origin labels in selection choice; Caucasian race incorporated "Caucasoid" and "Hispanic". For items with missing phenotype frequency, the formula $PF = 1 - (1 - AF)^2$ was used.

# 3 Results

A total of 752 publications of RCTs from databases and 325 RCTs from clinical trial registries were extracted. A total of 10 RCTs with reports of hypoglycemia were extracted. One trial that recruited less than 100 participants (NCT01864005), one trial that enrolled infants (NCT00396877), and two trials (14, 15) using prasugrel (containing sulfhydryl group in active metabolite) in the control arm were excluded. Six trials with 61,399 participants in total fulfilled the inclusion criteria and were included (Figure 1). Adverse events reported by 61,048 participants were included in the statistical analyses. Baseline characteristics are shown in Table 1. Risk-of-bias assessment identified some concerns for two of six trials (Supplementary Table 1). Egger's test yielded p = 0.453, and a funnel plot was graphed (Supplementary Figure 1). All trials were distributed symmetrically around the vertical stripped line, showing no underlying publication bias.

Compared to control groups, clopidogrel was associated with 5% decreased odds (OR 0.95, 95% CI 0.65 to 1.40) of hypoglycemia (Figure 2). Test of heterogeneity yielded $I^2$ = 0%, indicating low heterogeneity between trials. We undertook mono-factor meta-regression analyses to identify potential sources of trial heterogeneity. In subgroup analyses comparing clopidogrel with drugs in the control arm, clopidogrel versus ticagrelor yielded OR = 0.76 (95% CI 0.43 to

10.3389/fendo.2023.1091933

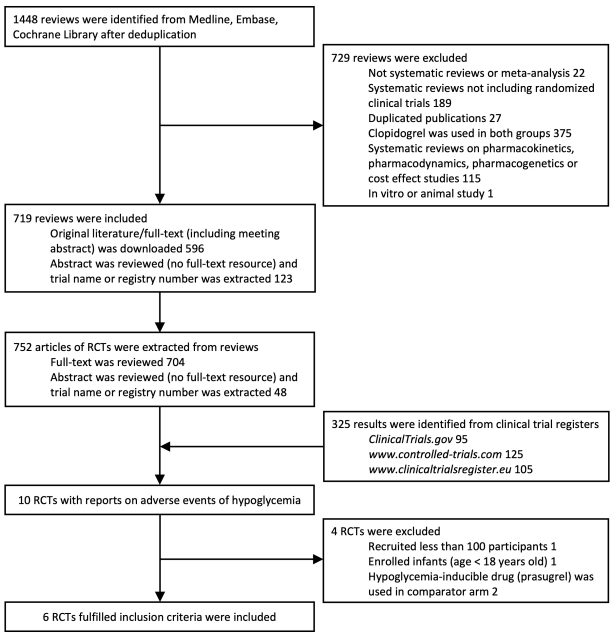



FIGURE 1
Flowchart of literature search to identify randomized controlled trials of clopidogrel.

1.34), clopidogrel versus edoxaban yielded OR = 3.00 (95% CI 0.12 to 74.53), clopidogrel versus aspirin + extended-released dipyridamole yielded OR = 1.17 (95% CI 0.67 to 2.02), and clopidogrel versus placebo yielded OR = 0.50 (95% CI 0.05 to 5.53) (Figure 3). Among the covariates, a higher Asian ratio was related to higher OR (p = 0.0437), whereas the others were not (age, p = 0.7486; female ratio, p = 0.5872; Caucasian ratio, p = 0.0501; follow-up duration, p = 0.8451) (Figure 4).

In sensitivity analyses, excluding trials with placebo (ACTIVE A) yielded OR 0.97 (95% CI 0.65 to 1.43); restricting to trials using aspirin as background treatment (ACTIVE A, PLATO, PHILO, and ePAD) yielded OR 0.87 (95% CI 0.39 to 1.93); using Peto method to calculate the random-effects estimate yielded OR 0.96 (95% CI 0.66 to 1.41); not using Hartung–Knapp method in meta-regression yielded no significant results of all covariates, with p = 0.7646 in age, p = 0.6157 in female ratio, p = 0.1315 in Caucasian ratio, p = 0.1246 in Asian ratio, and p = 0.8547 in follow-up duration.

Clopidogrel decreased the risk of hypoglycemia incidence by 0.023% (RD −0.00023, 95% CI −0.00077 to 0.00031) (Figure 5). In a subgroup of age older than 66.6 years, clopidogrel increased the risk by 0.210% (RD 0.00210, 95% CI −0.00494 to 0.00914). In a subgroup with an Asian ratio higher than 12%, clopidogrel increased the risk by 0.040% (RD 0.00040, 95% CI −0.00096 to 0.00177), whereas the other subgroups did not present increasing RD (Figure 6). To interpret, one in at least 3,220 patients taking clopidogrel was under threat of hypoglycemia manifesting in IAS. If the average age of the population exceeded 66.6 years, the number was 109. If the Asian ratio of the population increased, the number was one in at least 565. In subgroup analyses comparing clopidogrel with drugs in the control arm, clopidogrel versus ticagrelor yielded RD = −0.00038 (95% CI −0.00140 to 0.00063); clopidogrel versus edoxaban yielded RD = 0.00990 (95% CI −0.01717 to 0.03698), clopidogrel versus aspirin + extended-released dipyridamole yielded RD = 0.00029 (95% CI −0.00109 to 0.00167), and clopidogrel versus

TABLE 1 Baseline characteristics for participants in clopidogrel RCTs.

| Trial | Publication year | Indication | Treatment | Follow-up period | Mean age (years) | Female sex (%) | Race (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Caucasian | Asian |
| EUCLID (16) NCT01732822 (N = 13,885) | 2017 | Peripheral artery disease | Ticagrelor 90 mg bid vs clopidogrel 75 mg od | 30 months | 66.6 (8.4) | 28.0 | 0.81 | 0.12 |
| ACTIVE A (17) NCT00249873 (N = 7,554) | 2009 | Atrial fibrillation vascular risk | Clopidogrel 75 mg od vs placebo od, all received aspirin 75 to 100 mg od | 3.6 years | 71.0 (10.2) | 41.8 | 0.94 | 0.04 |
| PLATO (18) NCT00391872 (N = 18,624) | 2009 | Acute coronary syndrome | Ticagrelor 90 mg bid vs clopidogrel 75 mg od, all received aspirin 75 to 100 mg od | 13 months | 62.2 (22.4) | 28.4 | 0.92 | 0.06 |
| PHILO (19) NCT01294462 (N = 801) | 2015 | Acute coronary syndrome, percutaneous coronary intervention | Ticagrelor 90 mg bid vs clopidogrel 75 mg od, all received aspirin 75 to 100 mg od | 13 months | 67 (11) | 23.5 | 0.00 | 1.00 |
| PRoFESS (20) NCT00153062 (N = 20,332) | 2008 | Stroke | Aggrenox (aspirin 25 mg plus extended-release dipyridamole 200 mg) bid vs clopidogrel 75 mg qd, and telmisartan 80 mg/placebo qd | 2.5 years | 66.1 (8.6) | 36.0 | 0.62 | 0.33 |
| ePAD (21) NCT01802775 (N = 203) | 2018 | Peripheral arterial disease | Edoxaban 60 mg qd vs clopidogrel 75 mg qd, all received aspirin 100 mg qd | 6 months | 67.4 (9.5) | 28.6 | 0.94 | 0.00 |

N, the number of participants in actual enrollment; RCT, randomized controlled trial.

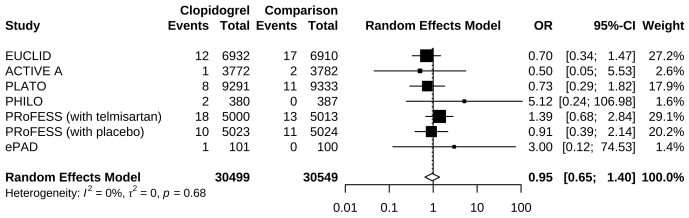



FIGURE 2
Forest plot of odds ratio estimation.

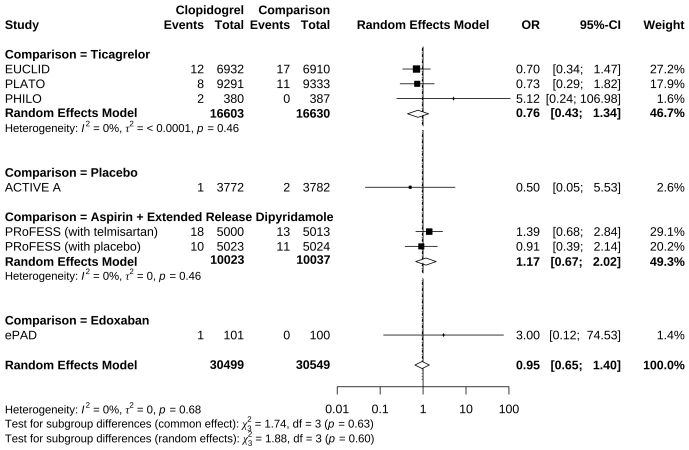



FIGURE 3
Forrest plot of odds ratio estimation in individual comparisons of clopidogrel versus drugs in the control arm.

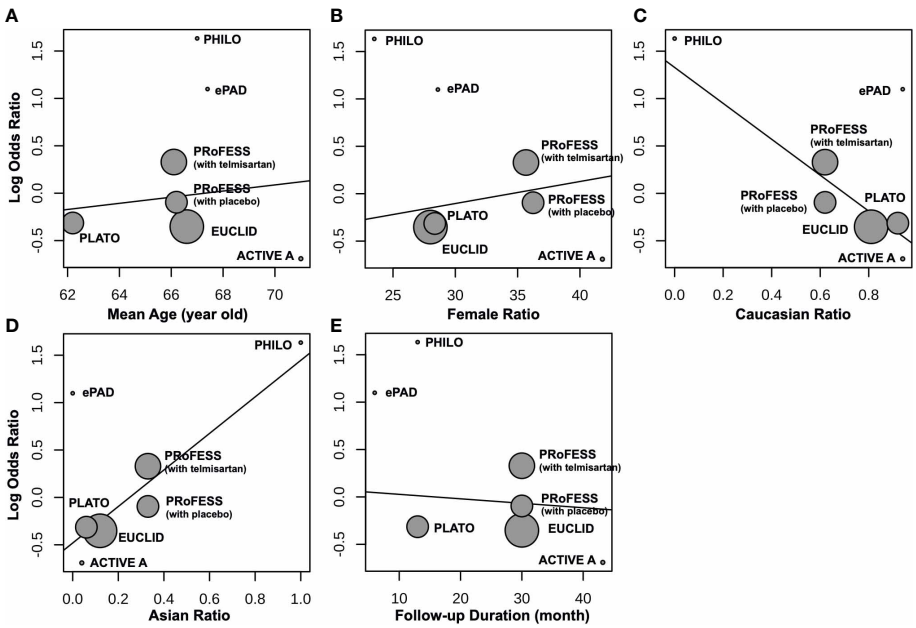



FIGURE 4
Meta-regression between clopidogrel-associated hypoglycemia and **(A)** mean age (years), **(B)** female ratio, **(C)** Caucasian ratio, **(D)** Asian ratio, and **(E)** follow-up duration (months).

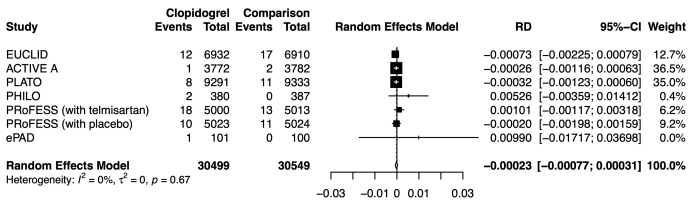



FIGURE 5
Forest plot of risk difference estimation.

placebo yielded RD = −0.00026 (95% CI −0.00116 to 0.00063) (Figure 7).

## 4 Discussion

This study focused on clopidogrel as a modifiable and causal risk of hypoglycemia manifesting in IAS. Since hypoglycemia deteriorates the risks of primary cardiovascular diseases, we conducted this research and generated the main findings, as follows: a) clopidogrel might not be associated with higher hypoglycemia risk, but Asians seemed more likely to develop clopidogrel-associated hypoglycemia; b) clopidogrel-associated hypoglycemia occurred at the highest rate of 0.03%, and this increased to 0.91% if it is an aging population and to 0.18% when Asian ratio of the population was elevated.

The mechanism for clopidogrel to trigger IAS has been revealed before. After ingestion, clopidogrel is transformed into a sulfhydryl derivative by hepatic enzymes CYP2C19, CYP2B6, CYP3A4, etc., which disrupts the disulfide bonds of the insulin molecule. The conformation-changed insulin molecule acquires immunogenicity and stimulates the proliferation of T cells and the production of IAAs. IAAs bind to insulin and dissociate in an unregulated way. When the insulin pool is discharged, blood glucose level decreases and induces hypoglycemia symptoms (22–24). In this autoimmune process, susceptible HLA alleles include HLA-DRB1*0406, DRB1*0403, and DQB1*0302 (25–27). In our calculation of HLA allele frequency (6), HLA-DRB1*0406 is carried by 3.59% Asians and 0.13% Caucasians, DRB1*0403 by 6.94% Asians and 2.07% Caucasians, and DQB1*0302 by 9.50% Asians and 30.40% Caucasians (calculation method is described in the Materials and Methods section).

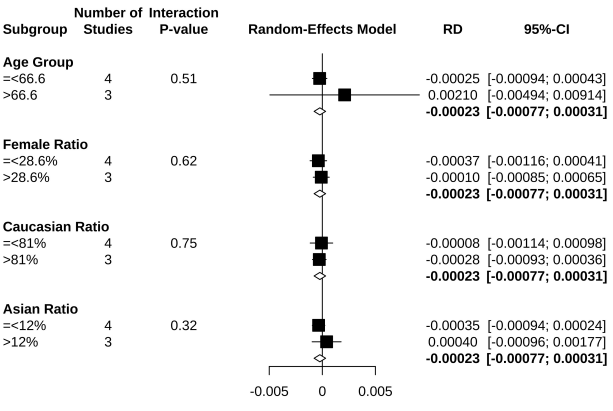



FIGURE 6
Forest plot of subgroup analysis in risk difference estimation.

Clopidogrel is commonly used in a massive number of cardiovascular, cerebral vascular, and peripheral artery disease patients. In the 2020 Medical Expenditure Panel Survey (MEPS) released by the Agency for Healthcare Research and Quality (AHRQ), the prescription of clopidogrel was 19,377,527, even higher than aspirin in the United States (28, 29). To multiply the prescription number and susceptible HLA allele frequency, the estimated number of victims is considerable. However, clopidogrel-induced IAS/hypoglycemia was rarely reported in either case or RCTs. Moreover, no systematic review or meta-analysis was conducted to reveal the association or evaluate the potential risks. The earliest case dated back to 2004, but the authors did not realize the triggering effect of clopidogrel on IAS (30). Until 2016, the concept of clopidogrel-induced IAS was formally proposed by Japanese scientists (22).

In the meta-analysis, clopidogrel appears to reduce the risk of hypoglycemia but with a non-confirmative trend. Result discrepancies between our postulation and meta-analysis could be attributed to several factors: a) hypoglycemia is a rare adverse event in traditional RCTs of non-diabetes mellitus participants and with strict recruitment criteria; b) clopidogrel metabolism depends on certain P450 enzyme, possibly not carried by all participants; c) participants of these trials were mainly Caucasians in their sixties, not balanced in age and race, whereas Asians are more susceptible to IAS due to genetic background; d) since hypoglycemia emerges asymptomatically or symptomatically with variable manifestation and onset time, it might not be identified during follow-up; e) some IAS are self-limiting and thus not reported (31).

Trials with higher Asian ratios seemed to have higher OR of clopidogrel-associated hypoglycemia, suggesting that Asians were probably more vulnerable. Consistently, geographic distribution was observed in IAS in previous studies (32). Asians, especially Japanese people, carrying HLA-DRB1*0406 at a higher frequency were more likely to develop IAS (33). This was validated by higher Asian carrier proportions of DRB1*0406 and DRB1*0403 alleles from our calculation (32). Although we also observed a considerable frequency of susceptible HLA allele carried by Caucasians, the odds

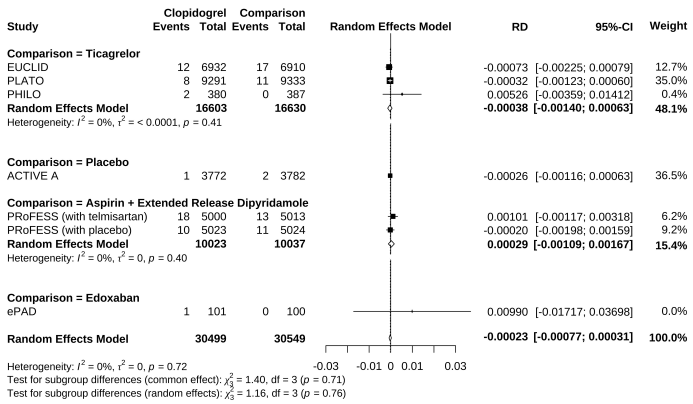



FIGURE 7
Forrest plot of risk difference estimation in individual comparisons of clopidogrel versus drugs in the control arm.

and risks of clopidogrel-induced hypoglycemia remain evidence-less in this population based on the meta-analysis result. Additional care might not be given to these patients.

As the association of clopidogrel and IAS/hypoglycemia is supported by biological plausibility and temporal relationship, we believe it necessary to estimate patients to be affected. One in at least 3,220 people (0.031%) in all patients, one in at least 109 people (0.91%) in the population with an average age above 66.6 years, and one in at least 525 people (0.18%) in higher-Asian-ratio subgroups might be affected by clopidogrel-induced hypoglycemia. Considering the massive sales of clopidogrel, we estimated that the actual number under threat was huge. Also, relatively rare adverse events have been drawing attention due to their serious outcomes, for instance, rhabdomyolysis induced by statins (at a risk of one in 66,667) (34) and allopurinol hypersensitivity syndrome (AHS) induced by allopurinol (at a risk of one in 1,000) (35).

Since clopidogrel administration is a modifiable and causal risk factor, drug discontinuation is the first-line therapy when hypoglycemia manifests. Drug substitution after clopidogrel discontinuation is also essential to maintain antiplatelet control over cardiovascular disease risk in treatment. The situation also depends on the clinical indications. For patients with non-ST-segment elevation acute coronary syndrome, a previous pharmacovigilance study based on data mining of the Food and Drug Administration Adverse Event Reporting System (FAERS) proposed ticagrelor as an alternative to substitute clopidogrel (36). Individual comparisons of clopidogrel versus drugs in the comparing arm in this meta-analysis might also provide information on which drug to switch to. Clopidogrel has higher (but not statistically significant) odds and risks than edoxaban and aspirin + dipyridamole to induce hypoglycemia while lower odds and risks (but not statistically significant) than ticagrelor. These results suggested that edoxaban is a possible substitute for patients with peripheral artery disease undergoing endovascular treatment and that aspirin + dipyridamole is a possible substitute for patients with ischemic stroke. Since our result lacked statistical significance, the inconsistent effect of clopidogrel *vs.* ticagrelor in the previous study probably needed further validation.

Although we revealed the association between hypoglycemia and clopidogrel according to widely acknowledged standards and guidelines, there were some limitations to mention. First, only aggregate data were available in included RCTs, whereas individual participant data (IPD) provide a more precise estimation of baseline characteristic effects. However, previous methodological research demonstrated that IPD and traditional aggregate data meta-analysis do not differ much in the combined effects (37). Second, observational studies were not included in our meta-analysis, so some cases with clopidogrel-induced IAS reported in previous publications may be missed in our evidence synthesis. Third, missing data might come from RCTs of clopidogrel published but not included in any of our searched clinical registries, meta-analysis or systematic review; or might have been published without reporting hypoglycemia. Fourth, the study population might contain more participants of certain races from developed countries that carried out more RCTs. Since hypoglycemia of IAS has geographical and racial distribution differences, the over-representation of participant races might lead to inaccurate estimation. Last, all included trials reported hypoglycemia instead of IAS. There was the possibility that some hypoglycemia events were not IAS manifestations. Moreover, to avoid duplication, reports on hypoglycemia-associated symptoms were excluded.

In summary, we were the first to explore the potential risk of IAS/hypoglycemia triggered by clopidogrel and estimated the incidence rate. We also pointed out that the Asian population might need extra attention. Clopidogrel substitution with alternative antithrombotic drugs might be effective in treatment. Future studies are expected to provide more robust evidence on the association between clopidogrel and IAS/hypoglycemia. We also recommend that hypoglycemia and IAS are specified as secondary endpoints in future large endpoint clopidogrel trials. We believe this study will arouse physicians' consciousness of hypoglycemia occurrence in prescribing clopidogrel and provide practical benefits for patients taking clopidogrel from a novel perspective.

## 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 author.

## Author contributions

SC, BZ, RT, TY, MiL, MeL, YL, and HZ conceived the study. SC, JQ, and YZ designed the study. JQ and YZ collected the systematic review data. SC, JQ, YZ, and HP analyzed and interpreted the data. SC, JQ, and YZ drafted the manuscript. All authors revised and approved the final version of the manuscript. All authors had full access to all the data in the study. HP accessed and verified the underlying data and had final responsibility for the decision to submit for publication. All authors contributed to the article and approved the submitted version.

## Funding

Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2021-I2M-1-023).

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1091933/full#supplementary-material

## References

1. Kittah NE, Vella A. Management of endocrine disease: Pathogenesis and management of hypoglycemia. *Eur J Endocrinol* (2017) 177:R37–47. doi: 10.1530/EJE-16-1062

2. Frier BM. Hypoglycaemia in diabetes mellitus: Epidemiology and clinical implications. *Nat Rev Endocrinol* (2014) 10:711–22. doi: 10.1038/nrendo.2014.170

3. Wright RJ, Frier BM. Vascular disease and diabetes: is hypoglycaemia an aggravating factor? *Diabetes Metab Res Rev* (2008) 24:353–63. doi: 10.1002/dmrr.865

4. Frier BM, Heller S, McCrimmon R. *Hypoglycaemia in clinical diabetes. 3rd Edition*. Chichester, West Sussex, UK: John Wiley & Sons Ltd., Wiley-Blackwell (2014). p. 392.

5. Rajpal A, Kassem LS, Moscoso-Cordero M, Arafah BM. Clopidogrel-induced insulin autoimmune syndrome: A newly recognized cause of hypoglycemia in a patient without diabetes. *J Endocr Soc* (2017) 1:1217–23. doi: 10.1210/js.2017-00316

6. Gonzalez-Galarza FF, McCabe A, Santos EJMD, Jones J, Takeshita L, Ortega-Rivera ND, et al. Allele frequency net database (AFND) 2020 update: Gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res* (2020) 48:D783–8. doi: 10.1093/nar/gkz1029

7. Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res* (2011) 39:D913–919. doi: 10.1093/nar/gkq1128

8. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. 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

9. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* (2018) 392:1736–88. doi: 10.1016/S0140-6736(18)32203-7

10. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* (2011) 343:d5928. doi: 10.1136/bmj.d5928

11. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* (1997) 315:629–34. doi: 10.1136/bmj.315.7109.629

12. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* (2002) 21:1539–58. doi: 10.1002/sim.1186

13. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with r: a practical tutorial. *Evid Based Ment Health* (2019) 22:153–60. doi: 10.1136/ebmental-2019-300117

14. Wiviott SD, Braunwald E, McCabe CH, Montalescot G, Ruzyllo W, Gottlieb S, et al. Prasugrel versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* (2007) 357:2001–15. doi: 10.1056/NEJMoa0706482

15. Roe MT, Armstrong PW, Fox KAA, White HD, Prabhakaran D, Goodman SG, et al. Prasugrel versus clopidogrel for acute coronary syndromes without revascularization. *N Engl J Med* (2012) 367:1297–309. doi: 10.1056/NEJMoa1205512

16. Hiatt WR, Fowkes FGR, Heizer G, Berger JS, Baumgartner I, Held P, et al. Ticagrelor versus clopidogrel in symptomatic peripheral artery disease. *N Engl J Med* (2017) 376:32–40. doi: 10.1056/NEJMoa1611688

17. ACTIVE Investigators, Connolly SJ, Pogue J, Hart RG, Hohnloser SH, Pfeffer M, et al. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. *N Engl J Med* (2009) 360:2066–78. doi: 10.1056/NEJMoa0901301

18. Wallentin L, Becker RC, Budaj A, Cannon CP, Emanuelsson H, Held C, et al. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* (2009) 361:1045–57. doi: 10.1056/NEJMoa0904327

19. Goto S, Huang C-H, Park S-J, Emanuelsson H, Kimura T. Ticagrelor vs. clopidogrel in Japanese, Korean and Taiwanese patients with acute coronary syndrome – randomized, double-blind, phase III PHILO study. *Circ J* (2015) 79:2452–60. doi: 10.1253/circj.CJ-15-0112

20. Sacco RL, Diener H-C, Yusuf S, Cotton D, Ounpuu S, Lawton WA, et al. Aspirin and extended-release dipyridamole versus clopidogrel for recurrent stroke. *N Engl J Med* (2008) 359:1238–51. doi: 10.1056/NEJMoa0805002

21. Moll F, Baumgartner I, Jaff M, Nwachuku C, Tangelder M, Ansel G, et al. Edoxaban plus aspirin vs dual antiplatelet therapy in endovascular treatment of patients with peripheral artery disease: Results of the ePAD trial. *J Endovasc Ther* (2018) 25:158–68. doi: 10.1177/1526602818760488

22. Yamada E, Okada S, Saito T, Osaki A, Ozawa A, Yamada M. Insulin autoimmune syndrome during the administration of clopidogrel. *J Diabetes* (2016) 8:588–9. doi: 10.1111/1753-0407.12385

23. Uchigata Y, Hirata Y, Iwamoto Y. Drug-induced insulin autoimmune syndrome. *Diabetes Res Clin Pract* (2009) 83:e19–20. doi: 10.1016/j.diabres.2008.10.015

24. Eisenbarth GS. *MEDICAL INTELLIGENCE UNIT 13: Molecular mechanisms of endocrine and organ specific autoimmunity*. Texas, U.S.A: R.G. Landes Company (1999).

25. Uchigata Y, Kuwata S, Tokunaga K, Eguchi Y, Takayama-Hasumi S, Miyamoto M, et al. Strong association of insulin autoimmune syndrome with HLA-DR4. *Lancet* (1992) 339:393–4. doi: 10.1016/0140-6736(92)90080-m

26. Uchigata Y, Omori Y, Nieda M, Kuwata S, Tokunaga K, Juji T. HLA-DR4 genotype and insulin-processing in insulin autoimmune syndrome. *Lancet* (1992) 340:1467. doi: 10.1016/0140-6736(92)92654-x

27. Uchigata Y, Tokunaga K, Nepom G, Bannai M, Kuwata S, Dozio N, et al. Differential immunogenetic determinants of polyclonal insulin autoimmune syndrome (Hirata's disease) and monoclonal insulin autoimmune syndrome. *Diabetes* (1995) 44:1227–32. doi: 10.2337/diab.44.10.1227

28. Agency for Healthcare Research and Quality (AHRQ), Rockville MD. Clopidogrel - drug usage statistics, ClinCalc DrugStats database, in: *Medical expenditure panel survey (MEPS) 2013-2020* (2022) (Accessed October 21, 2022).

29. Agency for Healthcare Research and Quality (AHRQ), Rockville MD. Aspirin - drug usage statistics, ClinCalc DrugStats database, in: *Medical expenditure panel survey (MEPS) 2013-2020* (2022) (Accessed October 21, 2022).

30. Yaturu S, DePrisco C, Lurie A. Severe autoimmune hypoglycemia with insulin antibodies necessitating plasmapheresis. *Endocr Pract* (2004) 10:49–54. doi: 10.4158/EP.10.1.49

31. Lupsa BC, Chong AY, Cochran EK, Soos MA, Semple RK, Gorden P. Autoimmune forms of hypoglycemia. *Med (Baltimore)* (2009) 88:141–53. doi: 10.1097/MD.0b013e3181a5b42e

32. Uchigata Y, Hirata Y, Omori Y, Iwamoto Y, Tokunaga K. Worldwide differences in the incidence of insulin autoimmune syndrome (Hirata disease) with respect to the evolution of HLA-DR4 alleles. *Hum Immunol* (2000) 61:154–7. doi: 10.1016/s0198-8859(99)00144-5

33. Censi S, Mian C, Betterle C. Insulin autoimmune syndrome: from diagnosis to clinical management. *Ann Transl Med* (2018) 6:335. doi: 10.21037/atm.2018.07.32

34. Francisco Lopez-Jimenez M.D. *Rhabdomyolysis from statins: What's the risk?*. (Oct 28th 2022). Available at: https://www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/expert-answers/rhabdomyolysis/faq-20057817.

35. Qurie A, Bansal P, Goyal A, Musa R. *Allopurinol*. Available at: https://www.ncbi.nlm.nih.gov/books/NBK499942/.

36. Chen S, Qiang J, Tian R, Yuan T, Li M, Li Y, et al. Clopidogrel-associated hypoglycemia and alternative antiplatelet therapy: A real-world, pharmacovigilance study. *Eur Heart J Cardiovasc Pharmacother* (2022) 9(1):8-9. doi: 10.1093/ehjcvp/pvac050

37. Tudur Smith C, Marcucci M, Nolan SJ, Iorio A, Sudell M, Riley R, et al. Individual participant data meta-analyses compared with meta-analyses based on aggregate data. *Cochrane Database Syst Rev* (2016) 9:MR000007. doi: 10.1002/14651858.MR000007.pub3

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 15 March 2023
DOI 10.3389/fendo.2023.1120119



OPEN ACCESS

EDITED BY
Xiao Wang,
Lund University, Sweden

REVIEWED BY
Akio Nakashima,
Jikei University School of Medicine, Japan
Slobodan Kapor,
University of Belgrade, Serbia
Milan Aksic,
Institute of Anatomy, University of Belgrade, Serbia

*CORRESPONDENCE
Li Ding
dinglitml@tmu.edu.cn
Ming Liu
mingliu@tmu.edu.cn

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

RECEIVED 09 December 2022
ACCEPTED 06 March 2023
PUBLISHED 15 March 2023

CITATION
Du Y, Li F, Li S, Ding L and Liu M (2023) Causal relationship between polycystic ovary syndrome and chronic kidney disease: A Mendelian randomization study. *Front. Endocrinol.* 14:1120119. doi: 10.3389/fendo.2023.1120119



# Causal relationship between polycystic ovary syndrome and chronic kidney disease: A Mendelian randomization study

Yufei Du, Fengao Li, Shiwei Li, Li Ding* and Ming Liu*

Department of Endocrinology and Metabolism, Tianjin Medical University General Hospital, Tianjin, China

**Objective:** Polycystic ovary syndrome is one of the most common endocrine disorders among women of childbearing age. The relationship between polycystic ovary syndrome and chronic kidney disease remains unclear and controversial. In this study, we investigated the causal role of polycystic ovary syndrome in the development of chronic kidney disease using the two-sample Mendelian randomization method.

**Methods:** Public shared summary-level data was acquired from European-ancestry genome wide association studies. We finally obtained 12 single nucleotide polymorphisms as instrumental variables, which were associated with polycystic ovary syndrome in European at genome-wide significance ($P < 5 \times 10^{-8}$). Inverse-variance weighted method was employed in the Mendelian randomization analysis and multiple sensitivity analyses were implemented. Outcome data were obtained from the Open GWAS database.

**Results:** A positive causal association was observed between polycystic ovary syndrome and chronic kidney disease (odds ratio [OR]=1.180, 95% confidence interval [CI]: 1.038-1.342; P=0.010). Further analyses clarified that causal relationship exist between polycystic ovary syndrome and some serological indicators of chronic kidney disease (fibroblast growth factor 23: OR= 1.205, 95% CI: 1.031-1.409, P=0.019; creatinine: OR= 1.012, 95% CI: 1.001-1.023, P=0.035; cystatin C: OR= 1.024, 95% CI: 1.006-1.042, P=0.009). However, there was no causal association of polycystic ovary syndrome with other factors in the data sources we employed.

**Conclusions:** Our results indicate an important role of polycystic ovary syndrome in the development of chronic kidney disease. This study suggests that regular follow-up of renal function in patients with polycystic ovary syndrome is necessary for the early treatment of chronic kidney disease.

KEYWORDS

**polycystic ovary syndrome, chronic kidney disease, Mendelian randomization, single nucleotide polymorphisms, causality**

# Introduction

Polycystic ovary syndrome (PCOS) is one of the most common endocrinopathies, affecting 4–21% of women among premenopausal women (1–3). The features of PCOS include oligomenorrhoea, infrequent ovulation, polycystic ovarian morphology (4) and a series of metabolic disorders including hyperandrogenism, hyperinsulinemia, gonadotropin imbalance, dyslipidemia, and often accompanied by increased visceral fat (5). The etiology and pathological mechanism of this disease remains unclear. Genome-wide association studies (GWAS) provided genetic evidence for further PCOS research (6–9). As mentioned above, patients with PCOS are accompanied by a series of metabolic disorders, and whether this leads to the occurrence of other chronic diseases is still lacking sufficient evidence.

Chronic kidney disease (CKD) is a common chronic disorder that refers to the long-term loss of renal function, eventually may progress into end-stage renal disease with a high rate of mortality (10). It affects millions of patients worldwide (11). The global burden of CKD is growing year after year, and almost 10% of adults around the world are affected by some subtypes of CKD, resulting in 1.2 million deaths and 28 million patients suffer from the disease each year (12, 13). CKD can occur for many reasons, including that more common and well-researched such as diabetes, glomerulonephritis, and cystic kidney diseases, however, the precise influencing factors of CKD are still not comprehensively studied (14). Recent studies have shown that long-term metabolic disorders, which could irreversibly impair the renal function and structure, are closely related to the occurrence of CKD (15).

The various metabolic disturbances caused by PCOS might lead to increasing incidence of CKD in later life in these PCOS patients, but few studies have focused on the long-term changes in renal function in patients with PCOS. Duleba et al. found that urinary albumin excretion (UAE) can occur in patients with PCOS with cardiovascular risk factors (16). Gozukara et al. found that although the test results of renal function were in the normal ranges in patients with PCOS, the GFR, urinary albumin excretion, and serum uric acid levels were all higher in PCOS patients than that in the controls (17). In previous animal experiments, Mohadetheh Moulana found that PCOS rats had higher renal mast cells infiltration which may lead to alteration of the immunological niche and persistent damage of kidney (18). Pruett et al. demonstrated that renal SGLT2 inhibitor could decreased the fat mass, plasma leptin, and additional therapies needed to improve renal injury in PCOS rats (19). These studies provide some supportive clues for us to understand the relationship between PCOS and CKD, but no human studies have been done to our knowledge to determine whether PCOS causes the occurrence of CKD or increase the risks for renal disease.

To gain further insight into the relationship between PCOS and CKD, we employed the Mendelian randomization (MR), a novel genetic epidemiological method that uses genetic variants as instrumental variables, to determine whether an observational association between a risk factor and an outcome is consistent with a causal effect (20). The MR methods capitalizes on the presumed random assortment of genes from parents to offspring, which provides unbiased detection of causal effects, since genetic variants are less susceptible to environmental factors (21). Using the findings for PCOS provided by the extensive GWAS data, we conducted two-sample (exposure and outcome measured in different samples) MR study to evaluate the causal effects of patients with PCOS on the risk of developing CKD, and further investigate the causal effects between genetic variants that determine the variation of PCOS and CKD risk factors, serological indicators of CKD, renal tubule injury biomarkers, separately.

# Methods

## Study design

We used a two-sample MR to investigate the causal relationship between PCOS and CKD or other factors including CKD risk factors, serological indicators of CKD and renal tubule injury biomarkers in the European population. For MR analysis, we used the inverse variance weighted (IVW) analysis accounting for correlations between instrument variables and outcomes. For the sensitivity analysis, weighted-median and MR-Egger analyses were performed. Exposure data derived from a meta-analysis, and ethics approval was obtained from the relevant research ethics committee (22). Outcome data can be obtained from the Open GWAS database (http://gwas.mrcieu.au.uk). Since this study was based on the public GWAS database, ethics approval was not required (23). This study was conducted based on the Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomization (STROBE-MR) guideline (24).

## Instrument variable selection

The SNPs associated to PCOS were derived from a large-scale genome-wide meta-analysis of polycystic ovary syndrome (22). The PCOS cohort included 113,238 European subjects (10,074 cases and 103,164 controls). We selected single nucleotide polymorphisms (SNPs) that have been shown to be significantly associated with PCOS (14 SNPs) at a genome-wide significance level ($P < 5 \times 10^{-8}$). Linkage disequilibrium analysis was performed based on a threshold of $r2 < 0.001$. In this study, 13 extracted SNPs passed the test and 1 was excluded (rs853854) for being palindromic with inter mediate allele frequencies (Supplementary Table 1). To avoid weak instrument bias, we calculated the F-statistic. In our study, F-statistics for the instrumental variables were strong (> 10) (25). These selected SNPs must meet the following three core assumptions. Firstly, SNPs are significantly correlated with exposure which means that the SNPs can predict exposure effectively. Secondly, the SNPs have to be independent of the outcome, namely the SNPs can only affect outcome through exposure. Thirdly, the SNPs must be independent of the confounding factors associated with exposure or outcome. Studies

have found that PCOS is associated with type 2 diabetes, obesity and hypertension, which may increase the risk of CKD. The above-mentioned risk factors may also be potential confounding factor between PCOS and CKD. In order to satisfy the assumptions, we conducted further analysis after taking into account confounding factors. To investigate whether PCOS associated SNPs are significantly correlated with other confounding factors at the genome-wide level, we searched the PhenoScanner database (http://www.phenoscanner.medschl.cam.ac.uk/) for these PCOS associated SNPs. We found that rs2271194 is associated with Body mass index (BMI), which is a risk factor for CKD. The remaining SNPs are not associated with common confounding factors interrelated to PCOS and CKD such as diabetes, hypertension, obesity, glomerulonephritis and polycystic kidney disease. To rule out potential confounding factors, rs2271194 was excluded and 12 SNPs were eventually included for further analyses.

## Outcome data source

The summary level dataset of CKD was retrieved from the Open GWAS database (https://gwas.mrcieu.ac.uk/), which included 216761 participants with CKD (3902 cases and 212841 controls). Summary-level statistics of the CKD risk factors were obtained from the Open GWAS database (23). we selected several well-established risk factors for CKD, including type 2 diabetes, autoimmune diseases, cystic kidney disease, obesity, and hypertension from a review of Chen et al. (26). In order to explore the relationship between PCOS and CKD serological indicators, we also obtained data from the database and selected some serological indicators that would be used in clinical practice for analysis, mainly including fibroblast growth factor 23, creatinine, cystatin C, estimated glomerular filtration rate, β2-microglobulin and urinary albumin-to-creatinine ratio. Finally, to explore whether renal tubular injury is associated with PCOS, we also obtained biomarkers of renal tubular injury from the Open GWAS database and selected a part of common biomarkers for analysis, including insulin-like growth factor-binding protein 7, epidermal growth factor, monocyte chemoattractant protein-1, tumor necrosis factor receptor 1, tumor necrosis factor receptor 2, chitinase-3-like protein 1, kidney injury molecule 1, neutrophil gelatinase-associated lipocalin, interleukin 6 and interleukin 10 from a review of Zhang et al. (27). All data for analysis are based on European population.

## Statistical analysis

We utilized 12 genetic variants as instrumental variables to estimate the causal association between PCOS and CKD and other factors. In this study, variable selection was relied on prior information. We selected outcome variables based on previous research reports to be as close to the clinical application as possible, and we also pay attention to the latest biomarkers to explore the possibility of new discoveries. In the present study, the robust methods: inverse variance weighting (IVW), MR Egger, and weighted median were employed to assess the causal relationship. IVW, the main dependent method, equates to implement a weighted linear regression of the associations between the outcome instrumental variables (IVs) and the exposure, merges the Wald ratio estimates of each IV in a meta-analysis manner (28, 29). Sensitivity analyses, including MR-Egger and weighted median. MR Egger considers the horizontal pleiotropic effects, in which the intercept may provide evidence for potential pleiotropic effects across IVs, although its estimates may be imprecise. The weighted median estimator is a method in which at least 50% of the weights are considered valid instrumental variables (30). Weighted median method allows some genetic variants are invalid, but only if at least half of them are valid instruments (31). Heterogeneity is another major concern in MR analyses, suggesting the possible concurrent presence of pleiotropy. To evaluate the heterogeneity between the IVs, IVW, MR-Egger, and Maximum likelihood methods were used. Cochran's Q statistic was employed to quantify the heterogeneity.

# Results

## The association between PCOS and CKD

We first used the selected 12 SNPs as the instrumental variables to assess the association between CKD and PCOS. The inverse variance weighted (IVW) method was used for the MR analysis, while the other four methods (MR Egger, Simple mode, Weighted median and Weighted mode) were applied to confirm the robustness of the results (Figure 1). In the IVW analysis, the effect odds ratio (OR) of PCOS on CKD was 1.180 (95% CI, 1.038-1.342; P = 0.010) (Table 1), which positive association exists between PCOS and CKD. MR-Egger do not detect potential horizontal pleiotropy for PCOS (P = 0.726), and there was no significant heterogeneity detected through Cochran's Q test with both MR-Egger and IVW methods (P = 0.361, P = 0.435, respectively).

## No causal relationship between PCOS and CKD risk factors

The occurrence of CKD is associated with many risk factors. We wondered whether PCOS could contribute to the occurrence of CKD by enhancing the risk factors of CKD. Several common CKD risk factors were selected and analyzed by multiple methods (MR Egger, WM, IVW), and none of them were statistically significant. The beta value and p value are as follows: Type 2 diabetes (0.964, 95% CI, -0.908-1.024; P = 0.235), Autoimmune diseases (1.012, 95% CI, 0.952-1.076; P = 0.704), Cystic kidney disease (1.089, 95% CI, 0.790-1.500; P = 0.604), Obesity class 1 (defined as BMI ≥ 30 kg/m2, -0.979, 95% CI, -0.906-1.058; P = 0.594), Obesity class 2 (defined as BMI ≥ 35 kg/m2, -0.904, 95% CI, -0.798-1.025; P = 0.117), Obesity class 3 (defined as BMI ≥ 40 kg/m2, -0.940, 95% CI, 0.736-1.200; P = 0.618), Hypertension (0.999, 95% CI, 0.994-1.004; P = 0.585) (Supplementary Figure 1). Significant heterogeneity was observed between hypertension and PCOS (IVW analysis: Q = 21.780,

10.3389/fendo.2023.1120119

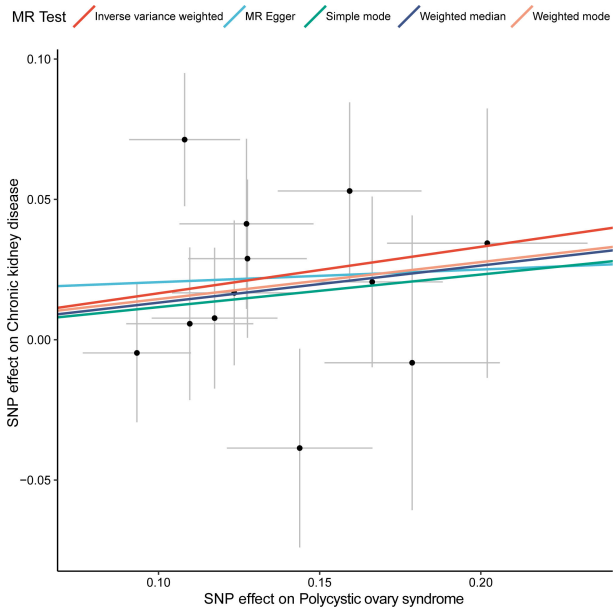



FIGURE 1
Two-sample MR analysis of the effect of PCOS on CKD using different methods. X-axis represents the effect of SNP on exposure (polycystic ovary syndrome) and the Y-axis represents the effect of SNP on outcome (chronic kidney disease), and the colored lines indicate the results of fitting different MR methods. MR, Mendelian randomization; SNP, single-nucleotide polymorphism.

P=0.026). No horizontal pleiotropy was identified by MR-Egger test.

## PCOS affects the changes of fibroblast growth factor 23, creatinine and cystatin C of CKD

Serological indicators are generally used to diagnose diseases or reflect the severity of diseases. We want to know whether there is a causal relationship between some serological indicators of CKD and PCOS. The IVW analysis showed that PCOS and some indicators have causal relationship such as fibroblast growth factor 23 (1.205, 95% CI, 1.031-1.409; P = 0.019), creatinine (0.012, 95% CI, 1.001-1.023; P = 0.035), cystatin C (1.024, 95% CI, 1.006-1.042; P = 0.009). There are no relationships between PCOS and other serological indicators: estimated glomerular filtration rate (-0.803, 95% CI, 0.470-1.371; P = 0.421), β2-microglobulin (1.233, 95% CI, 0.850-1.791; P = 0.270), urinary albumin-to-creatinine ratio (0.994, 95% CI, 0.925-1.067; P = 0.861) (Figure 2). The heterogeneity was observed between PCOS and serological indicators for CKD by the IVW analysis: fibroblast growth factor 23 (Q = 5.593, P= 0.848), creatinine (Q =12.209, P = 0.348), cystatin C (Q = 24.178, P= 0.012), estimated glomerular filtration rate (Q = 14.978, P= 0.092), β2-microglobulin (Q =1.172, P = 0.760), urinary albumin-to-creatinine ratio (Q = 33.101, P<0.001). For the horizontal pleiotropy, we found that there was no evidence between PCOS and serological indicators of CKD: fibroblast growth factor 23 (P= 0.971), creatinine (P = 0.212), cystatin C (P= 0.357), estimated glomerular filtration rate (P= 0.180), β2-microglobulin (P = 0.600), except for urinary albumin-to-creatinine ratio (P= 0.046).

## No causal association between PCOS and renal tubule injury biomarkers

An increasing number of renal tubule injury biomarkers are being discovered and applied in the diagnosis and treatment of CKD. However, whether there should be a relationship between PCOS and renal tubule injury biomarkers still remains a mystery.

TABLE 1 The causal effect between PCOS and CKD.

| Exposure | Outcome | Methods | nSNP | SE | OR (95%CI) | P value |
|---|---|---|---|---|---|---|
| Polycystic ovary syndrome | Chronic kidney disease | MR Egger | 12 | 0.341 | 1.046(0.536,2.042) | 0.898 |
| | | Weighted median | 12 | 0.086 | 1.141(0.963,1.352) | 0.127 |
| | | Inverse variance weighted | 12 | 0.065 | 1.180(1.038,1.342) | 0.011 |
| | | Simple mode | 12 | 0.146 | 1.123(0.844,1.495) | 0.443 |
| | | Weighted mode | 12 | 0.131 | 1.141(0.882,1.475) | 0.337 |

SNP, single-nucleotide polymorphism; SE, standard error; OR, odds ratio; 95%CI, 95% confidence interval.

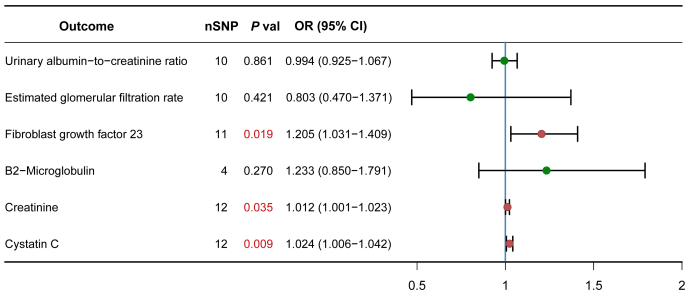



FIGURE 2
Causal effect of PCOS on CKD Serological indicators. Forest plots showing the range of OR values for different serological indicators. The green point (p>0.05) and red point (p<0.05) represents OR values of each CKD serological indicators, the vertical lines on either side of the point represent the 95% confidence interval. OR, odds ratio; 95%CI, 95% confidence interval.

We next performed IVW analyses for PCOS and those biomarkers. The result shows that there was no causal relationship between the two: insulin-like growth factor-binding protein 7 levels (0.918, 95% CI, 0.741-1.135; P = 0.429), epidermal growth factor levels (0.978, 95% CI, 0.902-1.060; P = 0.588), monocyte chemoattractant protein-1 levels (1.036, 95% CI, 0.959-1.118; P = 0.368), tumor necrosis factor receptor 1 levels (1.035, 95% CI, -0.970-1.104; P = 0.301), tumor necrosis factor receptor 2 levels (1.063, 95% CI, 0.974-1.160; P = 0.173), chitinase-3-like protein 1 levels (1.014, 95% CI, 0.944-1.088; P = 0.712), kidney injury molecule 1 levels (1.006, 95% CI, -0.937-1.080; P = 0.868), neutrophil gelatinase-associated lipocalin (1.033, 95% CI, 0.880-1.214; P = 0.688), interleukin 6 (1.118, 95% CI, 0.956-1.306; P = 0.163), interleukin 18 (1.042, 95% CI, 0.854-1.270; P = 0.688), interleukin 10 (0.985, 95% CI, 0.578-1.681; P = 0.957) (Supplementary Figure 2). For the heterogeneity, we found that monocyte chemoattractant protein-1 levels (IVW: Q = 25.601, P= 0.007), tumor necrosis factor receptor 2 levels (IVW: Q = 34.016, P< 0.001), chitinase-3-like protein 1 levels (IVW: Q = 22.586, P= 0.020) were statistically significant. MR-Egger do not detect potential horizontal pleiotropy.

## Discussion

Previous researches have demonstrated that comorbidities related to PCOS included obesity, type 2 diabetes, cardiovascular disease, etc., but the relationship between PCOS and CKD is still contradictory (32–34). Patil et al. showed that women with PCOS may be at increased risk for development of CKD with advanced age (32). Behboudi-Gandevan et al. found that the risk of CKD in patients with PCOS was like the common female population, and larger studies with long-term follow-up was needed (33). Since there is no consensus, this study used Mendelian randomization to explore the potential casual association of PCOS and CKD. Our result suggested that the association between PCOS and CKD was statistically significant when using IVW analysis, and the trend of other sensitivity analyses were consistent with the results of IVW, which indicated that the causal relationship between PCOS and CKD we obtained is relatively robust. Our findings indicated that PCOS is causally linked to CKD. To learn more about the role and impact of PCOS in CKD development, the risk factors were evolved as outcomes and MR analysis were conducted (26). Several studies suggest that PCOS causes CKD because of metabolic-related factors such as obesity and disordered hormone secretion (32, 33). However, our research did not find the causal relationship between PCOS and CKD risk factors. Future studies with larger sample sizes need to be performed to clarify the relationship between the two.

The clinical diagnosis of CKD mainly relies on the classical definition: estimated glomerular filtration rate (eGFR) <60 ml/min/1.73m$^2$, or markers of kidney damage, such as albuminuria, hematuria or abnormal structure on imaging, persistent for at least 3 months (15). With the deepening of research on CKD diagnosis and treatments, other indicators were validly used in fundamental and clinical research. Elevated levels of urinary albumin-to-creatinine ratio often indicates glomerular injury. Previous reports showed the high level of urinary albumin-to-creatinine ratio in patients with PCOS (35).β2-microglobulin is another early and sensitive diagnostic indicator of kidney disease. However, our results suggested no causal relationship between PCOS and those indicators mentioned above (eGFR, urinary albumin-to-creatinine ratio and β2-microglobulin). Creatinine was used as an important indicator for CKD patients, and it also formed a collaboration equation for estimating glomerular filtration rate (36). New eGFR equations that incorporate creatinine and cystatin C (37) could provide more accurate information for diagnosing CKD. Thus, we also studied Cystatin C (Cys C), another novel predictor for the development of CKD (38). Cys C has been proposed as a potential glomerular filtration rate (GFR) marker for the early detection of CKD and skin interstitial fluid (ISF) Cys C equipment has been developed (39). Cystatin-C levels was associated with an increase in IL-6 and a decrease in SOD in PCOS patients. This result suggested immune dysregulation exist in PCOS patients may affect the function of other organs (40). Fibroblast growth factor 23 (FGF23), a bone-derived phosphaturic hormone for regulating 1,25-dihydroxyvitamin D3 (calcitriol), could decrease the number of sodium-phosphate cotransporter 2a (NaPi-2a) on the basolateral

membrane of proximal tubule cells. In recent years, the role of FGF23 in CKD has gradually been discovered. Makoto Kuro-o considered that FGF23 increase is deemed necessary to compensate for the decrease in the nephron number during CKD progression (41). Previously studies have reported that FGF23 increasing, calcitriol decreasing, and parathyroid hormone increasing occur in this order during CKD progression (41). Another research suggested that decreased FGF23 could improve outcomes in CKD (42). Anyhow, FGF23 has been served as a marker and potential pathogenic factor for CKD progression (41). We performed MR analysis to determine whether PCOS influenced those indicators of CKD. These results showed that there might be a causal relationship between PCOS and those three indicators including creatinine, Cys C and FGF23 which hint a possible association between PCOS and CKD.

Renal tubular damage is the important pathological process leading from CKD to the end stage of various renal disease (43). We wondered whether PCOS might be directly targeting the renal tubular. We selected reported renal tubular injury markers to explore whether PCOS is causally related to them (27). The results showed that PCOS could not cause the changes of these biomarkers, which might imply that CKD induced by PCOS is not mediated by these biomarkers of renal tubular injury.

Overall, Mendelian randomization (MR) is a powerful tool in investigating the relationship between exposures and outcomes, it could help prioritize potential causal relationships. The advantage of MR is that confounding factors could be controlled, and the wastes of manpower and material resources could be overcome comparing with the randomized controlled trial. In addition, MR studies are generally considered to provide higher-quality evidence than observational studies. Our results suggested that PCOS can lead to CKD, might be an independent causal relationship. Nevertheless, the study also has certain limitations that must be acknowledged. The sample size is a limiting factor in such studies. The larger the sample size, the more reliable the conclusion. Gender factor is a possible bias in the current study because the CKD data included data on mixed-gender loci, but the PCOS data included only female data. This may have influenced the results to some extent, but it is difficult to isolate female CKD data alone. For further study, a long-term cohort stratified by age and sex is needed to validate our conclusions. In addition, our study sample was all European ancestry, so the generalizability of our results to other races/ethnicities is uncertain.

This is the first study to investigate the causal relation of PCOS and CKD according to our knowledge. In conclusion, our results revealed that there is a causal relationship between PCOS and CKD, but whether this causal relationship is direct or indirect is still unclear, large-scale prospective cohort studies are needed to validate the preliminary findings of our study, and even more basic experimental research need to be done with human tissue, to further clarify the deep molecular mechanisms. The clinical significance of this study is that in addition to the prevention of cardiovascular and diabetes complications in patients with PCOS, more attentions should also be paid to the long-term follow-up of renal function to provide an opportunity for early intervention to slow down the progression of CKD.

## 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

YD performed analyses and wrote the manuscript. SL, LD, and FL helped to analyze the results and edited the manuscript. ML guided the study and edited the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Natural Science Foundation of China (81830025, 81620108004); we acknowledge the support of the National Key R&D Program of China 2019YFA0802502, and Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-030A).

## 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.2023.1120119/full#supplementary-material

# References


1. Escobar-Morreale HF. Polycystic ovary syndrome: Definition, aetiology, diagnosis and treatment. *Nat Rev Endocrinol* (2018) 14(5):270–84. doi: 10.1038/nrendo.2018.24

2. March WA, Moore VM, Willson KJ, Phillips DI, Norman RJ, Davies MJ. The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. *Hum Reprod (Oxford England)* (2010) 25(2):544–51. doi: 10.1093/humrep/dep399

3. Tehrani FR, Simbar M, Tohidi M, Hosseinpanah F, Azizi F. The prevalence of polycystic ovary syndrome in a community sample of Iranian population: Iranian pcos prevalence study. *Reprod Biol Endocrinol RB&E* (2011) 9:39. doi: 10.1186/1477-7827-9-39

4. Azziz R, Carmina E, Chen Z, Dunaif A, Laven JS, Legro RS, et al. Polycystic ovary syndrome. *Nat Rev Dis Primers* (2016) 2:16057. doi: 10.1038/nrdp.2016.57

5. Goodarzi MO, Dumesic DA, Chazenbalk G, Azziz R. Polycystic ovary syndrome: Etiology, pathogenesis and diagnosis. *Nat Rev Endocrinol* (2011) 7(4):219–31. doi: 10.1038/nrendo.2010.217

6. Day FR, Hinds DA, Tung JY, Stolk L, Styrkarsdottir U, Saxena R, et al. Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat Commun* (2015) 6:8464. doi: 10.1038/ncomms9464

7. Hayes MG, Urbanek M, Ehrmann DA, Armstrong LL, Lee JY, Sisk R, et al. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations. *Nat Commun* (2015) 6:7502. doi: 10.1038/ncomms8502

8. Shi Y, Zhao H, Shi Y, Cao Y, Yang D, Li Z, et al. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet* (2012) 44 (9):1020–5. doi: 10.1038/ng.2384

9. Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Shi Y, et al. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet* (2011) 43(1):55–9. doi: 10.1038/ng.732

10. Fraser SD, Roderick PJ, May CR, McIntyre N, McIntyre C, Fluck RJ, et al. The burden of comorbidity in people with chronic kidney disease stage 3: A cohort study. *BMC Nephrol* (2015) 16:193. doi: 10.1186/s12882-015-0189-z

11. Thomas B, Matsushita K, Abate KH, Al-Aly Z, Ärnlöv J, Asayama K, et al. Global cardiovascular and renal outcomes of reduced gfr. *J Am Soc Nephrol JASN* (2017) 28(7):2167–79. doi: 10.1681/asn.2016050562

12. Xie Y, Bowe B, Mokdad AH, Xian H, Yan Y, Li T, et al. Analysis of the global burden of disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. *Kidney Int* (2018) 94(3):567–81. doi: 10.1016/j.kint.2018.04.011

13. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet (London England)* (2020) 395(10225):709–33. doi: 10.1016/s0140-6736(20)30045-3

14. Dincer N, Dagel T, Afsar B, Covic A, Ortiz A, Kanbay M. The effect of chronic kidney disease on lipid metabolism. *Int Urol Nephrol* (2019) 51(2):265–77. doi: 10.1007/s11255-018-2047-y

15. Kalantar-Zadeh K, Jafar TH, Nitsch D, Neuen BL, Perkovic V. Chronic kidney disease. *Lancet (London England)* (2021) 398(10302):786–802. doi: 10.1016/s0140-6736(21)00519-5

16. Duleba AJ, Ahmed IM. Predictors of urinary albumin excretion in women with polycystic ovary syndrome. *Fertil Steril* (2010) 93(7):2285–90. doi: 10.1016/j.fertnstert.2008.12.120

17. Gozukara IO, Gozukara KH, Kucur SK, Karakılıc EU, Keskin H, Akdeniz D, et al. Association of glomerular filtration rate with inflammation in polycystic ovary syndrome. *Int J Fertil Steril* (2015) 9(2):176–82. doi: 10.22074/ijfs.2015.4238

18. Moulana M. Immunophenotypic profile of leukocytes in hyperandrogenemic female rat an animal model of polycystic ovary syndrome. *Life Sci* (2019) 220:44–9. doi: 10.1016/j.lfs.2019.01.048

19. Pruett JE, Torres Fernandez ED, Everman SJ, Vinson RM, Davenport K, Logan MK, et al. Impact of sglt-2 inhibition on cardiometabolic abnormalities in a rat model of polycystic ovary syndrome. *Int J Mol Sci* (2021) 22(5):2576. doi: 10.3390/ijms22052576

20. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *Jama* (2017) 318(19):1925–6. doi: 10.1001/jama.2017.17219

21. 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

22. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, et al. Large-Scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PloS Genet* (2018) 14(12):e1007813. doi: 10.1371/journal.pgen.1007813

23. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, et al. The mrc ieu opengwas data infrastructure. *bioRxiv* (2020) 2020:08. 10.1101/2020.08.10.244293

24. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (Strobe) statement: Guidelines for reporting observational studies. *Int J Surg (London England)* (2014) 12(12):1495–9. doi: 10.1016/j.ijsu.2014.07.013

25. Burgess S, Thompson SG. Avoiding bias from weak instruments in mendelian randomization studies. *Int J Epidemiol* (2011) 40(3):755–64. doi: 10.1093/ije/dyr036

26. Chen TK, Knicely DH, Grams ME. Chronic kidney disease diagnosis and management: A review. *Jama* (2019) 322(13):1294–304. doi: 10.1001/jama.2019.14745

27. Zhang WR, Parikh CR. Biomarkers of acute and chronic kidney disease. *Annu Rev Physiol* (2019) 81:309–33. doi: 10.1146/annurev-physiol-020518-114605

28. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* (2013) 37 (7):658–65. doi: 10.1002/gepi.21758

29. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, et al. Recent developments in mendelian randomization studies. *Curr Epidemiol Rep* (2017) 4 (4):330–45. doi: 10.1007/s40471-017-0128-6

30. 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

31. Nowak C, Ärnlöv J. A mendelian randomization study of the effects of blood lipids on breast cancer risk. *Nat Commun* (2018) 9(1):3957. doi: 10.1038/s41467-018-06467-9

32. Patil CN, Racusen LC, Reckelhoff JF. Consequences of advanced aging on renal function in chronic hyperandrogenemic female rat model: Implications for aging women with polycystic ovary syndrome. *Physiol Rep* (2017) 5(20):e13461. doi: 10.14814/phy2.13461

33. Behboudi-Gandevani S, Amiri M, Cheraghi L, Amanollahi Soudmand S, Azizi F, Ramezani Tehrani F. The risk of chronic kidney disease among women with polycystic ovary syndrome: A long-term population-based cohort study. *Clin Endocrinol (Oxf)* (2020) 93(5):590–7. doi: 10.1111/cen.14284

34. Hoeger KM, Dokras A, Piltonen T. Update on pcos: Consequences, challenges, and guiding treatment. *J Clin Endocrinol Metab* (2021) 106(3):e1071–e83. doi: 10.1210/clinem/dgaa839

35. Song Y, Ye W, Ye H, Xie T, Shen W, Zhou L. Serum testosterone acts as a prognostic indicator in polycystic ovary syndrome-associated kidney injury. *Physiol Rep* (2019) 7(16):e14219. doi: 10.14814/phy2.14219

36. Butt JH, Adamson C, Docherty KF, Vaduganathan M, Solomon SD, Anand IS, et al. Eligibility for pharmacological therapies in heart failure with reduced ejection fraction: Implications of the new chronic kidney disease epidemiology collaboration creatinine equation for estimating glomerular filtration rate. *Eur J Heart Fail* (2022) 24 (5):861–6. doi: 10.1002/ejhf.2460

37. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, et al. New creatinine- and cystatin c-based equations to estimate gfr without race. *New Engl J Med* (2021) 385(19):1737–49. doi: 10.1056/NEJMoa2102953

38. Markwardt D, Holdt L, Steib C, Benesic A, Bendtsen F, Bernardi M, et al. Plasma cystatin c is a predictor of renal dysfunction, acute-on-Chronic liver failure, and mortality in patients with acutely decompensated liver cirrhosis. *Hepatology* (2017) 66 (4):1232–41. doi: 10.1002/hep.29290

39. Chen YJ, Hsu YP, Tain YL, Li NS, Pang HH, Kuo SW, et al. Microneedle patches integrated with lateral flow cassettes for blood-free chronic kidney disease point-of-Care testing during a pandemic. *Biosens Bioelectron* (2022) 208:114234. doi: 10.1016/j.bios.2022.114234

40. Özdemir Başer Ö, Göçmen AY, Aydoğan Kırmızı D. The role of inflammation, oxidation and cystatin-c in the pathophysiology of polycystic ovary syndrome. *Turk J Obstet Gynecol* (2022) 19(3):229–35. doi: 10.4274/tjod.galenos.2022.29498

41. Kuro OM. Klotho and endocrine fibroblast growth factors: Markers of chronic kidney disease progression and cardiovascular complications? *Nephrology dialysis Transplant Off Publ Eur Dialysis Transplant Assoc - Eur Renal Assoc* (2019) 34 (1):15–21. doi: 10.1093/ndt/gfy126

42. Courbon G, Francis C, Gerber C, Neuburg S, Wang X, Lynch E, et al. Lipocalin 2 stimulates bone fibroblast growth factor 23 production in chronic kidney disease. *Bone Res* (2021) 9(1):35. doi: 10.1038/s41413-021-00154-0

43. Fujiu K, Manabe I, Nagai R. Renal collecting duct epithelial cells regulate inflammation in tubulointerstitial damage in mice. *J Clin Invest* (2011) 121(9):3425–41. doi: 10.1172/jci57582

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 22 March 2023
DOI 10.3389/fendo.2023.1050049



OPEN ACCESS

EDITED BY
Haifeng Hou,
Shandong First Medical University, China

REVIEWED BY
Marijana Vujkovic,
University of Pennsylvania, United States
Xiuqing Guo,
Cedars Sinai Medical Center, United States

*CORRESPONDENCE
Kei Hang Katie Chan
kkhchan@cityu.edu.hk

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

RECEIVED 21 September 2022
ACCEPTED 09 March 2023
PUBLISHED 22 March 2023

CITATION
Tan Y, He Q and Chan KHK (2023) Identification of shared genetic architecture between non-alcoholic fatty liver disease and type 2 diabetes: A genome-wide analysis.
*Front. Endocrinol.* 14:1050049.
doi: 10.3389/fendo.2023.1050049



# Identification of shared genetic architecture between non-alcoholic fatty liver disease and type 2 diabetes: A genome-wide analysis

Yajing Tan[1], Qian He[1] and Kei Hang Katie Chan[1,2,3]*

[1]Department of Biomedical Sciences, City University of Hong Kong, Hong Kong, Hong Kong SAR, China, [2]Department of Electrical Engineering, City University of Hong Kong, Hong Kong, Hong Kong SAR, China, [3]Department of Epidemiology, Center for Global Cardiometabolic Health, Brown University, Providence, RI, United States

**Background:** The incidence of complications of non-alcoholic fatty liver disease (NAFLD) and type 2 diabetes (T2D) has been increasing.

**Method:** In order to identify the shared genetic architecture of the two disease phenotypes of NAFLD and T2D, a European population-based GWAS summary and a cross-trait meta-analysis was used to identify significant shared genes for NAFLD and T2D. The enrichment of shared genes was then determined through the use of functional enrichment analysis to investigate the relationship between genes and phenotypes. Additionally, differential gene expression analysis was performed, significant differentially expressed genes in NAFLD and T2D were identified, genes that overlapped between those that were differentially expressed and cross-trait results were reported, and enrichment analysis was performed on the core genes that had been obtained in this way. Finally, the application of a bidirectional Mendelian randomization (MR) approach determined the causal link between NAFLD and T2D.

**Result:** A total of 115 genes were discovered to be shared between NAFLD and T2D in the GWAS analysis. The enrichment analysis of these genes showed that some were involved in the processes such as the decomposition and metabolism of lipids, phospholipids, and glycerophospholipids. Additionally, through the use of differential gene expression analysis, 15 core genes were confirmed to be linked to both T2D and NAFLD. They were correlated with carcinoma cells and inflammation. Furthermore, the bidirectional MR identified a positive causal relationship between NAFLD and T2D.

**Conclusion:** Our study determined the genetic structure shared between NAFLD and T2D, offering a new reference for the genetic pathogenesis and mechanism of NAFLD and T2D comorbidities.

KEYWORDS

**non-alcoholic fatty liver disease, type 2 diabetes, GWAS, differential gene expression, shared genetics, mendelian randomization**

# 1 Introduction

Non-alcoholic fatty liver disease (NAFLD) is a clinicopathological syndrome that is characterized by hepatic parenchymal steatosis and fat storage in the absence of a history of binge drinking. In simple terms, NAFLD is usually benign, but if linked with an unhealthy lifestyle, obesity, and other metabolic syndromes, it may develop from simple fat accumulation to non-alcoholic steatohepatitis (NASH), liver fibrosis and cirrhosis, and, in rare cases, liver cancer (1). It is a complex disease that results from the interaction of environmental and genetic factors, so it has multiple pathogenic factors such as insulin resistance (IR), lipid metabolism disorders, oxidative stress, and cytokine effects (2). In addition to being considered a manifestation of IR in the liver, NAFLD often coexists with metabolic syndromes such as obesity, type 2 diabetes (T2D), and hyperlipidemia (3). Among them, T2D is a typical endocrine and metabolic disease that is affected by multiple pathogenic factors, and its incidence is increasing rapidly worldwide. T2D, the most prevalent type of diabetes, is characterized by hyperglycemia, IR, and lipid metabolism disorder as the pathological basis (4). In recent years, it has been found that NAFLD and T2D show similar pathological characteristics and often coexist as common diseases that seriously endanger public health. On one hand, T2D can lead to dysfunction of glycolipid metabolism in the body through the development of factors such as IR, chronic inflammation and oxidative stress, which results in NAFLD and further liver damage and worsens the prognosis of NAFLD (5)? On the other hand, through fat deposition, inflammation, endoplasmic reticulum stress, and oxidative stress, NAFLD can also exacerbate hepatic IR and promote metabolic abnormalities including hyperglycemia, creating the ideal environment for the development of T2D (6).

Related studies have shown that NAFLD and T2D interact with each other, and that there is a complex two-way relationship between the two that can accelerate deterioration. Targher et al. (7) found that NAFLD was ubiquitous in patients with T2D (7). Similarly, Jarvis et al. (8), in a meta-analysis of population-based cohort studies, found that the occurrence of T2D was associated with a more than two-fold increase in the risk of severe liver disease events among those at risk of or diagnosed with NAFLD (8). This finding was the same as that found by Mantovani et al. (9) in a study of the impact of NAFLD on the risk of development of T2D (9). Also, Pinero et al. (10) reported that the global incidence of NASH had reached 3% to 5%. NASH occurs in 20% to 30% of patients with T2D and obesity, and NAFLD occurs in 69% to 87% of those patients (10). Hence, the incidence of NAFLD combined with T2D is higher than that of NAFLD alone or T2D alone (5). Thus, the identification of the shared genetic architecture of NAFLD and T2D has important implications for the prevention and treatment of these diseases.

In recent years, the development of next-generation sequencing and high-throughput genotyping arrays has led to the GWAS and exome-wide association studies (EWAS), which are methods for the identification of genetic factors for many complex diseases (11). The use of GWAS, which is larger-scale than EWAS, has led to the identification of many polymorphisms and genetic variants that are associated with NAFLD and T2D and the investigation of new therapeutic targets. In addition, differentially expressed genes (DEGs) are key to learn about gene activity. They have now become one of the most important tools for the discovery of biomarkers (12). This method can be used to find genes that exhibit notable variations in expression, to analyze statistically the findings to pinpoint particular genes that are associated with those conditions, and then to analyze the biological importance of those particular genes. More importantly, DEGs can complement the knowledge of important target tissues and cell types that the GWAS approach lacks in disease pathogenesis, thus realizing the transformation of relevant gene loci into mechanisms. Hence it can be seen that the integration of GWAS summary statistics and gene expression data can identify disease-related tissues and cell types without bias, increase the credibility of the analysis results, and provide a sufficient basis to explain the pathogenesis.

Current genetic studies that target NAFLD and T2D require the discovery of more significant genetic association signals to support and translate the research through various novel analytical methods into biological and potentially therapeutic knowledge. Therefore, this study first conducted a comprehensive genetic analysis through the use of GWAS to identify susceptibility genes for NAFLD combined with T2D. The core shared genes were further screened as this information was combined with the results of DEG analysis. Subsequently, functional annotation analysis was performed to identify the underlying biological pathways of these core, shared genes. At the same time, a two-sample MR analysis was conducted to explore the causal relationship between NAFLD and T2D. The above analysis provided a robust theoretical basis for the study of NAFLD and T2D complications and new ideas and opportunities for the further development of prevention and treatment strategies.

# 2 Methods

## 2.1 Data summary

To identify genetic variants in NAFLD combined with T2D, the GWAS summary statistics for this study were obtained from the US National Human Genome Research Institute GWAS catalog (https://www.ebi.ac.uk/gwas/), including NAFLD (1,483 cases and 17,781 controls) and T2D (4,040 cases and 113,735 controls). (13, 14). For the DEG analysis, the organism was *Homo sapiens*, and the experiment type was expression profiling by array, which set the screening conditions of the dataset. The related gene expression datasets GSE48452, GSE25724, GSE17470, and GSE20966 were downloaded from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/), which included human liver biopsy (18 NASH, 14 NAF, 27 obese, 14 controls), human islets, and NASH liver biopsy (6 case and 4 controls) and beta-cells from pancreatic tissue (n=10) (15–18). All data sources can be found in Table 1.

TABLE 1 Data source information summary.

| Phenotype | Data Source | Population | Cases' size | Controls' size | Total' size | Number (SNPs) | PubMed ID | Download Link |
|---|---|---|---|---|---|---|---|---|
| NAFLD | GWAS Catalog | European | 1,483 | 17,781 | 19,264 | 6,797,908 | 32298765 | https://www.ebi.ac.uk/gwas/studies/GCST90011885 |
| T2D | GWAS Catalog | European | 4,040 | 113,735 | 117,775 | 8,404,432 | 26961502 | https://www.ebi.ac.uk/gwas/studies/GCST006801 |
| NAFLD (Discovery set) | GEO database (GSE48452, GPL11532) | Germany | 32 | 41 | 73 | – | 23931760 | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi |
| NAFLD (Validation set) | GEO database (GSE17470, GPL2895) | USA | 7 | 4 | 11 | – | 20221393 | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi |
| T2D (Discovery set) | GEO database (GSE25724, GPL96) | Italy | 6 | 7 | 13 | – | 21127054 | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi |
| T2D (Validation set) | GEO database (GSE20966, GPL1352) | USA | 10 | 10 | 20 | – | 20644627 | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi |

## 2.2 Study-level quality control

The "GWASInspector" R package was used to conduct harmonized quality control (QC) on the GWAS statistics of NAFLD and T2D phenotypes to ensure that false positive signals were eliminated and that low-quality data did not obscure actual signals (19, 20). NAFLD and T2D GWAS summary data were performed separately in the QC. The reference dataset was the 1000 genomes project reference panel (21), the specific genome build version was GRCh37, and we included the relevant information from the European population. The QC involved: the deletion of variants that contained missing fundamental values or were duplicated; the deletion of monomorphic variants; the checking of the consistency of allele frequencies with reference datasets; the alignment of alleles with reference datasets, and the comparison of those alleles to ensure that the resulting allele frequencies were correct; the removal of unverifiable mismatches and multi-allelic variants, and; the setting of a threshold plot_cut-off_$p$ = 0.01 to exclude low-significance SNPs.

## 2.3 Cross-trait meta-analysis

Cross-trait meta-analysis was performed using the CPASSOC software package. CPASSOC is a method for studying cross-phenotypic (CP) associations by using summary statistics from GWAS of multiple phenotypes. It combines effect estimates and standard errors of GWAS summary statistics to test the hypothesis of an association between SNPs and traits (22). Cross-phenotype associations increase statistical power when the traits analyzed share common variants or common genetic pathways, which are often associated with pleiotropy (23). CPASSOC includes two tests, SHom and SHet. In this study, R v.4.1.3 was used to perform the SHet test considering the effect of trait heterogeneity, which can increase the power when the genetic effect size of different traits is different (24). At this time, the gamma distribution parameters are estimated by setting $N = 1E4$ and calling the EstimateGamma function. Due to the many hypothesis tests that may be applied in GWAS studies, the threshold is strictly controlled to minimize the number of false positives reported. Currently, the most significant threshold is generally recognized as $p<5\times10^{-8}$, and this threshold also applies to CPASSOC (24, 25). Then, hg19 was used as the reference genome, and the refGene database was used to annotate the SNPs that reached the threshold of significance level using the ANNOVAR software (http://www.openbioinformatics.org/annovar/). Finally, the shared genes of NAFLD combined with T2D were obtained.

## 2.4 Enrichment analysis

In this study, SNPs that showed significant variation in meta-analysis and the genes from which they came were used for functional enrichment analysis to explore the potential biological function of shared susceptibility genes between NAFLD and T2D. The online tool Metaspace (https://metascape.org/gp/index.html#/main/step1) was used to analyze comprehensively these susceptibility genes. Metaspace integrates more than 40 gene function annotation databases such as Gene Ontology (GO) and DisGeNET, providing multiple functional and diversified visualization methods such as gene enrichment analysis and protein interaction network analysis, which can be used for easy exploration and analysis of gene function (26). The GO enrichment analysis of candidate genes was focused on the use of the "clusterProfiler" package of R v.4.1.3 (https://www.r-project.org/). In addition, the online platform TissueEnrich (https://tissueenrich.gdcb.iastate.edu/) was used as a calculating input-gene centralized organization-specific enrichment tool to complete the tissue-specific expression analysis (27).

## 2.5 Differential gene expression analysis and enrichment

To confirm which of the chosen genes were the core shared genes in NAFLD and T2D, four GEO datasets were selected for DEG analysis. The specific dataset information is shown in Table 1. Of the four, GSE48452 and GSE25724 were used as discovery sets, while GSE17470 and GSE20966 were used as validation sets to verify the validity and disease association of the identified genes. In addition, each dataset was divided into two groups of samples, with NAFLD patients or T2D patients as the experimental group and healthy people as the control group. GEO data were processed through the application of the online analysis tool GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) to identify DEGs. The visualization of overlapping genes in GEO datasets is realized through the online platform jvenn (http://www.bioinformatics.com.cn/static/others/jvenn/example.html) (28). After the discovery and validation sets were merged, the final DEGs were screened through use of adjust.P.Value, which is applied to adjust the *p*-value for multiple tests to control the false discovery rate (FDR). The FDR is calculated as expected rate x (false positive/(false positive + true positive)) (29). The value of *adj.p* was set at *<0.05* to screen out the DEGs of NAFLD and T2D. Then, the genes that overlapped with those found through the GWAS were identified as the shared core genes of NAFLD and T2D. Differential gene expression analysis is to identify shared genes through differential expression analysis between two phenotypes and to find overlapping genes with cross-trait analysis based on the GWAS summary statistics. Using multiple analytical methods to explore the reproducibility of our results between the two phenotypes can make the findings more reliable and robust. Next, the enrichment analysis described in section 2.4 was carried out for the genes that were found to overlap in the GWAS and DEG analyses.

## 2.6 Mendelian randomization analysis

The QC-processed GWAS data were used in the MR analysis. The potential causal effect between T2D and NAFLD was explored through the use of a bidirectional MR analysis, in which the two traits were evaluated alternately as exposure and outcome, and independent SNPs that were closely related to exposure and outcome traits were used as instrumental variables. Among them, the screening of exposure was essential. The parameters $p=5\times10^{-8}$ specified the *p*-value of the SNP in the exposure; that is, only SNPs with *p*-values of $<5\times10^{-8}$ were extracted (30). The NbDistribution simulation calculation was set to 1000 and the *p*-value threshold for judging whether the SNP was an outlier was set to 0.05 before the MR analysis was performed. Then, the calculation of MR pleiotropy residual sum and outlier (MR-PRESSO) was performed to identify the existence of the outliers (31). Once outliers were located, they were eliminated, and subsequent MR analysis was performed. MR and sensitivity analyses were performed through the use of the inverse variance weighted (IVW) method (32) with multiplicative random effects, supplemented by MR Egger (33, 34), weighted median (33), simple mode, and weighted mode methods (35). It is important to note that horizontal pleiotropy is a potential confounding factor in MR analysis; i.e., instrumental variable SNPs influence the outcome through a non-causal pathway, which may affect the measurement of the relationship between traits. To examine the impact of pleiotropy on the results of the MR analysis, MR-PRESSO was also used to test for horizontal pleiotropy for multiple instrumental variables. In addition, heterogeneity statistics and leave-one-out analyses were included in the MR analysis. Heterogeneity statistics mainly test the differences between individual SNPs, and leave-one-out analysis mainly tests the stability of MR results. The "TwoSampleMR" and "MRPRESSO" packages were used for MR analysis in R v.4.1.3.

# 3 Results

## 3.1 Study-level QC

QC was performed through the use of "GWASInspector". 100% of NAFLD GWAS summary data (6,797,908 SNPs) and 99.7% of T2D GWAS summary data (8,380,746 SNPs) passed the QC procedure (Table 2). SNPs that passed the QC were included in the subsequent cross-trait meta-analysis and MR analysis.

## 3.2 Cross-trait meta-analysis

In total, CPASSOC identified 241 SNPs that were significantly associated ($p<5\times10^{-8}$) between NAFLD and T2D (Supplementary Table 1). In the results of the cross-trait meta-analysis, the SNP with the most significant p-value is rs73233361 ($p=6.78\times10^{-11}$), which is located on chromosome 12. Most of the remaining SNPs are located on chromosome 6, chromosome 2, and chromosome 3. 115 genes were obtained by ANNOVAR annotation (Supplementary Table 1).

## 3.3 Enrichment analysis

The DisGeNET enrichment analysis revealed that 115 shared genes were enriched in physical activity measurement, substance-related disorders, lean body mass, smoking behaviors, substance abuse problem, etc. (Figure 1A). The relevant results that were identified based on DisGeNET enrichment analysis are listed in Supplementary Table 2. GO enrichment analysis (Supplementary Table 3; Supplementary Figure 1) showed that the shared genes of NAFLD and T2D were enriched in the biological processes of

TABLE 2 The number of SNPs after QC processing.

| | NAFLD | T2D |
|---|---|---|
| Input variant count | 6,797,908 | 8,404,432 |
| Missing crucial variable | 0 | 2 |
| Duplicated variants | 0 | 12,367 |
| Monomorphic variants | 0 | 0 |
| Output variant count | 6,797,908 | 8,380,746 |

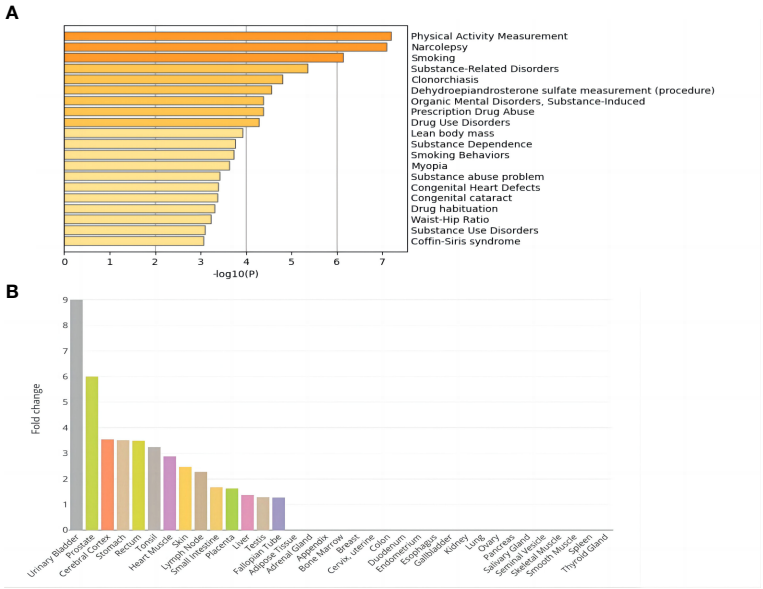



FIGURE 1
Enrichment analysis of shared genes in NAFLD and T2D. **(A)** DisGeNET enrichment analysis results. **(B)** Tissue enrichment results.

processes of glycerophospholipids, phospholipids, lipid decomposition, glycerolipid metabolism, and they also participated in the enrichment in the molecular function of activity of various enzymes. Among them, the most were associated with sensory system development, a total of 7 genes (*FASLG, ISL1, TULP1, TBC1D32, ATP8A2, MAX*, and *ADAMTS18*). In addition, tissue enrichment analysis showed that NAFLD and T2D shared genes were mainly enriched in 14 tissues: the urinary bladder, prostate, cerebral cortex, stomach, rectum, tonsil, heart muscle, skin, lymph node, small intestine, placenta, liver, testis, and fallopian tube (Figure 1B). Among these tissues, the liver is closely related to the pathogenesis of NAFLD and T2D.

## 3.4 Differential gene expression analysis and enrichment

The DEG analysis results for each dataset are shown in Supplementary Figures 2–5. The discovery sets GSE48452 and GSE25724 contained 11,795 shared genes for NAFLD and T2D (Supplementary Figures 6A, B), whereas the validation sets GSE17470 and GSE20966 contained 15,557 shared genes for NAFLD and T2D (Supplementary Figures 6C, D). A combination of all discovery and validation sets yielded 9711 DEGs (Supplementary Figure 6E). Subsequently, 5545 DEGs shared by NAFLD and T2D were screened by adj.P (Supplementary Table 4). Consideration of these genes with the candidate genes that had been obtained through the GWAS produced fifteen core genes that were shared by NAFLD and T2D, namely *DNAJB9, VPS53, SCGN, CMAS, RGS6, FASLG, ABHD10, ATRN, PLA2G2F, ITIH2, ROBO1, SGCG, SH3GL2, CNR1*, and *FOXN3* (Table 3).

DEG analysis of the above core genes (Supplementary Table 5) showed that seven genes were upregulated (*logFoldChange>0*) and eight genes were downregulated (*logFoldChange<0*) in disease. These fifteen core genes were subjected to enrichment analysis, and DisGeNET enrichment analysis revealed that they were enriched in carcinoma cells and inflammation (Figure 2A). Relevant findings from the DisGeNET enrichment analysis are provided in Supplementary Table 6.

GO enrichment analysis showed significant enrichment of several biological processes including regulation of endopeptidase and peptidase activity, lipid catabolic process, fatty acid metabolic process, response to lipopolysaccharide, and positive regulation of proteolysis; cellular components of the distal axon, endoplasmic reticulum lumen, and glutamatergic synapse; and molecular functions such as carboxylic ester hydrolase activity (Supplementary Table 7; Supplementary Figure 7). In addition, tissue enrichment analysis showed that NAFLD and T2D core shared genes were enriched in the urinary bladder, stomach, rectum, tonsil, heart muscle, lymph node, skeletal muscle, liver, skin, cerebral cortex, and testis (Figure 2B). The above enrichment results supported the earlier finding that the fifteen core shared genes, *DNAJB9, VPS53, SCGN, CMAS, RGS6, FASLG, ABHD10, ATRN, PLA2G2F, ITIH2, ROBO1, SGCG, SH3GL2, CNR1*, and *FOXN3*, were closely related to NAFLD and T2D.

## 3.5 Mendelian randomization analysis

No outliers were detected after processing with the "MRPRESSO" R package. The results of MR analysis in T2D and NAFLD are listed in Table 4. Among the results, regardless of

TABLE 3 Core genes after GWAS analysis and differential gene expression analysis combined.

| | Gene | Cytogenetic region | Description | Remark |
|---|---|---|---|---|
| 1 | *VPS53* | 17p13.3 | VPS53 subunit of GARP complex | Not novel gene |
| 2 | *SCGN* | 6p22.2 | secretagogin, EF-hand calcium binding protein | Not novel gene |
| 3 | *RGS6* | 14q24.2 | regulator of G protein signaling 6 | Not novel gene |
| 4 | *SGCG* | 13q12.12 | sarcoglycan gamma | Not novel gene |
| 5 | *FOXN3* | 14q32.11 | forkhead box N3 | Not novel gene |
| 6 | *DNAJB9* | 7q31.1 | DnaJ heat shock protein family (Hsp40) member B9 | Novel gene |
| 7 | *CMAS* | 12p12.1 | cytidine monophosphate N-acetylneuraminic acid synthetase | Novel gene |
| 8 | *FASLG* | 1q24.3 | Fas ligand | Novel gene |
| 9 | *ABHD10* | 3q13.2 | abhydrolase domain containing 10, depalmitoylase | Novel gene |
| 10 | *ATRN* | 20p13 | attractin | Novel gene |
| 11 | *PLA2G2F* | 1p36.12 | phospholipase A2 group IIF | Novel gene |
| 12 | *ITIH2* | 10p14 | inter-alpha-trypsin inhibitor heavy chain 2 | Novel gene |
| 13 | *ROBO1* | 3p12.3 | roundabout guidance receptor 1 | Novel gene |
| 14 | *SH3GL2* | 9p22.2 | SH3 domain containing GRB2 like 2, endophilin A1 | Novel gene |
| 15 | *CNR1* | 6q15 | cannabinoid receptor 1 | Novel gene |

whether NAFLD or T2D was used as exposure or outcome, the *p*-value obtained by the IVW method was less than 0.05, indicating a causal relationship between T2D and NAFLD; the related *beta*-value was more than zero, which indicated that the causal relationship between T2D and NAFLD was positive; this meant that increasing exposure (T2D) increased the risk of the outcome (NAFLD). From the scatter plot of the MR results (Figure 3), it can be seen that the IVW method yielded the most significant results among the five methods that were used for MR analysis. The plot also demonstrates the positive relationship between T2D and NAFLD, as did the forest plot (Supplementary Figure 8).

The statistical results of heterogeneity show that there was no heterogeneity between the instrumental variable SNPs (Q_pval was >0.05), which can be confirmed from the funnel plot (Supplementary Figure 9). The results of the pleiotropy test show that there was no statistical difference ($p$>0.05), which indicates that there was no horizontal pleiotropic effect. Through leave-one-out analysis (Supplementary Table 8; Supplementary Figure 10), it can

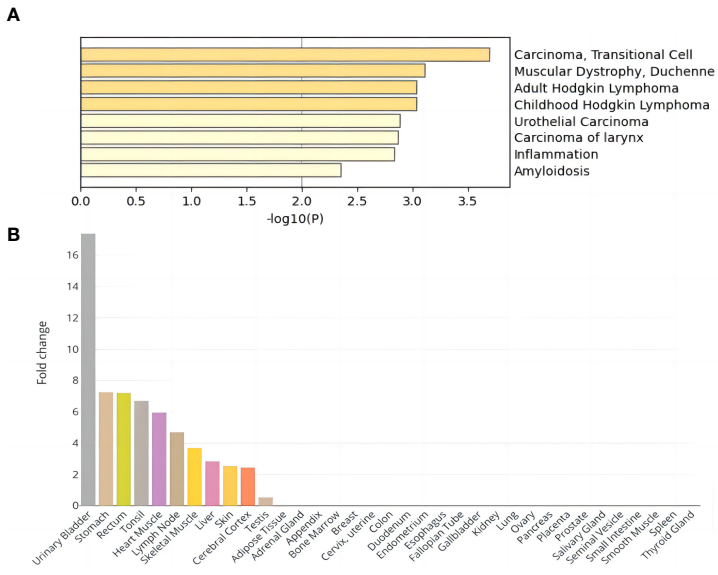



FIGURE 2
Enrichment analysis results of DEGs shared by NAFLD and T2D. **(A)** DisGeNET enrichment analysis results. **(B)** Tissue enrichment results.

TABLE 4 Results of two-sample MR analysis of NAFLD and T2D.

| Outcome | Exposure | Method | nsnp | b(exposure/outcome) | se | pval | OR (95%CI) |
|---|---|---|---|---|---|---|---|
| NAFLD | T2D | MR Egger | 62 | 0.01109 | 0.00652 | 0.09406 | 1.01115 (0.99832-1.02415) |
| | | Weighted median | 62 | 0.00278 | 0.00184 | 0.13017 | 1.00278 (0.99918-1.00639) |
| | | Inverse variance weighted | 62 | 0.00343 | 0.00129 | 0.00801 | 1.00344 (1.00090-1.00599) |
| | | Simple mode | 62 | -0.00016 | 0.00461 | 0.97296 | 0.99984 (0.99084-1.00892) |
| | | Weighted mode | 62 | -0.00016 | 0.00359 | 0.96526 | 0.99984 (0.99283-1.00690) |
| T2D | NAFLD | MR Egger | 217 | 0.00344 | 0.00573 | 0.54896 | 1.00344 (0.99224 -1.01477) |
| | | Weighted median | 217 | 0.01771 | 0.00389 | 5.36E-06 | 1.01787 (1.01013-1.02566) |
| | | Inverse variance weighted | 217 | 0.02414 | 0.00302 | 1.24E-15 | 1.02443 (1.01839-1.03051) |
| | | Simple mode | 217 | 0.02673 | 0.01887 | 0.15799 | 1.02709 (0.98980-1.06578) |
| | | Weighted mode | 217 | 0.02673 | 0.01909 | 0.16297 | 1.02709 (0.98936 -1.06625) |

be seen that no matter which SNP was removed, it would not have a fundamental impact on the results. So the MR results are robust.

# 4 Discussion

This study used GWAS summary data for 6,797,908 NAFLD and 8,404,432 T2D from European populations to determine the shared genetic architecture of these two phenotypes. A cross-trait meta-analysis identified 115 shared genes, and subsequent DEG analysis identified fifteen core shared genes: *DNAJB9, VPS53, SCGN, CMAS, RGS6, FASLG, ABHD10, ATRN, PLA2G2F, ITIH2, ROBO1, SGCG, SH3GL2, CNR1*, and *FOXN3*.

The liver is a vital organ that regulates glucose and lipid metabolism, and hepatic fat deposition is a critical factor in the pathogenesis of NAFLD and T2D (36). The twin-cycle hypothesis based on T2D explains that a gradual increase in the level of fat in the liver can lead to IR, which weakens the ability of insulin to suppress hepatic glucose production. This leads to an aggravation of hepatic gluconeogenesis and rises in blood sugar levels (37). The excess glucose is used to synthesize triglycerides, which results in increased levels of liver fat and reduced capacity to use glucose. These processes create a vicious circle between the liver and pancreas (38). At the same time, hepatic triglyceride synthesis is increased in NAFLD patients. When the level of free fatty acids (FFAs) produced by lipoprotein lipase exceeds the lipid storage capacity of adipose tissue, β-cells will take up many fatty acids and store them as triglycerides. This damages the β cells and causes IR (39). This may eventually promote the progression of liver damage to HCC.

Relevant studies to date have shown that the mechanism of action of the mechanisms mentioned above has become a tool for the conduct of research in clinical practice. Previous studies have suggested some potential links between these mechanisms and the

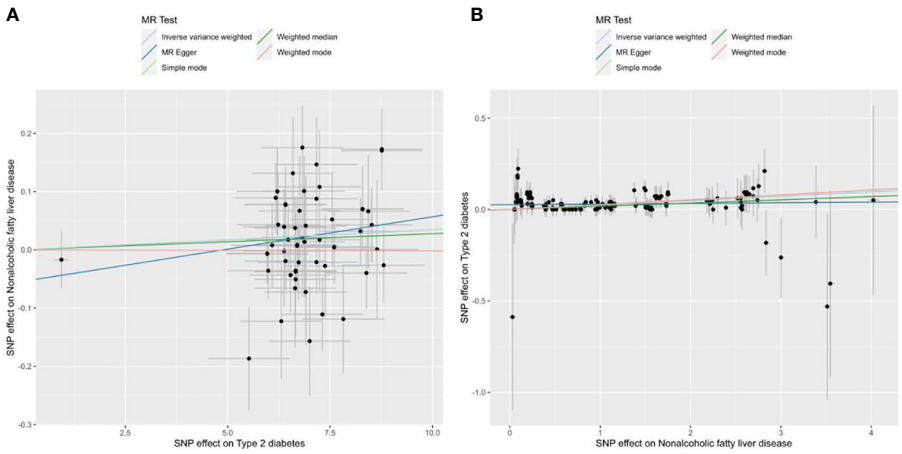



FIGURE 3
Scatter plots of the MR analysis. The light blue line shows the result of the IVW method, which has the most significant impact. The dark blue line shows the result of the MR Egger method; the light green is the result of the simple model method; the dark green is the result of the weighted median mode; and the red line represents the result of the weighted mode method. **(A)** Scatter plot of T2D as exposure and NAFLD as outcome. **(B)** Scatter plot of NAFLD as exposure and T2D as outcome.

identified, core shared genes. Forkhead box N3 (*FOXN3*), an important member of the FOX transcription factor family, is an important tumor suppressor gene that plays a crucial role in several cancers such as liver cancer, lung cancer, and colon cancer (40). The *FOXN3* gene locus is associated with fasting blood glucose levels. Hepatic *FOXN3* increases fasting blood glucose by inhibiting hepatic glucose utilization while also regulating the expression of amino acid transporters and catabolic enzymes (41, 42). Studies have shown that *FOXN3* suppresses the mRNA and protein expression of *E2F5* by inhibiting the promoter activity of potential oncogene *E2F5*, thereby inhibiting the proliferation of HCC cells *in vitro* and *in vivo* (43). Another tumor suppressor, regulator of G protein signaling 6 (*RGS6*), is upregulated in the liver of NAFLD patients, forms a complex with ATM in the liver, promotes ATM phosphorylation, and drives hepatic steatosis (44, 45). A study confirmed that hepatic *RGS6* increases oxidative stress and inflammation, which drive lipid deposition, fibrosis, and nonalcoholic fatty liver disease (46). In contrast, *RGS6* deficiency effectively ameliorated fat deposition, attenuated alcohol-dependent liver injury, and enhanced liver regeneration (47).

Other genes that may play a role in NAFLD and T2D include *SGCG*, a single-pass transmembrane glycoprotein implicated in the pathogenesis of obesity and T2D in humans (48). It has beneficial effects on glucose homeostasis, and elevated levels in diabetic patients may be compensatory for IR (49). Furthermore, *SCGN* is highly enriched in pancreatic β-cells and has pronounced effects on lipolysis and lipogenesis (50). It also regulates insulin expression and secretion, which is downregulated in type 2 diabetes (51, 52). Studies have shown that the *SCGN*-insulin interaction can stabilize insulin, enhance the hypoglycemic activity of insulin *in vivo*, and reduce hepatic steatosis and cholesterol metabolism disorders (51). In addition, the homologous gene *HCCS1* of *VPS53* also has a strong anti-tumor effect on liver cancer cells (53, 54).

Through a combination of the results of this study with the known mechanisms of action of NAFLD and T2D and related research findings, it can be shown that essential pathways affecting NAFLD and T2D include catabolism of lipids such as fatty acids, glycerides, and phospholipids. These biological processes affect lipid levels in tissues and hence affect hepatic fat accumulation and IR. Further research on this aspect of our findings should be considered.

In conclusion, the determination of triglyceride, FFA, and cholesterol levels can assist in the clinical observation of the dynamic changes in liver fat levels and IR and is of great significance in the prediction of comorbidities. At the same time, through the continuous deepening of genetic research, the development of targeted drugs to regulate the level of liver fat and the regulation of liver fat content is expected to become a key and effective treatment method for comorbidities. In addition, once the relevant mechanism of action is identified, specific gene therapy for NAFLD and T2D is expected to be realized.

One limitation of this study was that the shared genes were all screened from the results of GWAS studies in European populations, so other populations were not considered. Few replicated validation studies of the susceptibility loci associated with NAFLD and T2D have been conducted in other populations. Genetic and environmental factors influence the genetic backgrounds of populations and result in variations in allele frequencies, which affect illness incidence rates and the findings of GWAS analyses of susceptibility genes. Therefore it is uncertain whether the susceptibility genes identified in this study exist in other populations. However, the results of this study can provide a reference for research on NAFLD combined with T2D in other populations. GWAS research involves not only different populations but also different genders and different ages, and this richness of the data should be exploited for further exploration.

It is essential to note that this study cannot avoid the shortcomings of GWAS itself, such as the fact that the study is focused on the loci that achieve the significance threshold for genome-wide association, even though these loci account only partially for the complicated heredity of the disease (55). GWAS studies often overlook signals of mild or moderate association and ignore the effects of other variants such as gene deletions, copy number variations, etc. These neglected factors may involve underlying biological mechanisms that ultimately lead to the occurrence of disease. While NAFLD and T2D are complex diseases in which genetic and environmental factors interact, the pathogenesis is often caused by mutations or abnormalities of multiple genes, and each gene may play a part in a specific pathway but its role cannot explain the whole mechanism. Therefore, the study design can be effectively improved to make up for these issues with GWAS and the complexity of the disease. For example, the candidate gene method is used to find low-frequency variants, or the data from multiple studies can be combined in a meta-analysis to increase the sample size, and rare variants with substantial genetic effects can be found in this way (56, 57).

The strength of this study was that it involved the first comprehensive use of GWAS and DEG analysis to identify shared genes for NAFLD and T2D. During gene screening, strict thresholds were used to ensure the accuracy of the results, and significant shared genes were discovered efficiently. The study reconfirmed the association of the unveiled core genes, *VPS53, SCGN, RGS6, SGCG*, and *FOXN3*, with NAFLD and T2D, which had been reported in previous studies. The core genes *DNAJB9, CMAS, FASLG, ABHD10, ATRN, PLA2G2F, ITIH2, ROBO1, SH3GL2*, and *CNR1* were found to be related to NAFLD and T2D for the first time, and this provides a new research target for the precise treatment of NAFLD and T2D comorbidities.

## 5 Conclusion

In summary, this study found a causal relationship between NAFLD and T2D, which will be beneficial for the elucidation of the pathogenesis of NAFLD and T2D comorbidities. Fifteen core genes, *DNAJB9, VPS53, SCGN, CMAS, RGS6, FASLG, ABHD10, ATRN, PLA2G2F, ITIH2, ROBO1, SGCG, SH3GL2, CNR1, and FOXN3*, were identified as shared between NAFLD and T2D. This finding provided new ideas for the genetic study of NAFLD combined with T2D. Further gene expression verification and functional mechanism research should be carried out on these candidate

genes in the future to explore the specific biological mechanisms of NAFLD and T2D comorbidities and to provide new drug-targeting sites for the prevention and treatment of comorbidities.

## 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.

## Author contributions

YT participated in data analysis, drafting, writing, and revising the paper. QH participated in the revision of the manuscript and carried out a strict review of the manuscript. KC conceived, designed, coordinated the study, and revised the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the City University of Hong Kong new research initiatives/infrastructure support from central funding (APRC). (grant no. 9610401)

## Acknowledgments

In the study, GWAS analysis was performed using the GWAS Catalog database resource of the National Human Genome Research Institute. Furthermore, the differential gene expression analysis used the GEO database resource of the National Center for Biotechnology Information. We thank Anstee QM, Wood AR, Jochen H, Baker SS, Bugliani M, Marselli L, and other researchers for their GWAS summary statistics and GEO data. We also thank all participants who provided study samples. We thank the strong support of the Department of Biomedical Sciences, City University of Hong Kong.

## 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.2023.1050049/full#supplementary-material

## References

1. Hardwick RN, Fisher CD, Canet MJ, Lake AD, Cherrington NJ. Diversity in antioxidant response enzymes in progressive stages of human nonalcoholic fatty liver disease. *Drug Metab Dispos* (2010) 38:2293–301. doi: 10.1124/dmd.110.035006

2. Bellan M, Colletta C, Barbaglia MN, Salmi L, Clerici R, Mallela VR, et al. Severity of nonalcoholic fatty liver disease in type 2 diabetes mellitus: Relationship between nongenetic factors and PNPLA3/HSD17B13 polymorphisms. *Diabetes Metab J* (2019) 43:700–10. doi: 10.4093/dmj.2018.0201

3. Marchesini G, Brizi M, Bianchi G, Tomassetti S, Bugianesi E, Lenzi M, et al. Nonalcoholic fatty liver disease: A feature of the metabolic syndrome. *Diabetes* (2001) 50:1844–50. doi: 10.2337/diabetes.50.8.1844

4. Pan Q, Lin S, Li Y, Liu L, Li X, Gao X, et al. A novel GLP-1 and FGF21 dual agonist has therapeutic potential for diabetes and non-alcoholic steatohepatitis. *EBioMedicine* (2021) 63:103202. doi: 10.1016/j.ebiom.2020.103202

5. Shimizu M, Suzuki K, Kato K, Jojima T, Iijima T, Murohisa T, et al. Evaluation of the effects of dapagliflozin, a sodium-glucose co-transporter-2 inhibitor, on hepatic steatosis and fibrosis using transient elastography in patients with type 2 diabetes and non-alcoholic fatty liver disease. *Diabetes Obes Metab* (2019) 21:285–92. doi: 10.1111/dom.13520

6. Mu W, Cheng XF, Liu Y, Lv QZ, Liu GL, Zhang JG, et al. Potential nexus of non-alcoholic fatty liver disease and type 2 diabetes mellitus: Insulin resistance between hepatic and peripheral tissues. *Front Pharmacol* (2018) 9:1566. doi: 10.3389/fphar.2018.01566

7. Targher G, Bertolini L, Padovani R, Rodella S, Tessari R, Zenari L, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes Care* (2007) 30:1212–8. doi: 10.2337/dc06-2247

8. Jarvis H, Craig D, Barker R, Spiers G, Stow D, Anstee QM, et al. Metabolic risk factors and incident advanced liver disease in non-alcoholic fatty liver disease (NAFLD): A systematic review and meta-analysis of population-based observational studies. *PloS Med* (2020) 17:e1003100. doi: 10.1371/journal.pmed.1003100

9. Mantovani A, Byrne CD, Bonora E, Targher G. Nonalcoholic fatty liver disease and risk of incident type 2 diabetes: A meta-analysis. *Diabetes Care* (2018) 41:372–82. doi: 10.2337/dc17-1902

10. Pinero F, Pages J, Marciano S, Fernandez N, Silva J, Anders M, et al. Fatty liver disease, an emerging etiology of hepatocellular carcinoma in Argentina. *World J Hepatol* (2018) 10:41–50. doi: 10.4254/wjh.v10.i1.41

11. Du X, Deforest N, Majithia AR. Human genetics to identify therapeutic targets for NAFLD: Challenges and opportunities. *Front Endocrinol (Lausanne)* (2021) 12:777075. doi: 10.3389/fendo.2021.777075

12. Haynes WA, Higdon R, Stanberry L, Collins D, Kolker E. Differential expression analysis for pathways. *PloS Comput Biol* (2013) 9:e1002967. doi: 10.1371/journal.pcbi.1002967

13. Wood AR, Tyrrell J, Beaumont R, Jones SE, Tuke MA, Ruth KS, et al. Variants in the FTO and CDKAL1 loci have recessive effects on risk of obesity and type 2 diabetes, respectively. *Diabetologia* (2016) 59:1214–21. doi: 10.1007/s00125-016-3908-5

14. Anstee QM, Darlay R, Cockell S, Meroni M, Govaere O, Tiniakos D, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically characterised cohort(☆). *J Hepatol* (2020) 73:505–15. doi: 10.1016/j.jhep.2020.04.003

15. Baker SS, Baker RD, Liu W, Nowak NJ, Zhu L. Role of alcohol metabolism in non-alcoholic steatohepatitis. *PloS One* (2010) 5:e9570. doi: 10.1371/journal.pone.0009570

16. 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:e11499. doi: 10.1371/journal.pone.0011499

17. 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) 286:4216–25. doi: 10.1074/jbc.M110.200295

18. Ahrens M, Ammerpohl O, Von Schönfels W, Kolarova J, Bens S, Itzel T, et al. DNA Methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab* (2013) 18:296–302. doi: 10.1016/j.cmet.2013.07.004

19. Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, et al. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* (2010) 34:591–602. doi: 10.1002/gepi.20516

20. Ani A, van der Most PJ, Snieder H, Vaez A, Nolte IM. GWASinspector: Comprehensive quality control of genome-wide association study results. *Bioinformatics* (2021) 37:129–30. doi: 10.1093/bioinformatics/btaa1084

21. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature* (2015) 526:68–74. doi: 10.1038/nature15393

22. 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

23. Park H, Li X, Song YE, He KY, Zhu X. Multivariate analysis of anthropometric traits using summary statistics of genome-wide association studies from GIANT consortium. *PloS One* (2016) 11:e0163912. doi: 10.1371/journal.pone.0163912

24. Li X, Zhu X. Cross-phenotype association analysis using summary statistics from GWAS. *Methods Mol Biol* (2017) 1666:455–67. doi: 10.1007/978-1-4939-7274-6_22

25. Trepo E, Valenti L. Update on NAFLD genetics: From new variants to the clinic. *J Hepatol* (2020) 72:1196–209. doi: 10.1016/j.jhep.2020.02.020

26. Alexandrov T, Ovchinnikova K, Palmer A, Kovalev V, Tarasov A, Stuart L, et al. METASPACE: A community-populated knowledge base of spatial metabolomes in health and disease. *BioRxiv* (2019) 539478. doi: 10.1101/539478

27. Jain A, Tuteja G. TissueEnrich: Tissue-specific gene enrichment analysis. *Bioinformatics* (2019) 35:1966–7. doi: 10.1093/bioinformatics/bty890

28. Bardou P, Mariette J, Escudié F, Djemiel C, Klopp C. Jvenn: an interactive Venn diagram viewer. *BMC Bioinf* (2014) 15:1–7. doi: 10.1186/1471-2105-15-293

29. Jafari M, Ansari-Pour N. Why, when and how to adjust your p values? *Cell J (Yakhteh)* (2019) 20:604–7. doi: 10.22074/cellj.2019.5992

30. Xiang K, Hu Y-Q, HE Y-S, Wang P, Tao S-S, Chen Y, et al. Non-causal effect of circulating vitamin d levels on the risk of rheumatoid arthritis: A two-sample mendelian randomization study. (2021). doi: 10.21203/rs.3.rs-479402/v1

31. 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

32. Mounier N, Kutalik Z. Bias correction for inverse variance weighting mendelian randomization. *bioRxiv* (2021). doi: 10.1101/2021.03.26.437168

33. Bowden J, Smith G,D, 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

34. Burgess S, Thompson SG. Interpreting findings from mendelian randomization using the MR-egger method. *Eur J Epidemiol* (2017) 32:377–89. doi: 10.1007/s10654-017-0255-x

35. Hartwig FP, Davey SG, Bowden J. Robust inference in summary data mendelian randomization *via* the zero modal pleiotropy assumption. *Int J Epidemiol* (2017) 46:1985–98. doi: 10.1093/ije/dyx102

36. Dongiovanni P, Stender S, Pietrelli A, Mancina RM, Cespiati A, Petta S, et al. Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. *J Intern Med* (2018) 283:356–70. doi: 10.1111/joim.12719

37. Taylor R. Type 2 diabetes and remission: practical management guided by pathophysiology. *J Internal Med* (2021) 289:754–70. doi: 10.1111/joim.13214

38. Taylor R. Putting insulin resistance into context by dietary reversal of type 2 diabetes. *J R Coll Physicians Edinburgh* (2017) 47:168–71. doi: 10.4997/JRCPE.2017.216

39. Birkenfeld AL, Shulman GI. Nonalcoholic fatty liver disease, hepatic insulin resistance, and type 2 diabetes. *Hepatology* (2014) 59:713–23. doi: 10.1002/hep.26672

40. Lin Z, Chen M, Wan Y, Lei L, Ruan H. miR-574-5p targets FOXN3 to regulate the invasion of nasopharyngeal carcinoma cells *via* wnt/beta-catenin pathway. *Technol Cancer Res Treat* (2020) 19:1533033820971659. doi: 10.1177/1533033820971659

41. Erickson ML, Karanth S, Ravussin E, Schlegel A. FOXN3 hyperglycemic risk allele and insulin sensitivity in humans. *BMJ Open Diabetes Res Care* (2019) 7:e000688. doi: 10.1136/bmjdrc-2019-000688

42. Karanth S, Chaurasia B, Bowman FM, Tippetts TS, Holland WL, Summers SA, et al. FOXN3 controls liver glucose metabolism by regulating gluconeogenic substrate selection. *Physiol Rep* (2019) 7:e14238. doi: 10.14814/phy2.14238

43. Sun J, Li H, Huo Q, Cui M, Ge C, Zhao F, et al. The transcription factor FOXN3 inhibits cell proliferation by downregulating E2F5 expression in hepatocellular carcinoma cells. *Oncotarget* (2016) 7:43534. doi: 10.18632/oncotarget.9780

44. Mahata T, Sengar AS, Basak M, Das K, Pramanick A, Verma SK, et al. Hepatic regulator of G protein signaling 6 (RGS6) drives non-alcoholic fatty liver disease by promoting oxidative stress and ATM-dependent cell death. *Redox Biol* (2021) 46:102105. doi: 10.1016/j.redox.2021.102105

45. Wang Z, Chen J, Wang S, Sun Z, Lei Z, Zhang HT, et al. RGS6 suppresses TGF-beta-induced epithelial-mesenchymal transition in non-small cell lung cancers *via* a novel mechanism dependent on its interaction with SMAD4. *Cell Death Dis* (2022) 13:656. doi: 10.1038/s41419-022-05093-0

46. Dao W, Xiao Z, Yang W, Luo X, Xia H, Lu Z. RGS6 drives spinal cord injury by inhibiting AMPK pathway in mice. *Dis Markers* (2022) 2022:4535652. doi: 10.1155/2022/4535652

47. Stewart A, Maity B, Anderegg SP, Allamargot C, Yang J, Fisher RA. Regulator of G protein signaling 6 is a critical mediator of both reward-related behavioral and pathological responses to alcohol. *Proc Natl Acad Sci U.S.A.* (2015) 112:E786–95. doi: 10.1073/pnas.1418795112

48. Israeli D, Cosette J, Corre G, Amor F, Poupiot J, Stockholm D, et al. An AAV-SGCG dose-response study in a gamma-sarcoglycanopathy mouse model in the context of mechanical stress. *Mol Ther Methods Clin Dev* (2019) 13:494–502. doi: 10.1016/j.omtm.2019.04.007

49. Kuhn T, Kaiser K, Lebek S, Altenhofen D, Knebel B, Herwig R, et al. Comparative genomic analyses of multiple backcross mouse populations suggest SGCG as a novel potential obesity-modifier gene. *Hum Mol Genet* (2022) 31:4019–33. doi: 10.1093/hmg/ddac150

50. Chin YX, Mi Y, Cao WX, LiM PE, Xue CH, Tang QJ. A pilot study on anti-obesity mechanisms of kappaphycus alvarezii: The role of native kappa-carrageenan and the leftover sans-carrageenan fraction. *Nutrients* (2019) 11. doi: 10.3390/nu11051133

51. Sharma AK, Khandelwal R, Chadalawada S, Ram NS, Raj TA, Kumar MJM, et al. SCGN administration prevents insulin resistance and diabetic complications in high-fat diet fed animals. *bioRxiv* (2017), 189324. doi: 10.1101/189324

52. Liu Z, Tan S, Zhou L, Chen L, Liu M, Wang W, et al. SCGN deficiency is a risk factor for autism spectrum disorder. *Signal Transduction Targeted Ther* (2023) 8:3. doi: 10.1038/s41392-022-01225-2

53. Perez-Victoria FJ, Schindler C, Magadan JG, Mardones GA, Delevoye C, Romao M, et al. Ang2/fat-free is a conserved subunit of the golgi-associated retrograde protein complex. *Mol Biol Cell* (2010) 21:3386–95. doi: 10.1091/mbc.e10-05-0392

54. Peng H, Zheng J, Su Q, Feng X, Peng M, Gong L, et al. VPS53 suppresses malignant properties in colorectal cancer by inducing the autophagy signaling pathway. *Onco Targets Ther* (2020) 13:10667–75. doi: 10.2147/OTT.S254823

55. Hu C, Jia W. Diabetes in China: epidemiology and genetic risk factors and their clinical utility in personalized medication. *Diabetes* (2018) 67:3–11. doi: 10.2337/dbi17-0013

56. Quan C, Zhang X-J. Research strategies for the next step of genome-wide association study. *Yi Chuan= Hereditas* (2011) 33:100–8. doi: 10.3724/SP.J.1005.2011.00100

57. Tomlinson IP, Carvajal-Carmona LG, Dobbins SE, Tenesa A, Jones AM, Howarth K, et al. Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer. *PloS Genet* (2011) 7:e1002105. doi: 10.1371/journal.pgen.1002105

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 24 March 2023
DOI 10.3389/fendo.2023.1153205



OPEN ACCESS

EDITED BY
Xiao Wang,
Lund University, Sweden

REVIEWED BY
Mei Zhang,
Jining Medical University, China
Zhiyuan Wu,
Capital Medical University, China

*CORRESPONDENCE
Ming Liu
mingliu@tmu.edu.cn
Qingbo Guan
qingboguan@126.com
Lixing Ma
malixingsdfmu@163.com

†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 29 January 2023
ACCEPTED 06 March 2023
PUBLISHED 24 March 2023

CITATION
Yang S, Li H, Gu Y, Wang Q, Dong L, Xu C, Fan Y, Liu M, Guan Q and Ma L (2023) The association between total bile acid and bone mineral density among patients with type 2 diabetes.
*Front. Endocrinol.* 14:1153205.
doi: 10.3389/fendo.2023.1153205



# The association between total bile acid and bone mineral density among patients with type 2 diabetes

Song Yang[1,2†], Hongyun Li[3†], Yuanyuan Gu[4†], Qiang Wang[5], Li Dong[6], Chao Xu[7], Yuxin Fan[1], Ming Liu[1*], Qingbo Guan[7*] and Lixing Ma[8*]

[1]Department of Endocrinology, Tianjin Medical University General Hospital, Tianjin, China, [2]Department of Endocrinology, The Affiliated Taian City Central Hospital of Qingdao University, Taian, China, [3]School of Public Health, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, [4]Department of Pharmacy, The Affiliated Taian City Central Hospital of Qingdao University, Taian, China, [5]Department of Joint Surgery, The Affiliated Taian City Central Hospital of Qingdao University, Taian, China, [6]School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China, [7]Department of Endocrinology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China, [8]Department of Gastroenterology, The Second Affiliated Hospital of Shandong First Medical University, Taian, China

**Objective:** Bile acids have underlying protective effects on bones structure. Long-term diabetes also causes skeletal disorders including osteoporosis, Charcot arthropathy and renal osteodystrophy. Nevertheless, few studies have reported whether bile acid is associated with bone metabolism in diabetics. This study aimed to explore the relationship between total bile acid (TBA) and bone mineral density (BMD) among patients with type 2 diabetes mellitus (T2DM).

**Methods:** We retrospectively included 1,701 T2DM patients who were hospitalized in Taian City Central Hospital (TCCH), Shandong Province, China between January 2017 to December 2019. The participants were classified into the osteopenia (n = 573), osteoporosis (n= 331) and control groups (n= 797) according to BMD in the lumbar spine and femoral. The clinical parameters, including TBA, bilirubin, vitamin D, calcium, phosphorus and alkaline phosphatase were compared between groups. Multiple linear regression was used to analyze the relationship between TBA and BMD in lumbar spine, femoral, trochiter, ward's triangle region. A logistic regression was conducted to develop a TBA-based diagnostic model for differentiating abnormal bone metabolism from those with normal BMD. We evaluated the performance of model using ROC curves.

**Results:** The TBA level was significantly higher in patients with osteoporosis (Median[M]= 3.300 μmol/L, interquartile range [IQR] = 1.725 to 5.250 μmol/L) compared to the osteopenia group (M = 3.200 μmol/L, IQR = 2.100 to 5.400 μmol/L) and control group (M = 2.750 μmol/L, IQR = 1.800 to 4.600 μmol/L) ($P$ <0.05). Overall and subgroup analyses indicated that TBA was negatively associated with BMD after adjusted for the co-variates (i.e., age, gender, diabetes duration, BMI, total bilirubin, direct bilirubin, indirect bilirubin) ($P$ <0.05). Logistic regression revealed that higher TBA level was associated with increased risk for

abnormal bone metabolism (OR = 1.044, 95% CI = 1.005 to 1.083). A TBA-based diagnostic model was established to identify individuals with abnormal bone metabolism (T-score ≤ -1.0). The area under ROC curve (AUC) of 0.767 (95% CI = 0.730 to 0.804).

**Conclusion:** Our findings demonstrated the potential role of bile acids in bone metabolism among T2DM patients. The circulating TBA might be employed as an indicator of abnormal bone metabolism.

KEYWORDS

type 2 diabetes mellitus (T2DM), total bile acid (TBA), bone mineral density (BMD), osteoporosis, abnormal bone metabolism

# Introduction

Type 2 diabetes mellitus (T2DM), one of metabolic diseases, is mainly caused by insulin deficiency or resistance (1). More than 460 million persons suffer from T2DM globally, accounting for 6.28% of the world's population in 2020 (2). Long-term diabetes commonly induces dysfunctions in multiple tissues and organs, such as brain, cardiovascular system, kidneys and eyes (3). Besides, skeletal disorders have been observed in association with DM, including osteoporosis, Charcot arthropathy and renal osteodystrophy (4). It is believed that disorders of glucose metabolism can damage bone microstructure and increase the incidences of osteoporosis and osteoporosis-associated fracture (5, 6). Bone mineral density (BMD) is a key parameter of bone health and an osteoporosis predictor (7, 8). Clinical evidences have evidenced that T2DM increases the risk of low BMD, osteoporosis and fractures, particularly in older men and postmenopausal women (9).

Total bile acids (TBA), a series of signaling molecules synthesized by liver cells, display biological functions, such as metabolism of glucose and lipid, and regulation of intestinal flora (10). Bile acid-induced activation of G protein-coupled bile acid receptor (TGR5) promotes insulin secretion by increasing intracellular calcium concentration (11). Studies have also identified that circulating TBA was positively correlated with BMD, indicating the potential role of bile acids in the regulation of bone metabolism (12, 13). Bile acids regulate bone metabolism *via* the activation of nuclear receptor, farnesoid X receptor (FXR), membrane receptor, TGR5 and intestinal flora (14–16).

Since TBA regulating bone metabolism is one of the pathophysiological pathways of osteoporosis, we hypothesized that TBA is associated with osteoporosis in diabetic. However, to date, no studies have reported the association with TBA and BMD in diabetic patients. We conducted this retrospective study to identify the relationship between serum TBA and bone metabolism, and to explore the potential role of TBA in the development of osteoporosis in diabetics.

# Methods

## Study participants

A total of 550 T2DM patients who did not fulfill the inclusion criteria or lacked clinical data were excluded. Finally, 1701 patients with T2DM were included from Taian City Central Hospital (TCCH) between January 2017 and December 2019 (Figure S1). The participants were classified into three groups: (1) osteoporosis, (2) osteopenia, and (3) control groups.

The diagnosis of T2DM and osteoporosis was based on World Health Organization (WHO) criteria (17–19), T-score ≤ -2.5 for osteoporosis, between -2.5 to -1.0 for osteopenia, > -1.0 for normality and T-score ≤ -1.0 for abnormal bone metabolism.

Inclusion criteria were as follows: (1) Individuals diagnosed with T2DM; (2) No severe somatic disorders including cardiovascular diseases and cancers; (3) No mental disorders; (4) No diabetic acute complications, including ketoacidosis, lactic acidosis and diabetic hyperosmolarity; (5) Not taking any medications that affect bone metabolism and bile acid metabolism in 6 months. Exclusion criteria: (1) Patients diagnosed with T1DM, gestational diabetes mellitus or other specific types of diabetes; (2) Patients with chronic kidney insufficiency, chronic hepatic insufficiency, liver or renal dysfunction; (3) Patients with endocrine diseases that affect bone metabolism, including parathyroid dysfunction, gonadal diseases and adrenal diseases; (4) Patients with diseases that seriously affect bone metabolism and lead to secondary osteoporosis, such as rheumatic diseases, hematological diseases and digestive disease; (5) Individuals with family history of osteoporosis; (6) Patients with a history of recent exposure to radioactive materials; (7) Patients with history of prolonged bed rest.

This study has been reviewed and approved by the ethics committee of TCCH (No. 2021-05-001). As a retrospective study of clinical dataset, this research was exempt from the request of informed consent from subjects.

## Data collection

The characteristics of age, gender, height, weight, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP) and diabetes duration were collected from clinical records.

Total cholesterol (TC), triglyceride (TG), low density lipoprotein cholesterol (LDL), high density lipoprotein cholesterol (HDL), calcium ions, phosphorus, alkaline phosphatase (ALP), TBA, total bilirubin, direct bilirubin and indirect bilirubin were measured by Modular P800 automatic biochemical analyzer (Roche, German). Glycated hemoglobin A1c (HbA1c) was detected *via* high-performance liquid chromatography (Bio-Rad Laboratories, CA, USA). Fasting blood glucose (FBG) was measured with an automatic analyzer (Hitachi, Tokyo, Japan). Fins, C-peptide and vitamin D were determined by Cobas 6000 electro chemiluminescence (Roche, German). BMD of lumbar spine, femoral, trochiter and ward's triangle region were measured by dual energy X-ray absorptiometry (GE Lunar IDXA, USA).

## Statistical analysis

Continuous data were presented as means and standard deviations (SDs) when normally distributed, otherwise presented as median (M) and interquartile range (IQR). Categorical data were presented as frequencies. For comparisons between multiple groups, one-way analysis of variance (ANOVA) followed by Least-Significant Difference (LSD) test was used for normally distributed data. Kruskal-Wallis test followed by Bonferroni *post hoc* test was used for non-normal distributed data. Chi-square test was used for comparison of categorical data. Jonckheere–Terpstra test was used to assess the trend in TBA level between multiple groups. Multiple linear regression was used to analyze the associations between TBA and BMD. Logistic regression analysis was used to establish a TBA-based diagnostic model to identify individuals with abnormal bone metabolism from those with normal (T-score ≤ -1.0). The receiver operator characteristic (ROC) curve and the area under the curve (AUC) were employed to evaluate the model's performance. Subgroup analysis is performed to assess the association between TBA and BMD based on gender, age group, BMI and menstrual conditions.

A two-side $P$-value < 0.05 was considered statistically significant. Statistical analyses were performed using R packages 4.1.0 (R Core Team) and SPSS 25.0 (IBM, New York).

# Results

## Clinical characteristics of the participants

The basic characteristics of the 1,701 T2DM patients are listed in Table 1. They were classified as control group: 797 individuals with normal BMD, aged (54.8 ± 11.3) years, of whom 68.4% (545/797) were male; osteopenia group: 573 individuals with osteopenia, aged (61.9 ± 9.2) years, of whom 50.3% (288/573) were male; and osteoporosis group: 331 individuals with osteoporosis, aged (67.1 ± 8.7) years, of whom 21.1% (70/331) were male. The results of hepatobiliary metabolism indicators are shown in Table 2 and

TABLE 1 Characteristic description of T2DM patients.

| Indicators | Osteoporosis (n=331) | Osteopenia (n=573) | Control (n=797) | $\chi^2/F$ | $P$ |
|---|---|---|---|---|---|
| Male [n (%)] | 70 (21.1) | 288 (50.3) | 545 (68.4) | 212.273 | <0.001 |
| Female [n (%)] | 261 (78.9) | 285 (49.7) | 252 (31.6) | | |
| Age (year) | 67.1 ± 8.7 *# | 61.9 ± 9.2# | 54.8 ± 11.3 | 195.389 | <0.001 |
| Diabetes duration (month) | 120.0 (72.0,204.0) *# | 120.0 (48.0, 180.0) # | 84.0 (36.0, 144.0) | 46.854 | <0.001 |
| BMI<25 [n (%)] | 195 (85.5) | 269 (47.7) | 304 (38.8) | 41.020 | <0.001 |
| BMI≥25 [n (%)] | 33 (14.5) | 295 (52.3) | 480 (61.2) | | |
| SBP | 143 ± 19# | 141 ± 21# | 139 ± 20 | 4.665 | 0.010 |
| DBP | 76 ± 11*# | 78 ± 11# | 81 ± 11 | 22.305 | 0.001 |
| TC (mmol/L) | 4.540 (3.655,5.365) | 4.460 (3.600, 5.295) | 4.510 (3.700,5.300) | 0.905 | 0.636 |
| TG (mmol/L) | 1.310 (0.890,2.075) # | 1.380 (0.960,2.125) | 1.470 (1.010, 2.412) | 7.570 | 0.023 |
| LDL (mmol/L) | 2.780 (2.110,3.483) | 2.800 (2.010,3.390) | 2.870 (2.170,3.500) | 1.641 | 0.440 |
| HDL (mmol/L) | 1.370 (1.160,1.670) # | 1.320 (1.100,1.161) | 1.270 (1.080,1.510) | 17.978 | <0.001 |
| C-Peptide (ng/ml) | 0.910 (0.600,1.520) # | 1.050 (0.640,1.600) | 1.150 (0.708,1.760) | 8.265 | 0.016 |
| FINS (uIU/ml) | 3.300 (1.725,5.250) | 8.145 (5.538,12.335) | 8.750 (6.100,13.523) | 3.773 | 0.152 |
| HbA1C (%) | 8.786 ± 2.366 | 8.945 ± 2.257 | 8.957 ± 2.188 | 0.639 | 0.528 |
| FBG (mmol/L) | 9.250 (6.980,12.060)# | 9.235 (7.355,12.260) | 9.820 (7.630,13.185) | 11.942 | 0.003 |

BMI, body mass index; SBP, Systolic blood pressure; DBP, diastolic blood pressure; FINS, Fasting insulin; HbA1c, Hemoglobin A1c; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, cholesterol; TG, triglycerides; FBG, Fasting blood glucose; *$P$ < 0.05 compared with osteopenia group; #$P$ < 0.05 compared with the controls group.

TABLE 2 Hepatobiliary metabolism indicators of T2DM patients.

| Indicators | Osteoporosis (n=331) | Osteopenia (n=573) | Control (n=797) | $\chi^2$ | *P* |
|---|---|---|---|---|---|
| TBA (μmol/L) | 3.300 (1.725,5.250) | 3.200 (2.100,5.400) # | 2.750 (1.800,4.600) | 6.435 | 0.040 |
| Total bilirubin (μmol/L) | 8.600 (6.650,11.600) *# | 10.400 (7.775,13.500) | 10.800 (8.100,14.100) | 48.848 | <0.001 |
| Direct bilirubin (μmol/L) | 3.700 (3.000,4.750) *# | 4.300 (3.400,5.500) | 4.400 (3.400,5.500) | 35.291 | <0.001 |
| Indirect Bilirubin (μmol/L) | 4.800 (3.600,6.900) *# | 6.100 (4.250,8.000) | 6.200 (4.500,8.500) | 40.169 | <0.001 |

TBA, total bile acid; *$P < 0.05$ compared with osteopenia group; #$P < 0.05$ compared with the controls group.

Figure 1. The results of bone metabolism indicators are presented in Table 3.

Significant differences were identified between age, gender, BMI, diabetes duration, total bilirubin, direct bilirubin, indirect bilirubin, SBP, DBP, FBG, TG, LDL, C-peptide, ALP and BMD indicators between the multigroup (i.e., control, osteopenia and osteoporosis groups) ($P < 0.05$).

Age showed a gradual increase in the values and there was a significant difference in pairwise comparisons between the multigroup. In addition, DBP presented a decreasing trend among control, osteopenia and osteoporosis groups ($P < 0.05$) (Table 1).

## The relevance between serum TBA levels and BMD

The serum TBA levels in the osteopenia group were (3.200 μmol/L [IQR = 2.100 to 5.400 μmol/L]), which were significantly higher than that in the control group (2.750 μmol/L, [IQR = 1.800 to 4.600 μmol/L]) ($P < 0.05$). Furthermore, Jonckheere–Terpstra test found TBA levels presented a significant increasing trend among control, osteopenia and osteoporosis groups ($P< 0.05$) (Table 2).

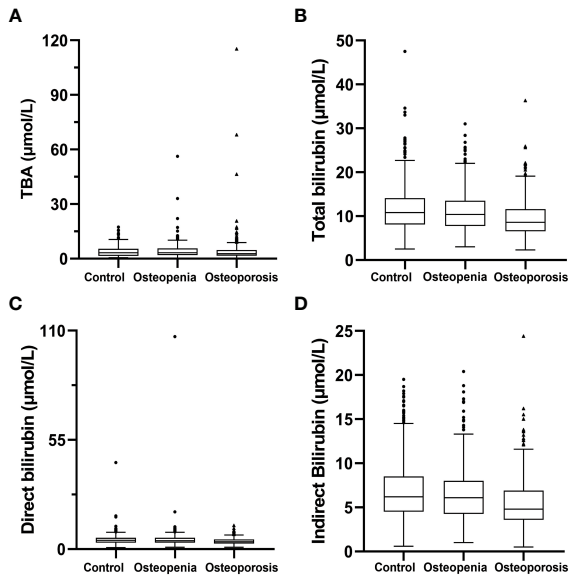



FIGURE 1
Hepatobiliary metabolism indicators. **(A)** total bile acid; **(B)** total bilirubin; **(C)** direct bilirubin; **(D)** indirect bilirubin.

Multiple linear regression revealed that TBA level were independent determinants associated with BMD in third lumbar vertebrae (L3), fourth lumbar vertebrae (L4), total lumbar spine, femoral neck, and ward's triangle region ($P< 0.05$) (Table 4).

## Subgroup analysis

Based on gender, age group (classify into < 60 and ≥ 60 years), BMI (classify into < 25 and ≥ 25) and menstrual conditions, a subgroup analysis is listed in Tables S1–S3.

Among the participants aged < 60 and BMI < 25, including men and women, TBA level was negatively associated with total lumbar spine BMD (Table S1). For men with normal BMI (BMI<25), TBA level was negatively associated with BMD levels in the femoral and ward's triangle region (Table S2). In postmenopausal population, TBA level was negatively associated with the BMD levels in first lumbar vertebrae(L1), second lumbar vertebrae(L2), third lumbar vertebrae (L3), fourth lumbar vertebrae (L4), femoral neck and total lumbar spine (Table S3).

## TBA-based diagnostic model for abnormal bone metabolism

Logistic regression analysis was used to establish a TBA-based diagnostic model to identify individuals with abnormal bone metabolism among T2DM patients (Table 5). A higher TBA level (OR = 1.044, 95% CI = 1.005 to 1.083) was associated with increased risk for abnormal bone metabolism. Older age (OR = 1.060, 95% CI = 1.042 to 1.079) and female (OR = 2.236, 95% CI = 1.619 to 3.087) correlated with a higher risk for abnormal bone metabolism. Higher BMI (OR =0.872, 95% CI =0.831 to 0.916) was associated with a lower risk for abnormal bone metabolism (Figure 2).

We then established a diagnostic model using TBA, age, gender and BMI. As depicted in Figure 3, a ROC curve was used to assess the performance of model, which showed the AUC of 0.767(95% CI = 0.730 to 0.804), with a sensitivity of 65.4% and a specificity of 77.2% at a cut-off-value of 0.656.

# Discussion

Our findings identify, for the first time in T2DM patients, the TBA level among diabetic patients with osteoporosis was higher than those with the normal BMD and osteopenia. TBA level

TABLE 3 Bone metabolism indicators of T2DM patients.

| Indicators | Osteoporosis (n=331) | Osteopenia (n=573) | Control (n=797) | $\chi^2/F$ | *P* |
|---|---|---|---|---|---|
| Calcium ion (mmol/L) | 2. 380 (2.300, 2.460) | 2.380 (2.310, 2.445) | 2.380 (2.310, 2.450) | 1.222 | 0.295 |
| Phosphorus (mmol/L) | 1.170 (1.065,1.280) | 1.180 (1.030, 1.280) | 1.170 (1.043, 1.290) | 0.608 | 0.738 |
| Vitamin D (ng/ml) | 17.150 (11.800,22.275) | 17.750 (12.600, 24.200) | 18.600 (13.750, 23.775) | 4.325 | 0.115 |
| ALP (u/l) | 71.00 (61.00,87.00) $^{\#}$ | 71.00 (60.00, 85.25) $^{\#}$ | 67.00 (56.00, 82.00) | 4.664 | 0.010 |
| L1BMD (g/cm²) | 0.787 ± 0.167$^{*\#}$ | 0.940 ± 0.150$^{\#}$ | 1.100 ± 0.177 | 433.627 | <0.001 |
| L2BMD (g/cm²) | 0.831 ± 0.135$^{*\#}$ | 1.033 ± 0.150$^{\#}$ | 1.203 ± 0.171 | 662.445 | <0.001 |
| L3BMD (g/cm²) | 0.911 ± 0.158$^{*\#}$ | 1.124 ± 0.162$^{\#}$ | 1.286 ± 0.186 | 554.659 | <0.001 |
| L4BMD (g/cm²) | 0.942 ± 0.172$^{*\#}$ | 1.135 ± 0.176$^{\#}$ | 1.281 ± 0.200 | 389.901 | <0.001 |
| Total lumbar spine BMD (g/cm²) | 0.872 ± 0.133$^{*\#}$ | 1.068 ± 0.147$^{\#}$ | 1.228 ± 0.176 | 590.632 | <0.001 |
| Femoral neck BMD (g/cm²) | 0.726 ± 0.309$^{*\#}$ | 0.821 ± 0.128$^{\#}$ | 1.001 ± 0.256 | 190.556 | <0.001 |
| Trochiter BMD (g/cm²) | 0.635 ± 0.113$^{*\#}$ | 0.749 ± 0.133$^{\#}$ | 0.881 ± 0.147 | 411.295 | <0.001 |
| Femoral shaft BMD (g/cm²) | 0.967 ± 0.194$^{*\#}$ | 1.133 ± 0.198$^{\#}$ | 1.288 ± 0.185 | 339.836 | <0.001 |
| Ward's triangle region BMD (g/cm²) | 0.514 ± 0.122$^{*\#}$ | 0.636 ± 0.115$^{\#}$ | 0.800 ± 0.290 | 109.966 | <0.001 |
| Total femoral BMD (g/cm²) | 0.827 ± 0.521$^{*\#}$ | 0.916 ± 0.139$^{\#}$ | 1.076 ± 0.197 | 25.255 | <0.001 |

ALP, Alkaline Phosphatase; L1BMD, first lumbar vertebra BMD; L2BMD, second lumbar vertebra BMD; L3 BMD, third lumbar vertebra BMD; L4 BMD, fourth lumbar vertebra BMD; $^{*}P < 0.05$ compared with osteopenia group; $^{\#}P < 0.05$ compared with the controls group.

negatively correlated with BMD in lumbar spine, femoral neck, femoral shaft and ward's triangle region. The serum TBA level could be employed as a predictor of BMD in T2DM patients. The TBA may be employed as a new therapeutic target for osteoporosis in diabetics (20), which has clinical significance for prevention of osteoporosis in diabetics.

Diabetes mellitus is a complex multifactorial disease (21). In mainland China, diabetes affects 11.2% of adults (22). An important complication of diabetes is osteoporosis. In diabetic, the disorder of glucose and lipid metabolism changes tissue structure and adversely affects bone metabolism, which increased the risk of osteoporosis and fracture (23). Diabetic people suffer a higher risk of fracture compared to the healthy (24, 25), which also correlated with diabetes duration (26–28). Circulating sclerostin level is significantly higher in diabetic persons, which inhibits the function of osteoblasts and bone formation, thus increasing the risk of osteoporosis (29). Osteocalcin is an essential protein for the process of bone formation. Hyperglycemia impairs the function of osteoblasts on synthesis of osteocalcin and then downregulates the osteocalcin level, leading to inhibition of bone formation (30). Increased oxidative stress of platelet mitochondria in T2DM patients interferes with physiological function of bone marrow cells and impairs bone metabolism (31).

TBA, including Chenodeoxycholic acid (CDCA), Tauroursodeoxycholic acid (TUDCA), Deoxycholic acid (DCA), Lithocholic acid (LCA), and 6-alpha-ethyl-chenodeoxycholic acid (6-ECDCA) and a series of endocrine substance with various physiological functions, is generally synthesized in the liver (32). Multiple pathophysiological mechanisms support the relation between TBA and bone metabolism, including FXR, intestinal flora and Oxidative stress. First, TBA facilitates the differentiation of bone marrow mesenchymal cells into osteoblasts *in vitro* (33, 34). After being treated with DCA *in vitro*, the activity of ALP in bone marrow stromal cells was improved, leading to bone erosion (35). Second, TBA have a positive regulatory effect on osteogenesis by FXR, the principle is CDCA and 6-ECDCA activates bile acid nuclear receptor FXR. FXR increases the activity of extracellular regulatory protein kinase (ERK) by upregulating runt-related transcription factor 2 (Runx2), which promotes differentiation of mesenchymal progenitor cells into osteoblast (36). FXR, provokes the expression of ALP, and DNA-binding activity of

TABLE 4 The association of TBA levels with BMD.

| Sites | $\beta$ | 95% CI of $\beta$ | *P* |
|---|---|---|---|
| L3 BMD | -0.003 | -0.003 ~ -0.001 | 0.019 |
| L4 BMD | -0.003 | -0.005 ~ -0.001 | 0.043 |
| Total lumbar spine BMD | -0.003 | -0.005 ~ -0.001 | 0.013 |
| Femoral neck BMD | -0.003 | -0.004 ~ -0.001 | 0.037 |
| Ward's triangle region BMD | -0.004 | -0.007 ~ -0.001 | 0.003 |

BMD, bone mineral density; TBA, total bile acid; $\beta$, regression coefficient; CI, confidence intervals; L3, third lumbar vertebra; L4, fourth lumbar vertebra; Adjusted for gender, age, BMI, diabetes duration, total bilirubin, direct bilirubin, and indirect bilirubin.

TABLE 5 Logistic regression of abnormal bone metabolism influence factors in T2DM patients.

| Factors | $\beta$ | SE | Walds $\chi 2$ | P | OR (95%CI) |
|---|---|---|---|---|---|
| Gender (male) | 0.805 | 0.165 | 23.909 | <0.001 | 2.236 (1.619~3.087) |
| Age | 0.059 | 0.009 | 42.355 | <0.001 | 1.060 (1.042~1.079) |
| Diabetes duration | 0.001 | 0.001 | 2.051 | 0.152 | 1.001 (0.999~1.003) |
| BMI | -0.136 | 0.025 | 30.765 | <0.001 | 0.872 (0.831~0.916) |
| TBA | 0.043 | 0.019 | 5.017 | 0.025 | 1.044 (1.005~1.083) |

BMI, body mass index; TBA, total bile acid; SE, standard error; OR, odds ratio; CI, confidence intervals; $\beta$, regression coefficient.

Runx2, the bone transcription factor (37). Third, intestinal flora regulates bone metabolism through its 7-dehydroxylation producing LCA, a ligand for the vitamin D receptor. Vitamin D regulates the gene coding of bone protein, osteocalcin and receptor activator of nuclear factor-kB ligand (RANKL) (38). Furthermore, LCA affect the formation of osteoblasts and osteoclasts by repressing the expression of calcitonin gene and RANKL gene (16). A dynamic balance exists between diet and intestinal flora-bile acid (39). However, high-fat and cholesterol diet can alter the composition of bile acids in the gut, causing imbalance of intestinal flora and aggravation of bile acid metabolism disorders (40). In addition, oxidative stress plays a role of inhibit osteogenesis by affecting the differentiation, proliferation and apoptosis of osteocytes, and TBA regulates bone metabolism through alleviating oxidative stress (41–44).

Thus far, the clinical evidence for the connection between TBA and bone metabolism is limited. Bile acid malabsorption(BAM) can reduce the absorption of vitamin D and then patients may develop low BMD (45). TUDCA enhance bone tissue regeneration in skull defect models, which can be used as a potential alternative drug for bone regeneration (46).

Following the STROBE guideline (47), we conducted subgroup analyses to make a better use of the data. A retrospective study in China of 2230 healthy persons with aged < 60 and BMI < 25 pointed out that serum TBA was positively correlated with BMD (13). We have different findings, which is that TBA and BMD are negatively related in diabetics with aged < 60 and BMI < 25. Study reported the TBA level was positively correlated with the BMD in postmenopausal healthy population (12). However, we found the TBA level was negatively related with the BMD in postmenopausal diabetics. The above indicated that the pathway of bile acids regulating bone metabolism might be interfered in diabetics. By reference to the mechanism of insulin resistance, we hypothesize that bile acids present a compensatory elevation and have an antagonist effect to osteoporosis. The protective role of TBA in bone metabolism is needed to be explored. In accord with our findings, study reported older age was correlated with decreased BMD and a positive correlation between BMI and BMD (48).

## Limitations

Our findings provided a novel insight into skeletal health in diabetics. Nevertheless, there are some limitations. First, this is a retrospective study based on clinical dataset, which cannot prove the causal relationship between TBA level and bone metabolism. Second, some of participants with osteoporosis has a history of supplementation of calcium and vitamin D, which might bias our findings. Third, although the sample size of this study is high, our evidence might lack of high generalizability and extrapolation due to a single-center study design. Therefore, multi-center prospective studies are needed to offer further identification.

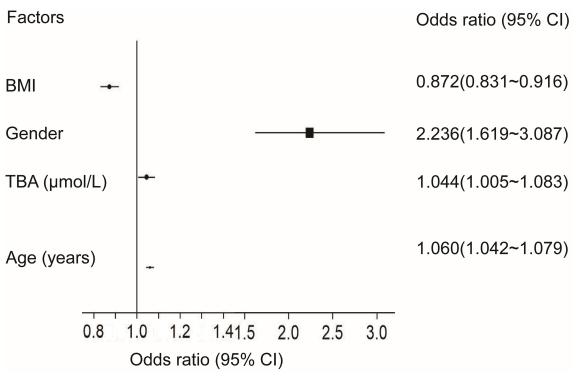



FIGURE 2
Forest plots of logistic regression analysis. CI, confidence interval.

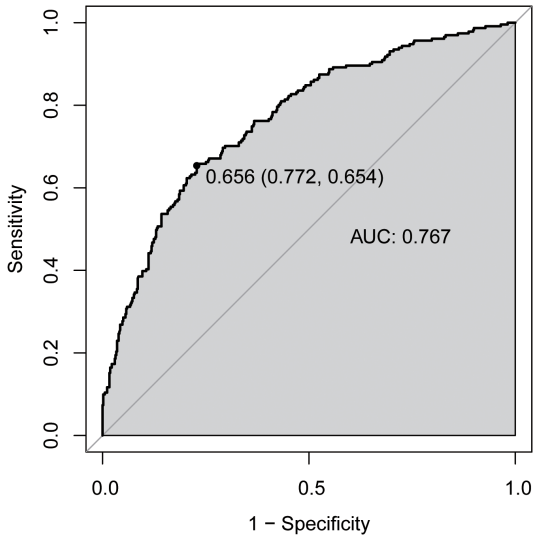



FIGURE 3
Receiver operator characteristic (ROC) curves for diagnostic model of abnormal bone metabolism. AUC, area under the curve.

## Conclusion

We observed the negative relevance between TBA and BMD in diabetics, suggesting that a role of bile acids in BMD and bone metabolism among T2DM patients. The circulating TBA level might be employed as an indicator of abnormal bone metabolism.

## Data availability statement

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

## Ethics statement

This study has been reviewed and approved by the ethics committee of TCCH (No. 2021-05-001). As a retrospective study of clinical dataset, this research was exempt from the request of informed consent from subjects.

## Author contributions

ML, QG, and LM designed the study. QW, LD, CX, and YF contributed to data collection. SY, HL, and YG made statistical analysis and manuscript writing. ML, QG, and LM revised the manuscript. All authors have approved the final version.

## Funding

This research was funded by the National Natural Science Foundation of China (81773527).

## 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.2023.1153205/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Flow-chart of the participants.

## References

1. Wang L, Li H, Hao J, Liu C, Wang J, Feng J, et al. Thirty-six months recurrence after acute ischemic stroke among patients with comorbid type 2 diabetes: A nested case-control study. *Front Aging Neurosci* (2022) 14:999568. doi: 10.3389/fnagi.2022.999568

2. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes - global burden of disease and forecasted trends. *J Epidemiol Glob Health* (2020) 10(1):107–11. doi: 10.2991/jegh.k.191028.001

3. Rines AK, Sharabi K, Tavares CD, Puigserver P. Targeting hepatic glucose metabolism in the treatment of type 2 diabetes. *Nat Rev Drug Discovery* (2016) 15 (11):786–804. doi: 10.1038/nrd.2016.151

4. Kurra S, Fink DA, Siris ES. Osteoporosis-associated fracture and diabetes. *Endocrinol Metab Clin North Am* (2014) 43(1):233–43. doi: 10.1016/j.ecl.2013.09.004

5. Wang W, Jiang B, Ye S, Qian L. (2018). Risk Factor Analysis of Bone Mineral Density Based on Feature Selection in Type 2 Diabetes. *IEEE International Conference on Big Knowledge (ICBK) Singapore*, (2018), pp. 221–6. doi: 10.1109/ICBK.2018.00037

6. Nguyen K-D, Bagheri B, Bagheri H. Drug-induced bone loss: A major safety concern in Europe. *Expert Opin Drug Saf* (2018) 17(10):1005–14. doi: 10.1080/14740338.2018.1524868

7. Zhu H, Fang J, Luo X, Yu W, Zhao Y, Li X, et al. A survey of bone mineral density of healthy han adults in China. *Osteoporosis Int* (2010) 21(5):765–72. doi: 10.1007/s00198-009-1010-2

8. Zhang Z, Ho S, Chen Z, Zhang C, Chen Y. Reference values of bone mineral density and prevalence of osteoporosis in Chinese adults. *Osteoporosis Int* (2014) 25 (2):497–507. doi: 10.1007/s00198-013-2418-2

9. Ren H, Ma X, Shao Y, Han J, Yang M, Wang Q. Correlation between serum mir-154-5p and osteocalcin in males and postmenopausal females of type 2 diabetes with different urinary albumin creatinine ratios. *Front Endocrinol (Lausanne)* (2019) 10:542. doi: 10.3389/fendo.2019.00542

10. Chavez-Talavera O, Tailleux A, Lefebvre P, Staels B. Bile acid control of metabolism and inflammation in obesity, type 2 diabetes, dyslipidemia, and nonalcoholic fatty liver disease. *Gastroenterology* (2017) 152(7):1679–94 e3. doi: 10.1053/j.gastro.2017.01.055

11. Kumar DP, Rajagopal S, Mahavadi S, Mirshahi F, Grider JR, Murthy KS, et al. Activation of transmembrane bile acid receptor Tgr5 stimulates insulin secretion in pancreatic beta cells. *Biochem Biophys Res Commun* (2012) 427(3):600–5. doi: 10.1016/j.bbrc.2012.09.104

12. Zhao YX, Song YW, Zhang L, Zheng FJ, Wang XM, Zhuang XH, et al. Association between bile acid metabolism and bone mineral density in postmenopausal women. *Clinics (Sao Paulo)* (2020) 75:e1486. doi: 10.6061/clinics/2020/e1486

13. Liu J, Chen Y, Luo Q. The association of serum total bile acids with bone mineral density in Chinese adults aged 20-59: A retrospective cross-sectional study. *Front Endocrinol (Lausanne)* (2022) 13:817437. doi: 10.3389/fendo.2022.817437

14. Zheng T, Kim NY, Yim M. Fexaramine inhibits receptor activator of nuclear factor-Kb ligand-induced osteoclast formation *Via* nuclear factor of activated T cells signaling pathways. *J Bone Metab* (2017) 24(4):207–15. doi: 10.11005/jbm.2017.24.4.207

15. Li Z, Huang J, Wang F, Li W, Wu X, Zhao C, et al. Dual targeting of bile acid receptor-1 (Tgr5) and farnesoid X receptor (Fxr) prevents estrogen-dependent bone loss in mice. *J Bone Mineral Res* (2019) 34(4):765–76. doi: 10.1002/jbmr.3652

16. Ruiz-Gaspa S, Guanabens N, Enjuanes A, Peris P, Martinez-Ferrer A, de Osaba MJ, et al. Lithocholic acid downregulates vitamin d effects in human osteoblasts. *Eur J Clin Invest* (2010) 40(1):25–34. doi: 10.1111/j.1365-2362.2009.02230.x

17. Kanis JA, Melton LJ3rd, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. *J Bone Miner Res* (1994) 9(8):1137–41. doi: 10.1002/jbmr.5650090802

18. Kanis JA. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: Synopsis of a who report. who study group. *Osteoporos Int* (1994) 4(6):368–81. doi: 10.1007/BF01622200

19. Resnick HE, Harris TB, Harris T, Cook D, Cappuccio FP, Davies M, et al. New diagnostic criteria for diabetes mellitus. *Bmj* (1999) 318(7182):531–. doi: 10.1136/bmj.318.7182.531

20. Wang J, Yan D, Zhao A, Hou X, Zheng X, Chen P, et al. Discovery of potential biomarkers for osteoporosis using lc-Ms/Ms metabolomic methods. *Osteoporosis Int* (2019) 30(7):1491–9. doi: 10.1007/s00198-019-04892-0

21. Koffi C, Soleti R, Nitiema M, Mallegol P, Hilairet G, Chaigneau J, et al. Ethanol extract of leaves of cassia siamea lam protects against diabetes-induced insulin resistance, hepatic, and endothelial dysfunctions in Ob/Ob mice. *Oxid Med Cell Longev* (2019) 2019:6560498. doi: 10.1155/2019/6560498

22. Li Y, Teng D, Shi X, Qin G, Qin Y, Quan H, et al. Prevalence of diabetes recorded in mainland China using 2018 diagnostic criteria from the American diabetes association: National cross sectional study. *BMJ* (2020) 369:m997. doi: 10.1136/bmj.m997

23. Ye Y, Zhao C, Liang J, Yang Y, Yu M, Qu X. Effect of sodium-glucose Co-transporter 2 inhibitors on bone metabolism and fracture risk. *Front Pharmacol* (2019) 9:1517. doi: 10.3389/fphar.2018.01517

24. Janghorbani M, Van Dam RM, Willett WC, Hu FB. Systematic review of type 1 and type 2 diabetes mellitus and risk of fracture. *Am J Epidemiol* (2007) 166(5):495–505. doi: 10.1093/aje/kwm106

25. Fan Y, Wei F, Lang Y, Liu Y. Diabetes mellitus and risk of hip fractures: A meta-analysis. *Osteoporos Int* (2016) 27(1):219–28. doi: 10.1007/s00198-015-3279-7

26. Forsen L, Meyer HE, Midthjell K, Edna TH. Diabetes mellitus and the incidence of hip fracture: Results from the nord-trondelag health survey. *Diabetologia* (1999) 42 (8):920–5. doi: 10.1007/s001250051248

27. Ivers RQ, Cumming RG, Mitchell P, Peduto AJ, Blue Mountains Eye S. Diabetes and risk of fracture: The blue mountains eye study. *Diabetes Care* (2001) 24(7):1198–203. doi: 10.2337/diacare.24.7.1198

28. Vestergaard P, Rejnmark L, Mosekilde L. Diabetes and its complications and their relationship with risk of fractures in type 1 and 2 diabetes. *Calcif Tissue Int* (2009) 84(1):45–55. doi: 10.1007/s00223-008-9195-5

29. Gennari L, Merlotti D, Valenti R, Ceccarelli E, Ruvio M, Pietrini MG, et al. Circulating sclerostin levels and bone turnover in type 1 and type 2 diabetes. *J Clin Endocrinol Metab* (2012) 97(5):1737–44. doi: 10.1210/jc.2011-2958

30. Kanazawa I, Yamaguchi T, Yamamoto M, Yamauchi M, Kurioka S, Yano S, et al. Serum osteocalcin level is associated with glucose metabolism and atherosclerosis parameters in type 2 diabetes mellitus. *J Clin Endocrinol Metab* (2009) 94(1):45–9. doi: 10.1210/jc.2008-1455

31. Avila C, Huang RJ, Stevens MV, Aponte AM, Tripodi D, Kim KY, et al. Platelet mitochondrial dysfunction is evident in type 2 diabetes in association with modifications of mitochondrial anti-oxidant stress proteins. *Exp Clin Endocrinol Diabetes* (2012) 120(4):248–51. doi: 10.1055/s-0031-1285833

32. Luo ZL, Cheng L, Wang T, Tang LJ, Tian FZ, Xiang K, et al. Bile acid transporters are expressed and heterogeneously distributed in rat bile ducts. *Gut Liver* (2019) 13(5):569–75. doi: 10.5009/gnl18265

33. Pellicciari R, Fiorucci S, Camaioni E, Clerici C, Costantino G, Maloney PR, et al. 6alpha-Ethyl-Chenodeoxycholic acid (6-ecdca), a potent and selective fxr agonist endowed with anticholestatic activity. *J Med Chem* (2002) 45(17):3569–72. doi: 10.1021/jm025529g

34. Costantino G, Macchiarulo A, Entrena-Guadix A, Camaioni E, Pellicciari R. Binding mode of 6ecdca, a potent bile acid agonist of the farnesoid X receptor (Fxr). *Bioorganic Med Chem Lett* (2003) 13(11):1865–8. doi: 10.1016/s0960-894x(03)00281-6

35. Id Boufker H, Lagneaux L, Fayyad-Kazan H, Badran B, Najar M, Wiedig M, et al. Role of farnesoid X receptor (Fxr) in the process of differentiation of bone marrow stromal cells into osteoblasts. *Bone* (2011) 49(6):1219–31. doi: 10.1016/j.bone.2011.08.013

36. Cho SW, An JH, Park H, Yang JY, Choi HJ, Kim SW, et al. Positive regulation of osteogenesis by bile acid through fxr. *J Bone Miner Res* (2013) 28(10):2109–21. doi: 10.1002/jbmr.1961

37. Caron S, Huaman Samanez C, Dehondt H, Ploton M, Briand O, Lien F, et al. Farnesoid X receptor inhibits the transcriptional activity of carbohydrate response element binding protein in human hepatocytes. *Mol Cell Biol* (2013) 33(11):2202–11. doi: 10.1128/MCB.01004-12

38. Kitazawa R, Mori K, Yamaguchi A, Kondo T, Kitazawa S. Modulation of mouse rankl gene expression by Runx2 and vitamin D3. *J Cell Biochem* (2008) 105(5):1289–97. doi: 10.1002/jcb.21929

39. Tang WHW, Bäckhed F, Landmesser U, Hazen SL. Intestinal microbiota in cardiovascular health and disease: Jacc state-of-the-Art review. *J Am Coll Cardiol* (2019) 73(16):2089–105. doi: 10.1016/j.jacc.2019.03.024

40. Rao Y, Kuang Z, Li C, Guo S, Xu Y, Zhao D, et al. Gut akkermansia muciniphila ameliorates metabolic dysfunction-associated fatty liver disease by regulating the metabolism of l-aspartate *Via* gut-liver axis. *Gut Microbes* (2021) 13(1):1–19. doi: 10.1080/19490976.2021.1927633

41. Perez MJ, Briz O. Bile-Acid-Induced cell injury and protection. *World J Gastroenterol* (2009) 15(14):1677–89. doi: 10.3748/wjg.15.1677

42. Pansini F, Mollica G, Bergamini CM. Management of the menopausal disturbances and oxidative stress. *Curr Pharm Des* (2005) 11(16):2063–73. doi: 10.2174/1381612054065819

43. Cervellati C, Bergamini CM. Oxidative damage and the pathogenesis of menopause related disturbances and diseases. *Clin Chem Lab Med* (2016) 54(5):739–53. doi: 10.1515/cclm-2015-0807

44. Bonaccorsi G, Piva I, Greco P, Cervellati C. Oxidative stress as a possible pathogenic cofactor of post-menopausal osteoporosis: Existing evidence in support of the axis oestrogen deficiency-redox imbalance-bone loss. *Indian J Med Res* (2018) 147 (4):341–51. doi: 10.4103/ijmr.IJMR_524_18

45. Hanly R, Ryan N, Snelling H, Walker-Bone K, Dizdarevic S, Peters AM. Association between bile acid turnover and osteoporosis in postmenopausal women. *Nucl Med Commun* (2013) 34(6):597–600. doi: 10.1097/MNM.0b013e3283608993

46. Cha BH, Jung MJ, Moon BK, Kim JS, Ma Y, Arai Y, et al. Administration of tauroursodeoxycholic acid enhances osteogenic differentiation of bone marrow-derived mesenchymal stem cells and bone regeneration. *Bone* (2016) 83:73–81. doi: 10.1016/j.bone.2015.10.011

47. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (Strobe) statement: Guidelines for reporting observational studies. *Int J Surg (London England)* (2014) 12(12):1495–9. doi: 10.1016/j.ijsu.2014.07.013

48. Mozaffarieh M, Fontana Gasio P, Schötzau A, Orgül S, Flammer J, Kräuchi K. Thermal discomfort with cold extremities in relation to age, gender, and body mass index in a random sample of a Swiss urban population. *Popul Health metrics* (2010) 8:17. doi: 10.1186/1478-7954-8-17

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 05 May 2023
DOI 10.3389/fendo.2023.1098078



OPEN ACCESS

EDITED BY
Di Liu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Roberto Scicali,
University of Catania, Italy
Giuseppe Della Pepa,
University of Naples Federico II, Italy

*CORRESPONDENCE
Liying Guan
Liliguansun001@163.com
Chao Xu
doctorxuchao@163.com
Fang Yan
yanfang19851234@163.com

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

RECEIVED 14 November 2022
ACCEPTED 23 March 2023
PUBLISHED 05 May 2023

CITATION
Cheng Y, Zhang Q, Li H, Zhou G, Shi P, Zhang X, Guan L, Yan F and Xu C (2023) Remnant cholesterol, stronger than triglycerides, is associated with incident non-alcoholic fatty liver disease. *Front. Endocrinol.* 14:1098078. doi: 10.3389/fendo.2023.1098078



# Remnant cholesterol, stronger than triglycerides, is associated with incident non-alcoholic fatty liver disease

Yiping Cheng[1,2], Qiang Zhang[3], Haizhen Li[4], Guangshuai Zhou[5], Ping Shi[1,2], Xu Zhang[1,2], Liying Guan[6]*, Fang Yan[7] and Chao Xu[1,2]*

[1]Department of Endocrinology and Metabolism, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, China, [2]Department of Endocrinology and Metabolism, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, China, [3]Department of Critical Care Medicine, Zibo Central Hospital, Zibo, Shandong, China, [4]Department of Endocrinology, Dongying City District People Hospital, Dongying, Shandong, China, [5]Department of Scientific Research and Cooperation, Zibo Central Hospital, Zibo, Shandong, China, [6]Department of Health Examination Center, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, China, [7]Department of Pain Management, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, China

**Introduction:** Non-alcoholic fatty liver disease (NAFLD) is characterized by excess accumulation of triglycerides within the liver. However, whether the circulating levels of triglycerides and cholesterol transported in triglyceride-rich lipoproteins (remnant cholesterol, remnant-C) are related to the occurrence of NAFLD has not yet been studied. This study aims to assess the association of triglycerides and remnant-C with NAFLD in a Chinese cohort of middle aged and elderly individuals.

**Methods:** All subjects in the current study are from the 13,876 individuals who recruited in the Shandong cohort of the REACTION study. We included 6,634 participants who had more than one visit during the study period with an average follow-up time of 43.34 months. The association between lipid concentrations and incident NAFLD were evaluated by unadjusted and adjusted Cox proportional hazard models. The potential confounders were adjusted in the models including age, sex, hip circumference (HC), body mass index (BMI), systolic blood pressure, diastolic blood pressure, fasting plasma glucose (FPG), diabetes status and cardiovascular disease (CVD) status.

**Results:** In multivariable-adjusted Cox proportional hazard model analyses, triglycerides (hazard ratio[HR], 95% confidence interval [CI]:1.080,1.047-1.113; p<0.001), high-density lipoprotein cholesterol (HDL-C) (HR, 95% CI: 0.571,0.487-0.670; p<0.001), and remnant-C (HR, 95% CI: 1.143,1.052-1.242; p=0.002), but not total cholesterol (TC) or low-density lipoprotein cholesterol (LDL-C), were associated with incident NAFLD. Atherogenic dyslipidemia (triglycerides>1.69 mmol/L, HDL-C<1.03 mmol/L in men or<1.29 mmol/L in women) was also associated with NAFLD (HR, 95% CI: 1.343,1.177-1.533; p<0.001). Remnant-C levels were higher in females than in males and increased with increasing BMI and in participants with diabetes and CVD compared with those without diabetes or CVD. After adjusting for other

factors in the Cox regression models, we found that serum levels of TG and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes in women group, non-cardiovascular disease status, non-diabetes status and middle BMI categories (24 to 28 kg/m2).

**Discussion:** In the middle aged and elderly subset of the Chinese population, especially those who were women, non-CVD status, non-diabetes status and middle BMI status (24 to 28 kg/m2), levels of triglycerides and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes independent of other risk factors.

KEYWORDS

Non-alcoholic fatty liver disease, Remnant cholesterol, Triglyceride, Fat metabolization, Longitudinal retrospective cohort study

# 1 Introduction

Non-alcoholic fatty liver disease (NAFLD), characterized by excessive intrahepatic lipid accumulation, is the most prevalent chronic liver disease in the world (1). In addition to progression from simple steatosis to nonalcoholic steatohepatitis (NASH), cirrhosis and hepatocellular carcinoma, NAFLD patients have an increased risk for cardiovascular disease (CVD) morbidity and mortality. Despite the huge investment in drug development, there are still no effective therapies targeting NAFLD. Clearly, identification and elimination of the risk factors that promote NAFLD development and progression are essential and promising therapeutic strategy that can reduce the incidence of NAFLD.

Because intrahepatic lipid accumulation results from lipid metabolism abnormalities, it takes for granted that dyslipidemia can cause NAFLD. However, there are very few studies on the role of different types of dyslipidemia in the development of NAFLD, and the research conclusions are inconsistent. Sun et al. have found the elevation of low-density lipoprotein cholesterol (LDL-C) level within the normal range appears to make a significant contribution to an increased risk of developing NAFLD (2) and previous studies have demonstrated that patients with NAFLD have significantly increased of oxidized LDL-C levels (3, 4). A cross-sectional and hospital-based study in Alexandria was performed to verify that NAFLD in outpatient schoolchildren aged 6-18years was significantly associated with high triglycerides (TG) and low high-density lipoprotein cholesterol (HDL-C) (5). A population-based study has shown NAFLD in 6814 participants aged 45-84 years was associated with higher fasting TG, lower serum HDL-C but no difference in total cholesterol (TC) or LDL-C (6). However, a detailed description, whether the progression of time affects the association between components of dyslipidemia and NAFLD, has been lacked in adults over the past decades.

Remnant cholesterol (Remnant-C) is the residue produced by triglyceride-rich lipoproteins (TRLs) metabolism, that is, chylomicrons (CM) and very low-density lipoproteins (VLDL) are lipolyzed by lipoprotein lipase (LPL) to lose TG and produce metabolic residues rich in cholesterol esters (7). Recently, studies have shown that remnant-C is highly correlated with coronary heart disease (CHD) and insulin resistance (IR) in the general population. As the fact that NAFLD is associated with CHD and IR has been widely recognized, exploration on the relationship between serum remnant-C levels and NAFLD developing is also sorely needed.

In this longitudinal retrospective cohort study, we aimed to explore the association of the components of dyslipidemia and serum remnant-C levels with the occurrence of NAFLD in a Chinese cohort of middle aged and elderly individuals(middle-aged, 45~59 years old; elderly, 60 years old and above; the average age is 56.94 years; mainly from Shandong Province), in an attempt to expand our understanding of the remnant-C as a possible risk factor of NAFLD. To our knowledge, our study is the first and largest analysis specifically led to evaluate the association between serum remnant-C levels and NAFLD risk in a longitudinal retrospective cohort. Knowledge of this association is important for perfect health care resource allocation and prevention and management of NAFLD-related diseases, in turn to attenuate the society medical burden.

**Abbreviations:** NAFLD, Non-alcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; CVD, cardiovascular disease; LDL-C, low-density lipoprotein cholesterol; TG, triglycerides; HDL-C, high-density lipoprotein cholesterol; TC, total cholesterol; Remnant-C, remnant cholesterol; TRLs, triglyceride-rich lipoproteins; CM, chylomicrons; VLDL, very low-density lipoproteins; LPL, lipoprotein lipase; CHD, coronary heart disease; IR, insulin resistance; BMI, body mass index; HC, hip circumference; FPG, fasting plasma glucose; SBP, systolic blood pressure, DBP, diastolic blood pressure; FFAs, free fatty acids; MACEs, major adverse cardiovascular events.

# 2 Methods

## 2.1 Ethical approval

The Ethics Committee of Shanghai Jiao Tong University and Shandong Provincial Hospital (NO.2021-323) approved this study

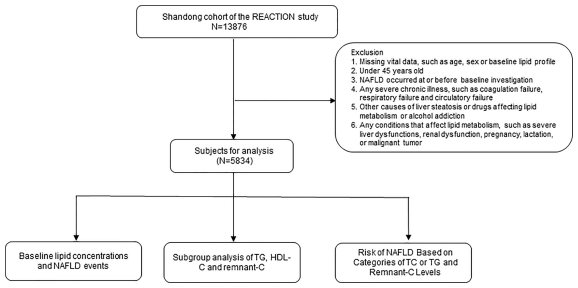



FIGURE 1
The flow chart of our study.

which was conducted with the principles in the Declaration of Helsinki. And the consent obtained from all subjects included in the study was both informed and written.

## 2.2 Subjects

This is a longitudinal study and the data were retrospectively reviewed. All subjects in the current study are from the 13,876 individuals who recruited in the Shandong cohort of the REACTION study, which is a prospective, multicenter, observational cohort study conducted from 2011 and has recruited more than 200, 000 people in China so far (8). We included 6,634 participants who had more than one visit during the study period with an average follow-up time of 43.34 months and assessed for eligibility.

As our main research objective was to assess the association of TG and remnant-C with the outcome of NAFLD, for the present analysis we have excluded individuals with (1) missing vital data, such as age, sex, body mass index (BMI), or baseline lipid profile (2); under 45 years old (3); NAFLD has occurred at or before the baseline investigation; (4) any severe chronic illness, such as coagulation failure, respiratory failure and circulatory failure; (5) other causes of liver steatosis or drugs affecting lipid metabolism or alcohol addiction; and (6) any conditions that affect lipid metabolism, such as severe liver dysfunctions, renal dysfunction, pregnancy, lactation, or malignant tumor (Figure 1). The date of the first recruited participant was April 2011, and the end of the follow-up was July 2017.

## 2.3 Measurements

Data was collected at local health stations by trained investigators to minimize instructor variability. Demographic characteristics were obtained from a well-established questionnaire through a face-to-face interview. All subjects were asked to fast for at least 10 hours before the health examinations. When subjects wore light clothing and took off their shoes, height and weight were measured in kilograms and centimeters, respectively. BMI (kg/m$^2$) was calculated by dividing weight by the square of the height. Waist circumference (WC) was measured at the midpoint between the lower rib cage and the iliac crest in centimeters. Hip circumference (HC) was measured at the widest protrusion of buttocks in centimeters. Using an electronic sphygmomanometer (HEM-7117; Omron, Kyoto, Japan), blood pressure was measured three times after 5 min rest, and the average of the three measurements was calculated.

Blood samples were collected in the morning after a minimum 10-hour fasting. Fasting plasma glucose (FPG) was measured within 2 hours using the glucose oxidase method. Using the VARIANT II Hemoglobin Testing System (Bio-Rad Laboratories), glycated hemoglobin (HbA1c) was measured by high-performance liquid chromatography. After serum and plasma samples were separated and then shipped by air to the Clinical Laboratory for Endocrinology, Shanghai Institute of Endocrine and Metabolic Diseases, the lipid profile measurements including TG, TC, LDL-C, and HDL-C were performed using an autoanalyzer (ADVIA-1650 Chemistry System, Bayer, Leverkusen, Germany). Remnant-C was estimated as TC minus LDL-C minus HDL-C.

Diabetes was defined as fasting blood glucose ≥7.0 mmol/L and/or HbA1c ≥6.2% and/or taking glucose- lowering medication and/or self- report of diabetes (9). The CHD definition and diagnostic criteria are included in the previous guidelines (10). In this study, CHD mainly refers to diagnosed angina pectoris, myocardial infarction, heart failure or coronary heart disease.

## 2.4 Outcome ascertainment

The primary outcome was NAFLD status at the end of the follow-up. The sources of information to identify outcome were yearly revisions of medical records by trained investigators and clinical technicians. All medical records related to outcome were evaluated by the outcome adjudication committee. As described by the Chinese Liver Disease Association, NAFLD was diagnosed by ultrasound (US) (2). In brief, the definition of NAFLD was a diffusion-enhanced near-field echo in the liver region and gradual decay of the far-field echo with one of the following conditions: mild to moderate hepatomegaly with peripheral and marginal passivation; the structure of the hepatic lacunae cannot clearly displayed; the blood flow distribution was normal, but the blood flow signal was reduced; or the unclear or incomplete right liver lobe and diaphragm muscle capsule (11).

## 2.5 Statistical analysis

The Kolmogorov-Smirnov Test was used to test the normality of all variables prior to performing parametric tests. Normally distributed continuous parameters were represented as the mean ± SD, while nonnormally distributed continuous variables were represented as the medians with interquartile ranges. Categorical variables were summarized as numbers (percentage). To analyze differences of remnant-C distribution at baseline between the sex, diabetes status, CHD status and BMI categories groups, data were tested by the nonparametric test.

TABLE 1 Description of study subjects.

| | Total | Female | Male | *p* |
|---|---|---|---|---|
| N | 5834 | 3555 | 2279 | |
| Age, yrs | 56.94 ± 7.38 | 56.43 ± 7.24 | 57.74 ± 7.51 | <0.001 |
| HC, cm | 97.00 (91.00,102.50) | 97.00 (91.00,102.00) | 97.00 (92.00,103.00) | 0.086 |
| BMI, kg/m$^2$ | 25.08 (22.76,27.46) | 25.20 (22.81,27.59) | 24.96 (22.64,27.34) | 0.024 |
| SBP, mmHg | 138 (125,153) | 136 (123,152) | 141 (128,155) | <0.001 |
| DBP, mmHg | 81 (74,89) | 79 (72,87) | 84 (76,92) | <0.001 |
| FPG, mmol/L | 5.83 (5.40,6.50) | 5.77 (5.36,6.39) | 5.94 (5.49,6.65) | <0.001 |
| Diabetes | 1701 (29.16) | 988 (27.79) | 713 (31.29) | 0.004 |
| CHD | 353 (6.05) | 219 (6.16) | 134 (5.88) | 0.661 |

Values are mean ± SD or median (IQR: interquartile range) and n. (%). HC, Hip circumference; BMI, Body mass index; SBP, Systolic blood pressure; DBP, Diastolic blood pressure; FPG, Fasting plasma glucose; CHD, Coronary heart disease.

Follow-up time was calculated as the interval between the date of recruitment in the study and the date of the incident NAFLD, the date of the last visit, or the last record of the deceased subjects while he or she was alive. The association between lipid concentrations (either as continuous or categorical variables) and incident NAFLD were evaluated by unadjusted and adjusted Cox proportional hazard models. The potential confounders that may affect the association between lipid concentrations and incident NAFLD were all adjusted in the Cox proportional hazard models, including age, sex, HC, BMI, systolic blood pressure (SBP), diastolic blood pressure (DBP), FPG, diabetes status and CHD status (12, 13). All *p* values were two-tailed, and *p* values less than 0.05 were considered statistically significant. SPSS version 25.0 (SPSS, Chicago, IL, USA) and R (version 4.1.1) was used for all parametric tests.

# 3 Results

## 3.1 Description of study subjects

Table 1 lists the baseline characteristics of the subjects (N=5,834) in the current study. The participants' average age was 56.94 years, 39.06% were men, median BMI was 25.08 kg/m$^2$, median HC was 97cm. Diabetes and CHD were present in 29.16% and 6.05% of participants, respectively.

## 3.2 Baseline lipid profile of study subjects

Table 2 depicts the lipid profile at baseline of study subjects. Median TC and TG was 5.12 and 1.19 mmol/L, respectively. The lipid alterations characteristic of atherogenic dyslipidemia (TG>1.69 mmol/L and HDL-C<1.03 mmol/L in men or<1.29 mmol/L in women) are present in 9.65% of the baseline population.

Median remnant-C was 0.58 mmol/L (Table 3), and its distribution was differed by BMI categories (<24 kg/m$^2$:0.53 (0.34,0.75) mmol/L; 24 to 28 kg/m$^2$: 0.60(0.40,0.88)mmol/L; ≥28 kg/m$^2$: 0.63(0.41,0.96) mmol/L), CHD status (no CHD: 0.57 (0.37,0.84) mmol/L; CHD:0.62(0.41,0.89) mmol/L) and diabetes status (no diabetes:0.56(0.36,0.79) mmol/L; diabetes:0.64 (0.41,0.97) mmol/L) (Figure 2).

## 3.3 Baseline lipid concentrations and NAFLD events

After adjusting for age, sex, HC, BMI, SBP, DBP, FPG, CHD and diabetes, the serum concentrations of TG were associated with

TABLE 2 Baseline lipid profile of study subjects.

| | All Participants | Female | Male | *p* |
|---|---|---|---|---|
| TC, mmol/L | 5.12 (4.41,5.88) | 5.20 (4.48,5.98) | 5.00 (4.29,5.75) | <0.001 |
| TG, mmol/L | 1.19 (0.86,1.72) | 1.18 (0.86,1.73) | 1.19 (0.85,1.70) | 0.801 |
| HDL-C, mmol/L | 1.40 (1.19,1.64) | 1.43 (1.22,1.67) | 1.34 (1.15,1.60) | <0.001 |
| LDL-C, mmol/L | 3.00 (2.44,3.63) | 3.02 (2.47,3.67) | 2.96 (2.4,3.57) | <0.001 |
| TG/HDL | 0.86 (0.57,1.35) | 0.83 (0.57,1.33) | 0.90 (0.58,1.39) | 0.001 |
| TG>1.69 mmol/l +HDL-C<1.03/1.29 mmol/L (in men/women) | 563 (9.65) | 454(12.77) | 109 (4.78) | <0.001 |

Values are median (IQR: interquartile range) or n (%). TC, Total cholesterol; TG, Triglycerides; HDL-C, High-density lipoprotein cholesterol; LDL-C, Low-density lipoprotein cholesterol.

10.3389/fendo.2023.1098078

TABLE 3 Baseline remnant-C and categories of TC and remnant-C levels of study subjects.

| | All Participants | Female | Male | *p* |
|---|---|---|---|---|
| Remnant-C, mmol/L | 0.58 (0.38,0.84) | 0.59 (0.39,0.86) | 0.56 (0.36,0.81) | <0.001 |
| **TC and remnant-C groups** | | | | <0.001 |
| TC≤ 6.2mmol/L& remnant-C≤ 0.58mmol/L | 2696 (46.21) | 1561 (43.91) | 1135 (49.80) | |
| TC≤ 6.2mmol/L& remnant-C>0.58mmol/L | 2146 (36.78) | 1309 (36.82) | 837 (36.73) | |
| TC>6.2mmol/L& remnant-C≤ 0.58mmol/L | 287 (4.92) | 196 (5.51) | 91 (3.99) | |
| TC>6.2mmol/L& remnant-C>0.58mmol/L | 705 (12.08) | 489 (13.76) | 216 (9.48) | |
| **TG and remnant-C groups** | | | | 0.004 |
| TG≤ 1.7mmol/L& remnant-C≤ 0.58mmol/L | 2833 (48.56) | 1665 (46.84) | 1168 (51.25) | |
| TG≤ 1.7mmol/L& remnant-C>0.58mmol/L | 1524 (26.12) | 981 (27.59) | 543 (23.83) | |
| TG>1.7mmol/L& remnant-C≤ 0.58mmol/L | 150 (2.57) | 92 (2.59) | 58 (2.54) | |
| TG>1.7mmol/L& remnant-C>0.58mmol/L | 1327 (22.75) | 817 (22.98) | 510 (22.38) | |

Values are median (IQR: interquartile range) or n (%). Remnant-C, Remnant-cholesterol; TC, Total cholesterol; TG, Triglycerides.

an 8.0% higher risk of NAFLD per every 1-SD increase, whereas serum remnant-C was associated with a 14.3% higher risk per every 1-SD increase (Table 4). Conversely, for every 1-SD increase in HDL-C, the risk of NAFLD decreased 42.9% (HR = 0.571, 95% CI:0.487-0.670, $p< 0.001$). The lipid alterations characteristic of atherogenic dyslipidemia (TG>1.69 mmol/L and HDL-C<1.03 mmol/L in men or<1.29 mmol/L in women) were also associated with a 34.3% higher risk of the incident NAFLD. No significant interactions were observed between TC ($p$=0.371) or LDL-C ($p$=0.597) and the risk of NAFLD.

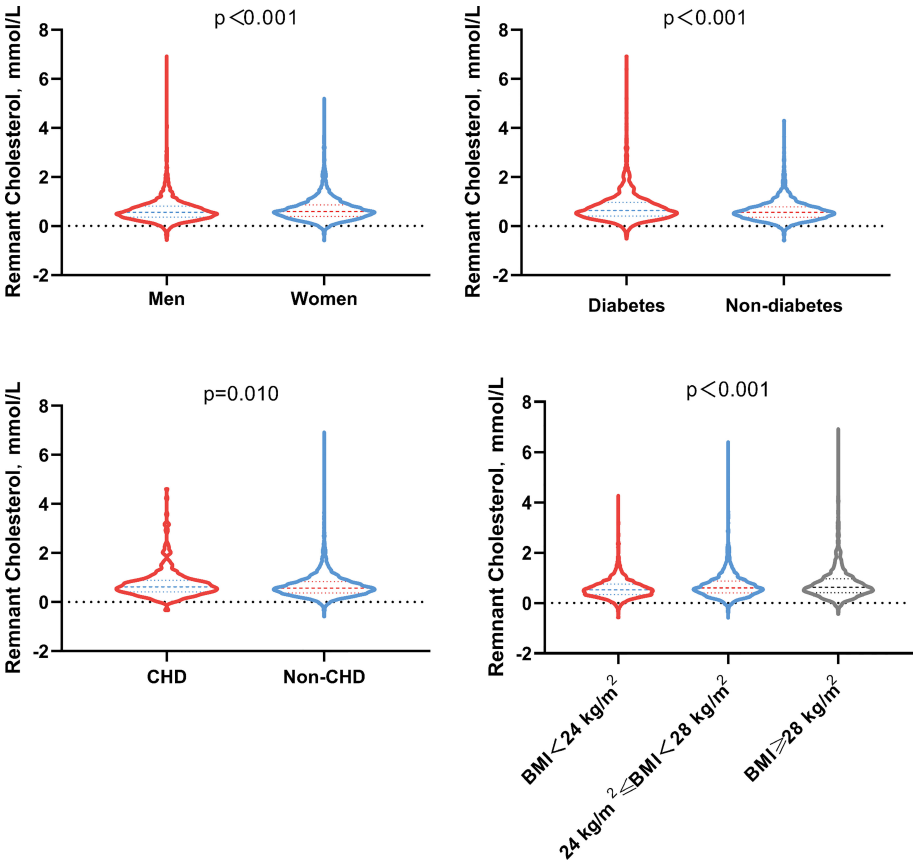



FIGURE 2
Remnant cholesterol distribution at baseline by sex, diabetes tatus, CHD status and BMI (kg/m$^2$) categories. BMI, Body mass index; CHD, Coronary heart disease.

10.3389/fendo.2023.1098078

TABLE 4 Association of baseline lipid values with NAFLD outcomes.

| | No Event (n=4,165) | Event (n=1,669) | Hazard Ratio (95% CI) | *p* |
|---|---|---|---|---|
| TG, mmol/L | 1.10 (0.81,1.57) | 1.43 (1.01,2.16) | 1.080 (1.047-1.113) | <0.001 |
| HDL-C, mmol/L | 1.43 (1.22,1.69) | 1.33 (1.12,1.53) | 0.571 (0.487-0.670) | <0.001 |
| LDL-C, mmol/L | 2.98 (2.43,3.59) | 3.06 (2.49,3.72) | 0.985 (0.931-1.042) | 0.597 |
| TC, mmol/L | 5.10 (4.40,5.84) | 5.21 (4.45,5.98) | 0.980 (0.938- 1.024) | 0.371 |
| Remnant-C, mmol/L | 0.56 (0.36,0.80) | 0.63 (0.42,0.94) | 1.143 (1.052-1.242) | 0.002 |
| TG/HDL | 0.77 (0.52,1.19) | 1.07 (0.73,1.76) | 1.072 (1.042,1.103) | <0.001 |
| TG>1.69 mmol/L +HDL-C< 1.03/1.29 mmol/L (in men/women) | 262 (6.29) | 301 (18.03) | 1.343 (1.177-1.533) | <0.001 |

Values are median (IQR: interquartile range) or n (%), unless otherwise indicated. Hazard ratios (HRs) were estimated by Cox proportional hazards regression models adjusted for age, sex, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, coronary heart disease and diabetes. The "TG", "HDL-C", "LDL-C", "TC", "Remnant-C", "TG/HDL" and the "TG>1.69 mmol/L +HDL-C< 1.03/1.29 mmol/L (in men/women)" are included in the Cox model, respectively. CI, confidence interval; other abbreviations as in Table 2.

## 3.4 Subgroup analysis of TG, HDL-C and remnant-C

In order to better understand the risk factors in the lipid profile that may affect NAFLD incidence and to further identify potential information, subgroup analysis of TG, HDL-C and remnant-C was performed. As the serum concentrations of TG increase, the risk of the incident NAFLD increases. On the contrary, the risk of NAFLD decreases as the HDL-C levels increase. Particularly, the incidence of NAFLD was high in subjects in the upper quartiles of the

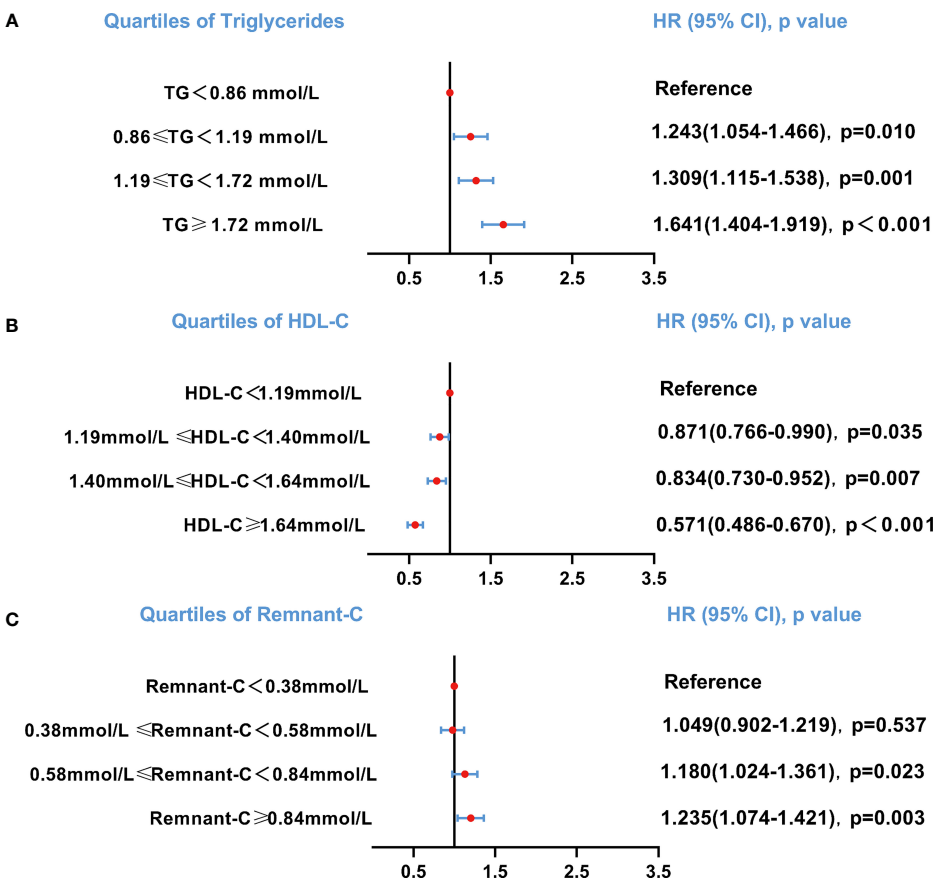



FIGURE 3
Risk of NAFLD across quartiles of baseline lipid parameters. To assess the risk of Non-alcoholic fatty liver disease (NAFLD) associated with the baseline lipid values of interest, we calculated hazard ratios (HR) for the second, third, and fourth quartiles (compared with the first quartile) of levels of **(A)** Triglycerides (TG), **(B)** High-density lipoprotein cholesterol (HDL-c), and**(C)** Remnant cholesterol (Remnant-C). Analyses were adjusted for age, sex, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, cardiovascular disease and diabetes. CI, confidence interval.

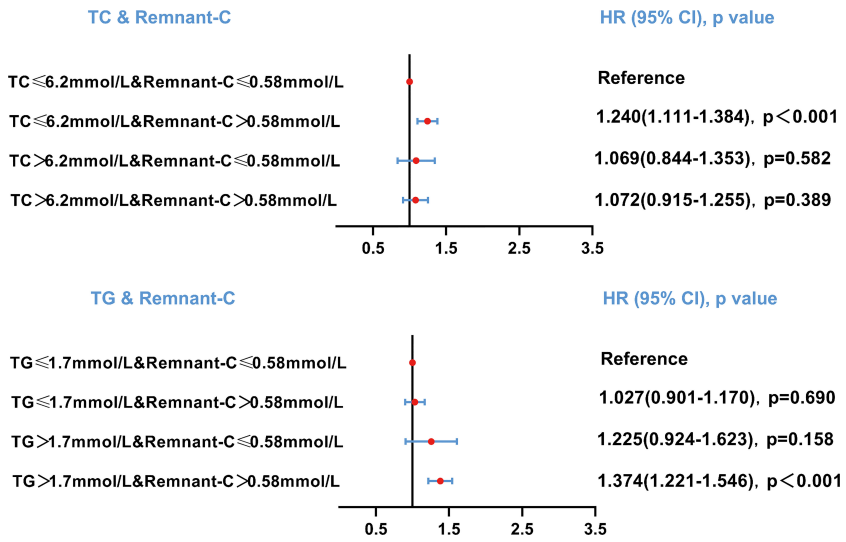



FIGURE 4
Risk of NAFLD based on categories of TC or TG and remnant-C levels. Data were adjusted for age, sex, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, cardiovascular disease and diabetes. HR, hazard ratio; CI, confidence interval; Remnant-C, Remnant-cholesterol; TC, Total cholesterol; TG, Triglycerides.

remnant-C compared with the lowest quartiles (HR = 1.235, 95% CI:1.074-1.421, $p$=0.003) (Figure 3).

## 3.5 Risk of NAFLD based on categories of TC or TG and remnant-C levels

As the residues rich in cholesterol esters and one of the components of serum TC, the relationship between the different combinations of remnant-C and TC levels and the occurrence of NAFLD is worth exploring. The abnormally high levels of remnant-C were defined to the remnant-C>50th percentile of the cohort (0.58mmol/L). Conventionally, high levels for TC were defined as >6.2mmol/L. When TC values ≤ 6.2mmol/L, high baseline remnant-C identified subjects are at a higher risk of NAFLD compared with those at lower concentrations (Figure 4). When TG values > 1.7mmol/L, high baseline remnant-C identified subjects are at a higher risk of NAFLD compared with those at lower concentrations (Figure 4). Data were adjusted for age, sex, HC, BMI, SBP, DBP, FPG, CHD and diabetes. Cumulative hazard curve was constructed to assess the incidence of NAFLD by categories of low and high TC or TG and remnant-C. NAFLD incidence was low in the low remnant-C groups, regardless of TC levels (Figure 5).

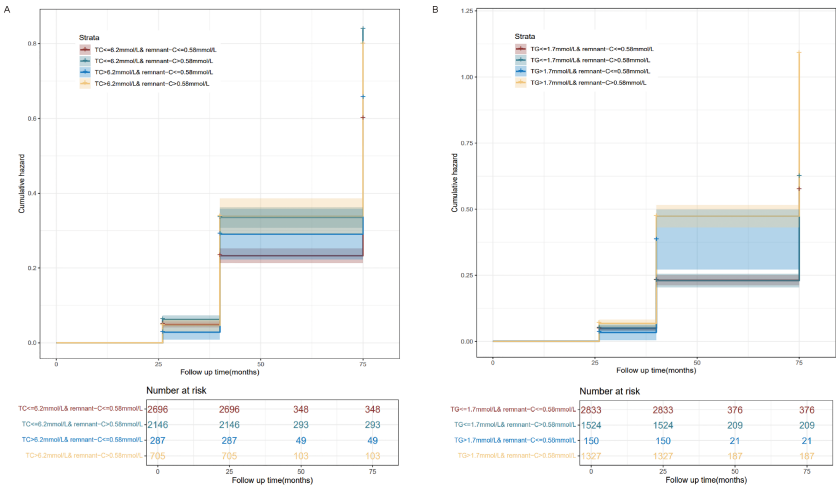



FIGURE 5
Incidence curves of NAFLD based on pre-defined categories of TC or TG and remnant-C levels. **(A)** Pre-defined categories of TC and remnant-C levels. **(B)** Pre-defined categories of TG and remnant-C levels. Remnant-C, Remnant-cholesterol; TC, Total cholesterol; TG, Triglycerides.

### 3.6 Sensitivity analysis

The distributions of remnant-C were differed by sex groups, BMI categories, cardiovascular disease status and diabetes status. Next, we supplemented some subgroup analyses for these factors to explore the robust associations of remnant-C or other lipids with NAFLD. After adjusting for other factors in the Cox regression models, we found that serum levels of TG and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes in women group (Supplementary Figure 1), non-cardiovascular disease status (Supplementary Figure 2), non-diabetes status (Supplementary Figure 3) and middle BMI categories (24 to 28 kg/m$^2$) (Supplementary Figure 4).

## 4 Discussion

Among the middle aged and elderly subset of the Chinese population in this longitudinal retrospective cohort study, the main findings were that serum levels of TG and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes, independent of other risk factors, regardless of age, sex, HC, BMI, SBP, DBP, FPG, CHD and diabetes. To the best of our knowledge, this is the first epidemiological and longitudinal study reporting the association between serum remnant-C and the risk of NAFLD outcomes.

Methods to assess risk factors for NAFLD may include autopsy studies, hospital population-based studies, and community population-based screening studies. Autopsy can easily provide evidence of NAFLD, but the main limitation of autopsy research is that it cannot provide the true prevalence of NAFLD in living people. Because NAFLD patients may be asymptomatic and often do not go to the hospital, studies based on the hospital population cannot provide true risk factors for NAFLD. In contrast, a community-based population-based screening study of random samples can more objectively assess the true risk factors of NAFLD. In the diagnosis of NAFLD, US has a sensitivity of 80%-95% and a specificity of 90%-95% and has a wide range of application value in the screening of NAFLD in the general population (14–16). In addition, reports from Japan showed that NAFLD was more prevalent in the middle-aged subjects (17). Hence, using a community US survey, we determined the risk factors of NAFLD in a Chinese cohort of middle aged and elderly individuals.

Our longitudinal cohort results confirmed previous cross-sectional evidence on the risk role of TG in NAFLD, that is, the increase concentrations of serum TG were associated with higher risk of NAFLD as time goes by. In the characterization of the pathogenic mechanisms of NAFLD, the 'first hit' is triggered by the lipid accumulation in hepatocytes, a trait in which exacerbated fat intake and IR would play a key role (18). Although NAFLD is associated with excess triglycerides in the liver, current evidence suggests that free fatty acids (FFAs), not TG, accumulate in lipid droplets to cause inflammatory liver damage in nonalcoholic steatohepatitis. The liver metabolism of FFAs leads to the formation of toxic metabolites, which are mainly responsible for the production of oxidative stress, inflammation and liver parenchymal damage (19–21). However, the accumulation of TG in the liver is currently considered to be a non-toxic and safer form of liver lipid storage, which is an epiphenomenon that reflects changes in the balance of hepatocyte FFA flux and cellular stress (22); therefore, steatosis can be recognized as an early adaptive response to hepatocyte stress as a result of increased caloric consumption. Through this process, potentially lipotoxic FFAs are segmented into relatively inert intracellular TG molecules (23). Many studies have found that IR is the most important and common potential factor involved in the accumulation of free fatty acids in the liver. The current dominant paradigm is that IR leads to dyslipidemia. Our results show that serum TG concentration is a risk factor for NAFLD over time. So, is hypertriglyceridemia triggering IR or IR causing changes in serum TG concentration in the occurrence of NAFLD? It is important to determine whether hypertriglyceridemia plays a causal role in the etiology of insulin resistance in NAFLD since it can reveal new avenues to combat NAFLD.

Remnant-C is the residue of the TRLs metabolism that consists of chylomicron remnants in the non-fasting state and VLDL and intermediate density lipoproteins in the fasting state. Previous evidence depicted remnant-C was associated with the increased risk of major adverse cardiovascular events (MACEs), but the relationship between serum remnant-C level and the occurrence of NAFLD has not been studied longitudinally (24). The study from Italian hospitals including 798 unselected patients with cardio-metabolic diseases and 79.2% with the presence of NAFLD showed that there was a correlation between values of the circulating remnant-C levels and liver disease severity in patients with NAFLD (25). Consistent with this, teenagers with high remnant-C levels had more severe fat accumulation in their livers compared to those with low remnant-C levels in the Raine Study (26).Particularly, in our longitudinal retrospective cohort study, we found a significant correlation between levels of serum remnant-C and the incidence of NAFLD after adjusting for age, sex, HC, BMI, SBP, DBP, FPG, CHD and diabetes. In other words, serum remnant-C was associated with a 14.3% higher risk per every 1-SD increase and the incidence of NAFLD was high in subjects in the upper quartiles of the remnant-C compared with the lowest quartiles (HR = 1.235, 95%CI:1.074-1.421, $p$=0.003), which can expand our knowledge of the remnant-C as a possible risk factor of NAFLD to some extent.

Atherogenic dyslipidemia, characterized by plasma hypertriglyceridemia, increased small dense LDL particles, and decreased serum HDL-C concentration, is often present in a wide range of chronic cardio-metabolic disorders within the NAFLD, overweight, obesity and diabetes and considered as one of the main causes of lipid-dependent residual risk, regardless of LDL-C concentration (27–29). It's worth noting that we also found the lipid alterations characteristic of atherogenic dyslipidemia were associated with a 34.3% higher risk of the incident NAFLD. Aside from IR, several other factors include diet composition, gut microbiota and genetic factors also contribute to the pathogenesis of atherogenic dyslipidemia in patients with NAFLD (30).

Furthermore, remnant-C, but not TC, was the major cholesterol fraction contributor to NAFLD in our cohort of participants who had no previous NAFLD. No significant interactions were observed between TC and the risk of NAFLD. Cumulative hazard curve, constructed to assess the incidence of NAFLD by categories of low and high TC and remnant-C, found that NAFLD incidence was low in the low remnant-C groups, regardless of TC levels. It is critical to identify potentially modifiable risk factors of NAFLD is of importance, so as to help develop targeted therapies that decrease the risk of NAFLD. In addition, consistent with the study by Catanzaro R et al. (31), we demonstrated that higher TG/HDL-C ratio is associated with NAFLD, so TG/HDL-C could be used as a reliable non-invasive marker in diagnostics of NAFLD in the future. Notably, our subgroup analyses found that serum levels of TG and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes in women group, non-cardiovascular disease status, non-diabetes status and middle BMI categories (24 to 28 kg/m$^2$). Therefore, more TG and remnant-C monitoring should be given to individuals with the above characteristics for early prevention and intervention in the occurrence and development of NAFLD.

Recently, the relevant data of Guideline for the Management of Diabetes Mellitus in the Elderly in China (2021 edition) (32) show that the prevalence of diabetes in the elderly in China is 30.2%, which is much higher than the 19.3% of diabetes prevalence in the elderly in the world. The number of patients suffering from diabetes in China reached 35.5 million, accounting for a quarter of the world's elderly diabetes and ranking first in the world. Our population fits this prevalence rate, so the results of the study are of great significance.

The study is not without its limitations. First, our research was observational and the causal role of remnant-C on the risk of NAFLD incident should be verified in further studies. Second, since we focused on plasma lipid related levels, we collected data on a patient's BMI, blood pressure, and other relevant indicators. We failed to measure or collect data on diet and physical exercise, which are very important influencing factors for NAFLD. And the lack of data on plasma insulin that could drive dyslipidemia or NAFLD may influencing the presented data. Third, the value of remnant-C in our study might have been overestimated by indirect calculation in comparison to direct measurement and more complicated and expensive measurement of remnant-C could be required for accurate results in vulnerable patients.

## 5 Conclusions

In summary, our study identified levels of TG and remnant-C, but not TC or LDL-C, were associated with NAFLD outcomes independent of other risk factors in the middle aged and elderly subset of the Chinese population, especially in those who were women, non-cardiovascular disease status, non-diabetes status and middle BMI status (24 to 28 kg/m$^2$). Consequently, the demonstration of an association between TG or remnant-C and NAFLD in those individuals could aid in the identification of subjects who might benefit from targeted risk factor assessment and management before the occurrence of adverse NAFLD outcomes.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Requests for access should be directed to Chao Xu, at doctorxuchao@163.com.

## Author contributions

YC, CX and LG design the study. YC and QZ performed statistical analyses. FY, HL and GZ contributed to the critical revision of the manuscript. FY, PS and XZ contributed to the statical review. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Natural Science Foundation (No.81974124) and the Taishan Scholar Project (No. tsqn20161071).

## Acknowledgments

We kindly acknowledge the participants and researchers who participated in this study.

## 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.2023.1098078/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Sensitivity analysis for the associations of remnant-C or other lipids with incident NAFLD in sex groups. Data were adjusted for age, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, cardiovascular disease and diabetes status. HR, hazard ratio; CI, confidence interval; other abbreviations as in Table 2.

SUPPLEMENTARY FIGURE 2
Sensitivity analysis for the associations of remnant-C or other lipids with incident NAFLD in different diabetes status. Data were adjusted for age, sex, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose and cardiovascular disease status. HR, hazard ratio; CI, confidence interval; other abbreviations as in Table 2.

SUPPLEMENTARY FIGURE 3
Sensitivity analysis for the associations of remnant-C or other lipids with incident NAFLD in different cardiovascular disease status. Data were adjusted for age, sex, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose and diabetes status. HR, hazard ratio; CI, confidence interval; other abbreviations as in Table 2.

SUPPLEMENTARY FIGURE 4
Sensitivity analysis for the associations of remnant-C or other lipids with incident NAFLD in different body mass index categories. Data were adjusted for age, sex, hip circumference, body mass index, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, cardiovascular disease and diabetes status. HR, hazard ratio; CI, confidence interval; BMI, body mass index; other abbreviations as in Table 2.

# References


1. Sîrbu O, Floria M, Dăscăliţa P, Şorodoc V, Şorodoc L. Non-alcoholic fatty liver disease-from the cardiologist perspective. *Anatolian J Cardiol* (2016) 16(7):534–41. doi: 10.14744/AnatolJCardiol.2016.7049

2. Sun DQ, Liu WY, Wu SJ, Zhu GQ, Braddock M, Zhang DC, et al. Increased levels of low-density lipoprotein cholesterol within the normal range as a risk factor for nonalcoholic fatty liver disease. *Oncotarget* (2016) 7(5):5728–37. doi: 10.18632/oncotarget.6799

3. Chalasani N, Deeg MA, Crabb DW. Systemic levels of lipid peroxidation and its metabolic and dietary correlates in patients with nonalcoholic steatohepatitis. *Am J gastroenterol* (2004) 99(8):1497–502. doi: 10.1111/j.1572-0241.2004.30159.x

4. Norris AL, Steinberger J, Steffen LM, Metzig AM, Schwarzenberg SJ, Kelly AS. Circulating oxidized LDL and inflammation in extreme pediatric obesity. *Obes (Silver Spring Md).* (2011) 19(7):1415–9. doi: 10.1038/oby.2011.21

5. Alkassabany YM, Farghaly AG, El-Ghitany EM. Prevalence, risk factors, and predictors of nonalcoholic fatty liver disease among schoolchildren: a hospital-based study in Alexandria, Egypt. *Arab J* (2014) 15(2):76–81. doi: 10.1016/j.ajg.2014.05.002

6. DeFilippis AP, Blaha MJ, Martin SS, Reed RM, Jones SR, Nasir K, et al. Nonalcoholic fatty liver disease and serum lipoproteins: the multi-ethnic study of atherosclerosis. *Atherosclerosis* (2013) 227(2):429–36. doi: 10.1016/j.atherosclerosis.2013.01.022

7. Varbo A, Nordestgaard BG. Remnant lipoproteins. *Curr Opin lipidol* (2017) 28 (4):300–7. doi: 10.1097/MOL.0000000000000429

8. Bi Y, Lu J, Wang W, Mu Y, Zhao J, Liu C, et al. Cohort profile: risk evaluation of cancers in Chinese diabetic individuals: a longitudinal (REACTION) study. *J diab* (2014) 6(2):147–57. doi: 10.1111/1753-0407.12108

9. Xu C, Zhou G, Zhao M, Zhang X, Fang L, Guan Q, et al. Bidirectional temporal relationship between obesity and hyperinsulinemia: longitudinal observation from a Chinese cohort. *BMJ Open Diabetes Res Care* (2021) 9(1):e002059. doi: 10.1136/bmjdrc-2020-002059

10. Fihn SD, Blankenship JC, Alexander KP, Bittl JA, Byrne JG, Fletcher BJ, et al. ACC/AHA/AATS/PCNA/SCAI/STS focused update of the guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American college of Cardiology/American heart association task force on practice guidelines, and the American association for thoracic surgery, preventive cardiovascular nurses association, society for cardiovascular angiography and interventions, and society of thoracic surgeons. *J Am Coll Cardiol* (2014) 64 (18):1929–49. doi: 10.1016/j.jacc.2014.07.017

11. Zeng MD, Li YM, Chen CW, Lu LG, Fan JG, Wang BY, et al. Guidelines for the diagnosis and treatment of alcoholic liver disease. *J digest diseases* (2008) 9(2):113–6. doi: 10.1111/j.1751-2980.2008.00332.x

12. Kasper P, Martin A, Lang S, Demir M, Steffen HM. Hypertension in NAFLD: An uncontrolled burden. *J hepatol* (2021) 74(5):1258–60. doi: 10.1016/j.jhep.2021.01.019

13. Tanase DM, Gosav EM, Costea CF, Ciocoiu M, Lacatusu CM, Maranduca MA, et al. The intricate relationship between type 2 diabetes mellitus (T2DM), insulin resistance (IR), and nonalcoholic fatty liver disease (NAFLD). *J Diabetes Res* (2020) 2020:3920196. doi: 10.1155/2020/3920196

14. Joseph AE, Saverymuttu SH, al-Sam S, Cook MG, Maxwell JD. Comparison of liver histology with ultrasonography in assessing diffuse parenchymal liver disease. *Clin radiol* (1991) 43(1):26–31. doi: 10.1016/S0009-9260(05)80350-2

15. Mathiesen UL, Franzén LE, Aselius H, Resjö M, Jacobsson L, Foberg U, et al. Increased liver echogenicity at ultrasound examination reflects degree of steatosis but not of fibrosis in asymptomatic patients with mild/moderate abnormalities of liver transaminases. *Digest liver Dis* (2002) 34(7):516–22. doi: 10.1016/S1590-8658(02)80111-6

16. Layer G, Zuna I, Lorenz A, Zerban H, Haberkorn U, Bannasch P, et al. Computerized ultrasound b-scan texture analysis of experimental diffuse parenchymal liver disease: correlation with histopathology and tissue composition. *J Clin ultrasound JCU.* (1991) 19(4):193–201. doi: 10.1002/jcu.1870190402

17. Kojima S, Watanabe N, Numata M, Ogawa T, Matsuzaki S. Increase in the prevalence of fatty liver in Japan over the past 12 years: analysis of clinical background. *J gastroenterol* (2003) 38(10):954–61. doi: 10.1007/s00535-003-1178-8

18. Bessone F, Razori MV, Roma MG. Molecular pathways of nonalcoholic fatty liver disease development and progression. *Cell Mol Life Sci CMLS.* (2019) 76(1):99–128. doi: 10.1007/s00018-018-2947-0

19. Mota M, Banini BA, Cazanave SC, Sanyal AJ. Molecular mechanisms of lipotoxicity and glucotoxicity in nonalcoholic fatty liver disease. *Metabol: Clin experiment* (2016) 65(8):1049–61. doi: 10.1016/j.metabol.2016.02.014

20. Hirsova P, Ibrabim SH, Gores GJ, Malhi H. Lipotoxic lethal and sublethal stress signaling in hepatocytes: relevance to NASH pathogenesis. *J Lipid Res* (2016) 57 (10):1758–70. doi: 10.1194/jlr.R066357

21. Malhi H, Gores GJ. Molecular mechanisms of lipotoxicity in nonalcoholic fatty liver disease. *Semin liver disease.* (2008) 28(4):360–9. doi: 10.1055/s-0028-1091980

22. Cortez-Pinto H, de Moura MC, Day CP. Non-alcoholic steatohepatitis: from cell biology to clinical practice. *J hepatol* (2006) 44(1):197–208. doi: 10.1016/j.jhep.2005.09.002

23. Tilg H, Moschen AR. Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. *Hepatol (Baltimore Md).* (2010) 52 (5):1836–46. doi: 10.1002/hep.24001

24. Sandesara PB, Virani SS, Fazio S, Shapiro MD. The forgotten lipids: Triglycerides, remnant cholesterol, and atherosclerotic cardiovascular disease risk. *Endocr Rev* (2019) 40(2):537–57. doi: 10.1210/er.2018-00184

25. Pastori D, Baratta F, Novo M, Cocomello N, Violi F, Angelico F, et al. Remnant lipoprotein cholesterol and cardiovascular and cerebrovascular events in patients with non-alcoholic fatty liver disease. *J Clin Med* (2018) 7(11):378. doi: 10.3390/jcm7110378

26. Chin J, Mori TA, Adams LA, Beilin LJ, Huang RC, Olynyk JK, et al. Association between remnant lipoprotein cholesterol levels and non-alcoholic fatty liver disease in adolescents. *JHEP Rep Innovation hepatol* (2020) 2(6):100150. doi: 10.1016/j.jhepr.2020.100150

27. Brunzell JD, Davidson M, Furberg CD, Goldberg RB, Howard BV, Stein JH, et al. Lipoprotein management in patients with cardiometabolic risk: consensus conference report from the American diabetes association and the American college of cardiology foundation. *J Am Coll Cardiol* (2008) 51(15):1512–24. doi: 10.1016/j.jacc.2008.02.034

28. Chapman MJ, Ginsberg HN, Amarenco P, Andreotti F, Borén J, Catapano AL, et al. Triglyceride-rich lipoproteins and high-density lipoprotein cholesterol in patients at high risk of cardiovascular disease: evidence and guidance for management. *Eur Heart J* (2011) 32(11):1345–61. doi: 10.1093/eurheartj/ehr112

29. Joshi PH, Khokhar AA, Massaro JM, Lirette ST, Griswold ME, Martin SS, et al. Remnant lipoprotein cholesterol and incident coronary heart disease: The Jackson heart and framingham offspring cohort studies. *J Am Heart Assoc* (2016) 5(5):e002765. doi: 10.1161/JAHA.115.002765

30. Akhtar DH, Iqbal U, Vazquez-Montesino LM, Dennis BB, Ahmed A. Pathogenesis of insulin resistance and atherogenic dyslipidemia in nonalcoholic fatty liver disease. *J Clin Trans hepatol* (2019) 7(4):362–70. doi: 10.14218/JCTH.2019.00028

31. Catanzaro R, Selvaggio F, Sciuto M, Zanoli L, Yazdani A, He F, et al. Triglycerides to high-density lipoprotein cholesterol ratio for diagnosing nonalcoholic fatty liver disease. *Minerva gastroenterol* (2022) 68(3):261–8. doi: 10.23736/S2724-5985.21.02818-X

32. National Center of Gerontology CSoG and Diabetes Professional Committee of Chinese Aging Well Association. Guideline for the management of diabetes mellitus in the elderly in China (2021 edition). *Chin J Diabetes Mellitus* (2021) 13(1):14–46.

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 15 May 2023
DOI 10.3389/fendo.2023.1157416



OPEN ACCESS

EDITED BY
Haifeng Hou,
Shandong First Medical University, China

REVIEWED BY
Rachelle Elizabeth Irwin,
Ulster University, United Kingdom
Sinan Tanyolac,
Istanbul University, Türkiye

*CORRESPONDENCE
Michael A. Zulyniak
m.a.zulyniak@leeds.ac.uk

RECEIVED 02 February 2023
ACCEPTED 01 May 2023
PUBLISHED 15 May 2023

CITATION
Fuller H, Iles MM, Moore JB and Zulyniak MA (2023) Metabolic drivers of dysglycemia in pregnancy: ethnic-specific GWAS of 146 metabolites and 1-sample Mendelian randomization analyses in a UK multi-ethnic birth cohort.
*Front. Endocrinol.* 14:1157416.
doi: 10.3389/fendo.2023.1157416



# Metabolic drivers of dysglycemia in pregnancy: ethnic-specific GWAS of 146 metabolites and 1-sample Mendelian randomization analyses in a UK multi-ethnic birth cohort

Harriett Fuller[1,2], Mark M. Iles[3,4], J. Bernadette Moore[1] and Michael A. Zulyniak[1]*

[1]School of Food Science and Nutrition, University of Leeds, Leeds, United Kingdom, [2]Public Health Science Division, Fred Hutchinson Cancer Center, Seattle, WA, United States, [3]Leeds Institute of Medical Research, University of Leeds, Leeds, United Kingdom, [4]Leeds Institute for Data Analytics, University of Leeds, Leeds, United Kingdom

**Introduction:** Gestational diabetes mellitus (GDM) is the most common pregnancy complication worldwide and is associated with short- and long-term health implications for both mother and child. Prevalence of GDM varies between ethnicities, with South Asians (SAs) experiencing up to three times the risk compared to white Europeans (WEs). Recent evidence suggests that underlying metabolic difference contribute to this disparity, but an investigation of causality is required.

**Methods:** To address this, we paired metabolite and genomic data to evaluate the causal effect of 146 distinct metabolic characteristics on gestational dysglycemia in SAs and WEs. First, we performed 292 GWASs to identify ethnic-specific genetic variants associated with each metabolite ($P \leq 1 \times 10\text{-}5$) in the Born and Bradford cohort (3688 SA and 3354 WE women). Following this, a one-sample Mendelian Randomisation (MR) approach was applied for each metabolite against fasting glucose and 2-hr post glucose at 26-28 weeks gestation. Additional GWAS and MR on 22 composite measures of metabolite classes were also conducted.

**Results:** This study identified 15 novel genome-wide significant (GWS) SNPs associated with tyrosine in the FOXN and SLC13A2 genes and 1 novel GWS SNP (currently in no known gene) associated with acetate in SAs. Using MR approach, 14 metabolites were found to be associated with postprandial glucose in WEs, while in SAs a distinct panel of 11 metabolites were identified. Interestingly, in WEs, cholesterols were the dominant metabolite class driving with dysglycemia, while in SAs saturated fatty acids and total fatty acids were most commonly associated with dysglycemia.

**Discussion:** In summary, we confirm and demonstrate the presence of ethnic-specific causal relationships between metabolites and dysglycemia in mid-pregnancy in a UK population of SA and WE pregnant women. Future work will aim to investigate their biological mechanisms on dysglycemia and translating this work towards ethnically tailored GDM prevention strategies.

KEYWORDS

genetics, GDM, gestational diabetes, South Asian, metabolism, glucose, pregnancy

# 1 Introduction

Pregnancy is accompanied by a period of intense maternal metabolic adaptation to meet the energy demands of the foetus (1–3). Mild maternal insulin resistance (IR) is a natural adaption to prioritise adequate glucose for the growing foetus (4). However, if gestational IR exceeds healthy levels and glycaemia is uncontrolled, moderate IR can progress to gestational diabetes mellitus (GDM) (1, 5). GDM is characterised by persistent maternal and foetal exposure to elevated levels of glucose, and places the mother and offspring at risk during pregnancy (i.e., macrosomia and haemorrhaging) and in later life (i.e. from obesity, type 2 diabetes (T2D) and cardiovascular disease) (6).

Globally, GDM is the most common pregnancy complication, affecting up to one in every seven births; however, its prevalence varies between ethnic groups, with South Asian (SA) women at 3-fold greater risk compared to white European (WE) women, irrespective of BMI and country of residence (5, 7). Furthermore, SA women are more likely to develop T2D in later life following a GDM diagnosis (8). Factors driving this disparity in prevalence are not fully understood but metabolism is thought to play a key role (6), given emerging evidence demonstrating (i) differences in metabolic profiles between GDM and non-GDM pregnancies in an ethnic-specific manner (2, 9); and (ii) that a single dietary strategy to manage GDM across all ethnic groups appears ineffective (10, 11). However, heterogeneity of reported metabolite-GDM associations between studies (due to differences in quantification methods, GDM diagnostic criteria (12, 13), ethnic and cultural groups), as well as residual confounding in observational studies have prevented complete understanding and, moreover, advancement to improved equitable care. In short, the field requires a clear and accurate understanding of ethnic-specific metabolic drivers of gestational dysglycemia to inform appropriate and effective prevention and management strategies across ethnic groups.

Mendelian Randomisation (MR) is an instrumental variable technique that can provide estimates of causal associations between an exposure (such as metabolites) and outcome (such as dysglycemia) (14–16). However, no study has yet used MR to establish the presence of casual associations between metabolites and measures of glycaemia at or before the 28 week of pregnancy in an ethnic-specific manner. The present study aimed to address this using the multi-ethnic Born in Bradford (BiB) cohort to identify ethnic-specific metabolic drivers of GDM.

# 2 Material and methods

## 2.1 Exposure data

BiB is a prospective longitudinal birth cohort that aimed to recruit all mothers receiving maternity care in the Bradford Royal Infirmary between 2007-2010 (17). Bradford, a large city in the north of England, has high levels of deprivation and a large SA population, predominantly of Pakistani ancestry. A total of 12,453 women (mean maternal age 27.8) were recruited, 45% of which were of SA ancestry (17, 18). The study was not pre-registered but (SP622) was approved by Born in Bradford. All participants provided written consent and ethnical approval was obtained from the Bradford Research Ethnics Committee (ref07/H1302/112) (17).

Fasted plasma sample collection and high-throughput metabolite quantification by automated NMR (Nightingale Health©; Helsinki, Finland) has been previously described and validated to a high accuracy (2). Briefly, samples were taken by trained phlebotomists from BiB participants (26-28 weeks' gestation) and were processed in the absence of freeze-thaw cycles within 2.5 hours before storage at -80°C. One hundred and forty-six absolute measures of metabolites were included in the analysis following the removal of metabolites expressed as a percentage or ratio to minimise redundancy. In total, 10 overarching classes of metabolites were included in the analysis: lipoproteins (n=97), amino acids (n=9), apolipoproteins (n=2), cholesterols (n=8), fatty acids (n=8), glycerides and phospholipids (n=8), glycolysis related metabolites (n=4), ketone bodies (n=2), measures of fluid balance and inflammation (n=3) and measures of lipoprotein particle diameter (n=3). A full list of included metabolites can be found in Supplementary Table 1.

## 2.2 Outcome data

All participants were assesed prior to GDM diagnosis and the 28$^{th}$ week of pregnancy. Individuals were diagnosed with GDM if

either their fasting glucose or if 2-hour post-load glucose concentration exceeded 6.1 mmol/L or 7.8 mmol/L following a 75g oral glucose tolerance test (OGTT) (19). The OGTT was performed in the morning following an overnight fast. To maximise power, MR analyses were performed using continuous metabolite values and fasting glucose and 2-hour post glucose. Fasting glucose and 2-hour post glucose values were log normalised prior to analysis.

## 2.3 Metabolite data

Information on metabolite data preparation has been described in full elsewhere (9). In brief, 11,480 blood samples were metabolically profiled from BiB, 54 of which were excluded due to failure of any one of five Nightingale© quality control measures leaving 11,426 samples for imputation. Missing data was imputed *via* multiple imputation using the missMDA package in R (20).

After combining with postprandial glucose data, 3,693 SA and 3,377 individuals whose samples were taken before the 28$^{th}$ week of pregnancy were retained before outlier removal (Supplementary Figure 1). Outliers were removed (those outside of 1.5 x IQR) for each metabolite in each ethnicity separately and metabolite values were normalised by taking the log, square root or normal score transformation (NST) as appropriate following the visual inspection of histograms and QQ plots. Following the removal of outliers, the number of individuals available for GWAS analysis of each metabolite varied (Supplementary Table 2) but was relatively consistent: for SAs the average sample size for each metabolite was 3622 (range 3472-3688), while for WEs it was 3301 (range 3158-3345). Information on gestational age at sample collection and parity was obtained from obstetric records. Ethnicity was self-reported or obtained from primary care records if missing. Individuals of a SA descent other than Pakistani were excluded from the analysis due to the small sample sizes of these populations. Differences in the distribution of continuous variables between ethnic groups were assessed by Mann-Whitney tests, while differences in the distribution of categorical variables were assessed by chi-squared tests. Women of SA ancestry tended to be older than WE women (27.9 ± 0.1 vs 26.7 ± 0.1 years) and were more likely to be overweight/obese (64.5% vs 53.4%), and be on their ≥2$^{nd}$ pregnancy (67.1% vs 48.8%), but were less likely have smoked during pregnancy (2.9% vs 30.9%) (Supplementary Table 3).

## 2.4 Genetic data

Imputed genetic data were obtained from BiB. All samples were genotyped using two chips: the Infinium Global Sequencing Array-24 v.1 (GSA) (~640K SNPs) and the Infinium CoreExome-24 v1.1 BeadChip (~550K SNPs) (21). Genetic data from the Illumina Global Sequencing Array (GSA) and Illumina CoreExome SNPs were combined. Where SNPs were missing in >5% of individuals, they were excluded (21). When evaluating imputed data, the $R^2$ value can be a measure of quality control as it reflects to the estimated proportion of genetic variation maintained in the imputed data. As a result, SNPs with an $R^2$ <0.9 were excluded prior to analysis.

## 2.5 GWAS analysis

Conventionally a GWAS assumes individuals are unrelated and the inclusion of related individuals can potentially lead to spurious associations (22, 23). However, the removal of individuals from the BiB sample with close ancestry would substantially reduce the sample size. In addition, high rates of consanguinity in the SA stratum of the cohort makes relatedness difficult to assess (24). As such, a GWAS mixed linear model association (MLMA) analysis was conducted in PLINK (version 1.9) that allowed for the inclusion of related individuals (23, 25–27). MLMA models include a fixed effect, adjusted covariates, and an additional random effect comprised of a variance-covariance matrix that models the correlation (here relatedness) between individuals to be accounted for (23, 25). GWAS MLMA models were implemented using the *GCTA* (Genome-wide Complex Trait Analysis) command line tool for each metabolite in both ethnicities (28). To increase power, MLMA-loo (leave-one-out) analysis was used, preventing a SNP from being included in both the fixed and random effects concurrently, thereby avoiding double fitting (25). MLMA models also included parity and principal components (PC) 1 and 2 to account for population stratification (Supplementary Figure 2, Supplementary Table 4). Gestational age, which showed little variation, was not included in the modelling (median gestational age SA = 184 days, IQR= 182-186.7, median gestational age WE = 184 days, IQR= 182-187). Genomic inflation factors (λ) were calculated for all models for a range of minor allele frequency (MAF) cut-offs (MAF <0.001, 0.001≤ MAF < 0.005, 0.005 ≤ MAF < 0.01, 0.01 ≤MAF < 0.05, 0.05 ≤ MAF < 0.1, and MAF ≥ 0.1) to minimise data loss while also minimising false positives. λ ≥ 1.1 was considered indicative of genomic inflation (29, 30). A MAF cut-off of <0.05 was the least stringent cut-off found to reduce λ to ~1 meaning this cut off was used in the analysis (Supplementary Table 5).

When a SNP was found to be associated with a metabolite value in only one ethnicity, a fixed effect inverse-variance weighted meta-analysis was implemented to assess the heterogeneity (via the $I^2$ statistic) of identified associations between ethnicities and to see if the SNP retained significance in a larger sample. Meta-analyses were conducted within the command-line tool METAL (31) and supplemented with FUMA (v1.5.2) (32) to investigate SNP function based on their effect on phenotypes.

## 2.6 One-sample MR

### 2.6.1 Genetic instruments

One-Sample MR was conducted for all 146 metabolite values in both ethnic groups using SNPs identified as significant at a genome-wide suggestive level (p-value ≤ 1 x $10^{-5}$) in the GWAS. All variants were also entered into MR-base (33) to test for other reported known associations that may be in horizontal or vertical pleiotropy. Metabolites were grouped into their overall classes and SNPs in

each class were thinned by linkage disequilibrium (LD) ($R^2$<0.2) *via* the NIH LDlink online tool (https://ldlink.nci.nih.gov) reducing the overlap of instruments in each class (34, 35). For individuals of WE ancestry, all European (EUR) and South Asian (SAS) populations in LD link (software that utilises 1000 Genome data) were used to estimate LD due to the expected similarity in their LD structure allowing for an increase sample size and resultant improvement in the accuracy of LD estimates (36). Similarity between 1000 Genome SA samples and BiB samples was assessed using principal components analysis (PCA) in PLINK (version 1.9) because Pakistani samples from BiB originate from a different region of Pakistan (the Mirpur Region) from the 1000 Genome SA samples (26, 27). This is of particular importance in SA as even geographically close populations can have differing allele frequencies due to differing Biraderi ('Brotherhood') membership between population subgroups. Biraderi membership is assigned at birth, is an indicator of male lineage as well as social-occupational status which largely governs partner choice and can result in higher levels of consanguinity in the population (21). PCA plots were created using the *ggplot2* package in R studio (version 4.0.2) (37, 38). No clear separation in SA BiB samples and SA 1000 Genome (1000G) samples was identified indicating that LD estimates obtained from 1000G was suitable for use in BiB (Supplementary Figures 3-5).

### 2.6.2 Analysis

Genetic Risk Scores (GRS) were created in PLINK (version 1.9) for each metabolite with each SNP receiving a weight based on its estimated effect size on the metabolite obtained from the GWAS (39). One-sample MR was then performed by Two-Stage Least Squares regression (TSLS; *ivpack*, ivreg, and *AER* packages in R version 4.0.2) to obtain a causal estimate for the effect of each metabolite value on the log-normalised continuous measures of fasting glucose and 2-hour post glucose following a 75g oral glucose tolerance test (OGTT) (37, 40, 41). Here, the level of a metabolite is regressed on its respective GRS and, subsequently, the outcome is regressed onto these fitted GRS-metabolite values in the second stage. All MR results have been reported according to STROBE-MR guidelines (42).

When significant associations were identified, leave-one-out analysis was performed. For this, SNPs were removed sequentially from the instrument and changes to the effect estimate and F-statistic was assessed. If the exclusion of a SNP was found to alter either the effect estimates or F-statistic (through the visualisation of forest plots) it is possible that the SNP is influencing the outcome *via* an alternative pathway to other SNPs, potentially highlighting a violation of the 2$^{nd}$ or 3$^{rd}$ MR assumption. To further test for violations of these assumptions, included SNPs were searched for in both the Phenoscanner and GWAS Catalogue databases to identify previously identified associations (43–45) with potential confounders in horizontal (i.e., multiple pregnancies, type-1 diabetes, deprivation index, parity) rather than vertical pleiotropy (i.e., along causal pathway, such anthropometrics). In both databases a p-value ≤ 1 x $10^{-5}$ was interpreted as indicative of an additional association. Differences between MR and linear regression results were also evaluated *via* the Wu-Hausman statistic to assess deviation of the instrumental variable estimate from the ordinary least squares (OLS) estimate (46). Deviations in these two measures can indicate either confounding in the OLS estimate (indicating a need for MR) or violations of the MR assumptions due to pleiotropy.

### 2.6.3 *Post-hoc* power analyses

For metabolites that associated with a measure of postprandial glucose in only one ethnic group, *post-hoc* power analyses were performed using the mRnd CNS genomics tool (https://shiny.cnsgenomics.com/mRnd/) to assess whether the absence of an association in the alternate ethnicity was due to limited power (47). Observational and 'true' associations required by the tool were obtained by performing linear regression of the outcome on the metabolite and obtaining unadjusted and adjusted estimates (adjusted for maternal age (years), BMI (continuous), smoking status, multiple pregnancy, parity, and gestational age) respectively. Due to the *post-hoc* nature of this analysis, additional power analyses could be conducted assuming the MR estimate to be the true causal effect in the MR calculation. This analysis was performed in the non-significant population for each metabolite associated in only one ethnicity. If power was found to be adequate (80%) at the 5% level ($\alpha = 0.05$) power was also assessed at the 1% level ($\alpha = 0.01$).

# 3 Results

## 3.1 GWAS of metabolite measures

A total of 6184 SNPs were associated with at least one metabolite in WEs at the suggestive level (1 x $10^{-5}$), with 2616 (42.3%) SNPs being associated with a single metabolite measure. However, no SNPs were identified below the genome-wide significant level (p-value <5 x $10^{-8}$) in WEs.

Fewer SNPs were identified at the suggestive level in SAs, with 3685 SNP-metabolite associations in total, of which 1544 (41.9%) SNPs being associated with only one metabolite measure. SNP associations were identified for 138/146 (94.5%) metabolite exposures in SAs (Supplementary Table 6). No SNP was identified as being associated at the suggestive level in both ethnicities, although shared genomic regions were identified between ethnicities (Supplemental Excel). Using FUMA to investigate SNP function based on their effect on phenotypes in both ancestries (32), of the 85 genetic variants meta-analysed that surpassed suggestive GWAS significance: 54 were intergenic (i.e., between genes), 21 were intronic (i.e., between exons of a gene), 5 were upstream (i.e., within 250 bps before transcription start site), 3 were downstream (i.e., within 500 bps after transcription start site), and 2 (2%) were exonic (i.e., within protein coding region).

To evaluate the possibility of shared genomic predictors of metabolites, a pooled meta-analysis of effect estimates in both ethnicities was performed. For 90 metabolite values, no associations were found to exceed the genome-wide suggestive level ($p<10^{-5}$) following meta-analyses of both ethnicities. SNP associations were identified at the suggestive level for four metabolite measures (concentration of XL-HDL, total lipids in M-VLDL, mean density of VLDL and citrate) despite these differing in direction of effects in SAs and WEs. In addition, 4 SNPs were associated with alanine, despite these SNPs initially being associated with alanine only in the SA population. These SNPs (rs12256633, rs17121228, rs7096521, rs12240368) are all found on chromosome 10 in the receptor gene *SORCS1* and have not been associated with alanine levels previously (48, 49).

## 3.2 MR Results

After LD thinning, genetic instruments were available for all 146 metabolites in WEs and for 136/146 (93.2%) metabolites in SAs. In WEs, 1040 SNPs were retained following LD thinning including 423 (40.67%) that were unique to an individual metabolite. Fewer SNPs were identified in SAs, where 383 SNPS remained after LD thinning, 195 (50.9%) of which were unique to a single metabolite. Only 2.7% of included genetic instruments (4 metabolites) in WEs and 12.5% (9 metabolites) of included genetic instruments in SAs had an F-statistic < 10, indicating that most instruments were at low risk of weak instrument bias (50). The average F-statistic for WEs instruments was 72.4, while in SAs it was considerably lower at 26.7 (Supplementary Figure 6). Screening of genomic predictors using Phenoscanner and GWAS databases did not raise major concerns for horizonal pleiotropy (Supplementary Table 7) but did suggest that modification of anthropometrics is a common pathways by which metabolites elicit their effect on dysglycemia – i.e., vertical pleiotropy. However, where horizontal pleiotropy was a possibility (e.g., cholesterol levels), sensitivity analyses were performed (*see 3.2.1.1 and 3.2.2.1 Sensitivity Analyses*). Using MR-Base, we report that [in agreement with recent GWAS (51)], almost all SNPs included in an instrument have been previously associated with dysglycemia metrices or diabetes (Supplementary Table 8). However, since most evidence of genomics-diabetes associations are sourced from non-SA populations, these results may not accurately reflect genetic associations in SAs.

### 3.2.1 White Europeans

Two metabolite values, leucine and mean density of HDL lipoproteins (HDL_D), were associated with both fasting glucose and 2-hour post glucose (Table 1; Supplementary Figures 7, 8). Specifically, a 1mmol/L increase in blood leucine associated with lower fasting glucose (-0.193 mmol/L, 95% CI -0.069, -0.319) and 2-hour post glucose (-0.443 mmol/L, 95% CI -0.113, -0.774). Likewise, a 1nm increase in mean diameter of HDL associated with lower fasting glucose (-0.082 mmol/L, 95% CI 0.026, 0.138) and 2-hour post glucose (-0.191, 95% CI 0.043, 0.339 mmol/L). No other metabolites were associated with both measures of glucose in WEs.

For fasting glucose, an increase of 1mmol/L total cholesterol in M-HDL (M-HDL-C) and cholesterol esters in M-HDL (M-HDL-CE) were associated with lower fasting glucose measures (-0.189 mmol/L, 95% CI -0.021, -0.358, and -0.327 mmol/L, 95% CI -0.069, -0.586 respectively). For 2-hr post-glucose, 8 metabolite values were positively associated with this (HDLC, HDL2C, HDL3C, triglycerides in XS-VLDL, cholesterol esters in XL-HDL, total concentration of L-HDL, total lipids in L-HDL and cholesterol esters in S-HDL) and one (total concentration of S-LDL) was negatively associated (Table 1). Cholesterol metabolites, measures of total cholesterols in lipoproteins and total cholesterols in lipoproteins were the most common types of metabolite class to be associated with postprandial glucose in WEs, with leucine being the only amino acid identified. Wu-Hausman p-values < 0.05 indicate deviations of the instrumental variable estimate from the OLS estimate (Table 1).

#### 3.2.1.1 Sensitivity analyses

For 6 of 13 metabolite values, leave-out one analyses maintained significance ($P\leq 0.05$) indicating that no individual SNP was driving the identified associations in WEs: leucine, mean diameter of HDL, total lipids in L-HDL, cholesterol esters in S-HDL and cholesterol esters in M-HDL (Supplementary Figure 9). For the remaining 8 metabolites, β values were consistent across leave-one-out analyses although not all associations remained significant. Additionally, for 12/13 metabolites (all but M-HDL-CE), the F-statistic did not substantially differ through the exclusion of individual SNPs from the instruments, which suggests they were not substantially driven by a single SNP (Supplementary Figure 9). The exception to this was cholesterol esters in M-HDL, where the exclusion of rs2138011 or rs739018 increased the F-statistic.

Three of the metabolites (leucine, L-HDL-L, L-HDL-C) that were associated with postprandial glucose in WEs included a SNP that previous studies have associated ($p \leq 1 \times 10^{-5}$) with at least one potential confounder (BMI, hypertension or waist circumference) (Supplementary Table 7). The removal of these SNPs from the instrument did not impact the significance of the associations identified for leucine or L-HDL-L (Table 2). However, for L-HDL-C instrument, the exclusion of two SNPs (rs5576825 and rs6811162) previously associated with a potential confounder (waist circumference and hypertension respectively) resulted in non-significant association between L-HDL-C and 2-hour post glucose. Importantly, for both SNPs, it is conceivable that the confounders could reside on their causal pathway (i.e., vertically pleiotropic, where L-HDL-C effects 2-hr post-prandial glucose through its effect on weight gain) rather than be in horizontal pleiotropy and may, therefore, violate the 2$^{nd}$ MR assumption (50).

### 3.2.2 South Asians

No metabolite was associated with both fasting glucose and 2-hour post glucose in SAs (Supplementary Figures 7, 8). Although, for fasting glucose, a 1 mmol/L increase in either total FAw3 or S-HDL-C was associated with an increase of fasting glucose by 0.432 mmol/L (95% CI 0.063 – 0.798) and 1 mmol/L (95% CI 0.116 –

10.3389/fendo.2023.1157416

TABLE 1 Significant MR results in white Europeans.

| | Class | Metabolite | Outcome | F statistic | β estimate (95% CI) | WuH |
|---|---|---|---|---|---|---|
| Lower Glucose | S-LDL | S-LDL-P | 2-hour post | 41.7 | -1000 (-20, -1984) | 0.017 |
| | Amino Acids | Leucine | Fasting glucose | 67.3 | -0.193 (-0.069, -0.319) | 0.005 |
| | | | 2-hour post | | -0.443 (-0.113, -0.774) | 0.008 |
| | M-HDL | M-HDL-CE | Fasting glucose | 62.4 | -0.327 (-0.069, -0.586) | 0.043 |
| | | M-HDL-C | Fasting glucose | 117 | -0.189 (-0.021, -0.358) | 0.117 |
| Increase Glucose | Lipoprotein Density | HDL_D | Fasting glucose | 131 | 0.082 0.026, 0.138) | 0.004 |
| | | | 2-hour post | | 0.191 (0.043, 0.339) | 0.024 |
| | L-HDL | L-HDL-P | 2-hour post | 108 | 220 (41.3, 397) | 0.02 |
| | | L-HDL_L | 2-hour post | 131 | 0.264 (0.062, 0,464) | 0.014 |
| | | L-HDL-C | 2-hour post | 120 | 0.279 (0.012, 0.544) | 0.048 |
| | Cholesterol | HDL2C | 2-hour post | 103 | 0.288 (0.007, 0.583) | 0.025 |
| | | HDLC | 2-hour post | 90.6 | 0.296 (0.007, 0.583) | 0.047 |
| | | HDL3C | 2-hour post | 66.6 | 1.58 (0.002, 3.15) | 0.074 |
| | XL-HDL | XL-HDL-CE | 2-hour post | 109 | 0.541 (0.079, 1.00) | 0.039 |
| | XSVLDL | XS-VLDL-TG | 2-hour post | 87.8 | 0.841 (0.098, 1.58) | 0.042 |
| | S-HDL | S-HDL-CE | 2-hour post | 41.9 | 1.78 (0.448, 3.11) | 0.007 |

Glucose measures are expressed as mmol/L. HDL_D, mean diameter of HDLs (nm); HDLC, total cholesterol in HDL (mmol/L); HDL2C, total cholesterol in HDL2 (mmol/L); HDL3C, total cholesterol in HDL3 (mmol/L); L-HDL-C, total cholesterols in L-HDL (mmol/L); L-HDL_L, total lipids in L-HDL (mol/L); L-HDL-P, concentration of L-HDL (mol/L); M-HDL-C, total cholesterol in M-HDL (mmol/L); M-HDL-CE, cholesterol esters in M-HDL (mmol/L); S-HDL-CE, cholesterol esters in S-HDL (mmol/L); S-LDL-P, concentration of S-LDL (mol/L); XL-HDL-CE, cholesterol esters in XL-HDL (mmol/L); XS-VLDL-TG, triglycerides in XSVLDL (mmol/L); WuH, Wu-Hausman p-value.

1.882) respectively. No metabolite associated with a decrease in fasting glucose in SAs.

Nine metabolites associated with 2-hour post glucose levels in SAs. Of these, 4 metabolites, LA, FAw6, total lipids in M-VLDL (M-VLDL-L) and total phospholipids in L-HDL (L-HDL-PL), associated with an increase in with 2-hour post glucose, with the largest effect being identified for L-HDL-PL. Specifically, a 1mmol/L increase in L-HDL-PL associated with a 0.692 mmol/L increase (95% CI 0.106 - 1.280) in 2-hour post glucose. A further 5 additional metabolites were associated with a decrease in 2-hour post glucose: concentration of L-LDL (L-LDL-P), total cholesterols in IDL (IDL-C), cholesterol esters in IDL (IDL-CE) concentration, total cholesterols in IDL (IDL-C), total lipids in small S-LDL (S-LDL-L), and total lipids in small S-HDL (S-HDL-L). The largest decrease in 2-hour post glucose was observed for L-LDL-P where a 1mmol/L increase in L-LDL-P associated with a 3.86 mmol/L decrease (95% 0.467 - 7.27) in 2-hour post glucose levels (Table 3).

Fatty acids were the class of metabolites most frequently associated with postprandial glucose in SAs. All three fatty acids (LA, FAw3 and FAw6) associations identified in SAs were of similar magnitude: a 1 mmol/l increase of FAw3 associated with a +0.4 mmol/l increase in fasting glucose or and a 1 mmol/increase of FAw6 and LA associated with a +0.4 mmol/l increase of 2-hour post glucose.

No metabolite found to be associated with postprandial glucose measures in WEs was found to be associated with postprandial glucose in

10.3389/fendo.2023.1157416

TABLE 2 Removal of potentially pleiotropic SNPs.

| Ethnicity | Metabolite | SNP | Gene | Associated confounder | Initial | | Confounder removal | |
|---|---|---|---|---|---|---|---|---|
| | | | | | β estimate (95% CI) | WuH | β estimate (95% CI) | WuH |
| WE | Leucine | rs2984433 | ACTG1P9 | BMI, Obesity class 1, weight | -0.193 (-0.319, -0.068) $^{FG}$ | 0.004 | -0.203 (-0.339, -0.068) $^{FG}$ | 0.006 |
| | | | | | -0.443 (-0.774, -0.113) $^{2H}$ | 0.024 | -0.547 (-0.909, -0.185) $^{2H}$ | 0.002 |
| | L-HDL_L | rs6811162 | ENPEP | Self-reported hypertension, diagnosed high blood pressure | 0.264 (0.062, 0.464) | 0.014 | 0.301 (0.092, 0.510) | 0.006 |
| | L-HDL-C | rs5576825 | LINC01621 ELOVL4 | Waist circumference | 0.279 (0.012, 0.544) | 0.048 | 0.323 (0.0456, 0.602) | 0.107 |
| | | rs6811162 | ENPEP | Self-reported hypertension, diagnosed high blood pressure | | | 0.241 (-0.037, 0.519) | 0.026 |
| | | rs5576825 + rs6811162 | - | - | | | 0.287 (-0.004, 0.578) | 0.063 |
| | XL-HDL class | rs5576825 | LINC01621 ELOVL4 | Waist circumference | -0.285 (-0.552, -0.018) | 0.015 | -0.244 (-0.528, 0.040) | 0.135 |
| SA | LA | rs12720820 | APOB | Self-reported high cholesterol, coronary artery disease, treatment with cholesterol lowering medication | 0.477 (0.013, 0.939) | 0.030 | 0.335 (-0.982, 0.763) | 0.459 |
| | FAw6 | rs12720820 | APOB | Self-reported high cholesterol, coronary artery disease, treatment with cholesterol lowering medication | 0.445 (0.094, 0.794) | 0.007 | 0.223 (-0.398, 0.843) | 0.105 |
| | L-HDL-PL | rs7486176 | C12orf76 | Systolic blood pressure, diagnosed high blood pressure, hypertension | 0.692 (0.106, 1.28) | 0.021 | 0.853 (0.170, 1.54) | 0.013 |
| | Fatty Acid class | rs12720820 | APOB | Self-reported high cholesterol, coronary artery disease, treatment with cholesterol lowering medication | 0.172 (0.018, 0.327) | 0.018 | 0.142 (-0.049, 0.334) | 0.119 |

2H, 2-hour post glucose; FAw6, Total n-6 fatty acids; FG, fasting glucose; LA, 18,2 linoleic acid (mmol/L); L-HDL-C, total cholesterols in L-HDL (mmol/L); L-HDL_L, total lipids in L-HDL (mmol/L); L-HDL-PL, phospholipids in L-HDL (mmol/L); SA, South Asian; WE, White European; WUH, Wu-Hausman p-value.

SAs. However, in both populations, members of the S-HDL and L-HDL class were found to be associated with increased postprandial glucose.

#### 3.2.2.1 Sensitivity analyses

Six instruments in SAs were comprised of a single SNP meaning it was not possible to perform a leave-one-out analysis for these metabolites. For the remaining 5 metabolites, associations were consistent across each leave-one-out analyses (Supplementary Figure 9). Likewise, no large differences in F-statistics following the removal of individual SNPs were identified (Supplementary Figure 10).

Just as in WEs, 3 metabolites identified in SAs included SNPs associated with cholesterol or hypertension, which are potential confounders of the association between metabolites and dysglycemia (Supplementary Table 7). Significance was maintained following the removal of SNP rs7486176 (found within the *C12orf76* gene) from the total phospholipids in L-HDL instrument. For the LA and FAw6 exposures, the removal of SNP rs12720820 (found within the *APOB* gene) resulted in a non-significant association indicating that this SNP was the main driver of the identified association (Table 2). In leave-one-out analyses, the removal of SNP rs58865405 from the FAw6 instrument resulted in non-significance, although the biological role of this SNP remains unknown.

## 3.3 *Post-hoc* analysis: analysis of metabolite classes

Numerous SNPs were found to be associated with more than one metabolite measure, particularly for metabolites in the same metabolite class (Supplementary Figures 11-12). This was anticipated since many metabolomic pathways are biologically

10.3389/fendo.2023.1157416

TABLE 3 Significant MR results in South Asians.

| | Class | Metabolite | Outcome | F statistic | β estimate (95% CI) | WuH |
|---|---|---|---|---|---|---|
| Lowers Glucose | L-LDL | LDL-P | 2-hour post | 11.7 | -2238 (-4828 -193.7) | 0.024 |
| | S-LDL | SLDL-L | 2-hour post | 11.1 | -3.86 (0.467, -7.27) | 0.015 |
| | IDL | IDL-C | 2-hour post | 10.9 | -1.19 (-0.12, -2.27) | 0.021 |
| | | IDL-CE | 2-hour post | 11.8 | -1.34 (-0.144, -2.55) | 0.023 |
| | S-HDL | S-HDL-L | 2-hour post | 11.4 | -1.23 (-0.137, -2.32) | 0.012 |
| Increases Glucose | Fatty Acids | LA | 2-hour post | 20.9 | 0.477 (0.013, 0.939) | 0.030 |
| | | FAw3 | Fasting glucose | 10 | 0.432 (0.063, 0.798) | 0.008 |
| | | FAw6 | 2-hour post | 33.4 | 0.445 (0.094, 0.794) | 0.007 |
| | M-VLDL | M-VLDL-L | 2-hour post | 68.4 | 0.046 (0.009, 0.083) | 0.008 |
| | L-HDL | L-HDL-PL | 2-hour post | 43.3 | 0.692 (0.106, 1.28) | 0.021 |
| | S-HDL | S-HDL-C | Fasting glucose | 22 | 1 (0.116,1.882) | 0.012 |

Glucose measures are expressed as mmol/L. FAw3, total n-3 fatty acids; FAw6, total n-6 fatty acids; IDL-C, total cholesterols in LDL (mmol/L); IDL-CE, cholesterol esters in LDL (mmol/L); LA, 18,2 Linoleic Acid (mmol/L); LDL_P, concentration of LDL particles (mol/L); L-HDL-PL, phospholipids in L-HDL (mmol/L); M-VLDL-L, total lipids in M-VLDL (mmol/L); S-HDL-C, total cholesterols in S-HDL (mmol/L); S-HDL-L, total lipids in S-HDL (mmol/L); S-LDL-L, total lipids in S-LDL (mmol/L); WuH, Wu-Hausman p-value.

intertwined. To minimise the risk of violation of the third MR assumption (that the genetic instrument must only influence the outcome *via* the exposure and not *via* an alternative biological pathway) (14), the collective effect of an entire class of metabolites on postprandial glucose measures was examined. A composite score for each metabolite class was created by placing all metabolites in a single class (e.g. all LDL metabolites), conducting a PCA and extracting PC1. This was only possible for 20 classes in WEs and 21 classes in SAs that had > 2 metabolites and ≥70% of the class variation was explained by PC1 (Supplementary Table 9). To assess the impact of outliers on PCA, outliers were defined and removed based on two cut-offs: standard (1.5 X IQR from the median) and stringent (3 x IQR from the median). For all classes, PC1 and PC2 scores were comparable after removal of both types of outliers so only 3xIQR outliers were removed prior to analyses (Supplementary Table 10).

138 SNPS remained after LD thinning in WEs, 87 (63.04%) of which were unique to a single metabolite class exposure. 19/20 (95%) of the metabolite classes examined in WEs had an F-statistic ≥10 (the only exception being the MHDL class). 54 SNPS remained after LD thinning in SAs, 42 (77.78%) of which were unique to a single metabolite class. Screening of metabolite class predictors in Phenoscanner and GWAS databases did not raise concerns for horizonal (Supplementary Table 11). Despite the lower number of SNPs identified in SAs,17/20 (85%) of the metabolite classes examined had a F-statistic ≥10 indicating that instrument strength was still sufficient, for all classes except for the non-branched amino acids, LLDL class, and all VLDL classes. On average, the mean F-statistic of metabolite class instruments was 19.83%. On average genetic instruments of composite measures of the metabolite classes were weaker than the instruments for individual metabolites in both WEs and SAs (Supplementary Figure 13).

In WEs, 4 metabolite classes were associated with a glucose measure: S-LDLs associated with fasting and 2-hour post glucose; M-LDL and all LDLS (i.e. the collective grouping of HDLs, MDLs and SDLs) were associated with fasting glucose, and XL-HDLs were associated with 2-hour post glucose (Table 4). In SAs, the fatty acid metabolite class were associated with 2-hour post glucose levels. No other associations were identified in SAs.

As with the sensitivity analyses of individual metabolites, the removal of individual SNPs was not found to greatly impact the F-statistic of most instruments (Supplementary Figures 14-15).

TABLE 4 Significant MR results from the analysis of metabolite classes.

| Ethnicity | Metabolite class | Outcome | F statistic | β estimate (95% CI) | WUH |
|---|---|---|---|---|---|
| **White European** | XL-HDL | 2-hour post glucose | 96.2 | -0.285 (-0.018, -0.552) | 0.015 |
| | M-LDL | Fasting glucose | 95.5 | -0.048 (-0.004, - 0.091) | 0.024 |
| | S-LDL | Fasting glucose | 98 | 0.084 (0.007, 0.162) | 0.024 |
| | | 2-hour post glucose | 98 | -0.249 (-0.016, -0.482) | 0.066 |
| | All LDL | Fasting glucose | 95 | 0.038 (-0.005, -0.068) | 0.016 |
| **South Asian** | Fatty Acids | 2-hour post glucose | 32.8 | 0.172 (0.018, 0.327) | 0.0184 |

Glucose measures are expressed as mmol/L. WuH: Wu-Hausman p-value.

However, for the fatty acid metabolites class, the exclusion of rs12720820 or rs7159441, and for 'XL-HDL', the removal of rs55768285, resulted in non-significant associations, suggesting that these SNPs are key drivers of the association. (Table 4; Supplementary Table 10).

## 3.4 Power analysis

$R^2$ values were consistently lower in the ethnic group where an effect was not detected but all genetic instruments for the metabolite values had an F-statistic ≥ 10, indicating that weak instrument bias was not responsible for the absence of significant effects. Where an association was identified in one ethnic group but not another, to determine whether the absence of an association was potentially due a lack of power in the other ethnicity rather than an ethnic-specific effect, *post-hoc* power analyses were performed.

When using MR estimates as an estimate for the true causal effect both the analyses of FAw3 and the overall fatty acid class in WEs were adequately powered to detect the observed MR effect in both populations. Therefore, the absence of an effect of FAw3 in WEs is unlikely due to inadequate power (Supplementary Table 12). The analysis of HDL2C and HDL3C in SAs was also sufficiently powered to detect the observed MR effect in WEs.

# 4 Discussion

This study has identified ethnically distinct associations between a range of metabolites and postprandial glucose measures taken during pregnancy in SAs and WEs, with notably no shared associations were identified. Fourteen metabolites were found to be associated with postprandial glucose measures in WEs. Whereas, a distinct set of 11 metabolites were associated in SAs. In WEs, cholesterols and lipoproteins were the metabolite classes associated with postprandial glucose measures, while in SAs fatty acids were the most commonly associated.

Furthermore, through an extensive GWAS of metabolites, this study identified novel genome-wide significant associations in relation to acetate (1 SNP, rs10945476) and tyrosine (15 SNPs, all on chromosome 17) in SAs. No previous associations have been identified for SNP rs10945476, found within the non-coding transcript gene *PRDM15* in relation to acetate or any other exposure.

Interestingly, 3 of the 15 SNPs associated with tyrosine are found in a transmembrane transporter gene, *SLC13A2*. Moreover, an additional 10 of the newly identified 15 SNPs associated with tyrosine were found in the *FOXN* gene, a transcription factor that has previously been identified to be associated with ceramide levels (a lipid metabolite) in a GWAS from a Chinese cohort (52). Moreover, ceramide has been shown to induce tyrosine phosphorylation in membrane proteins meaning it is plausible that a gene associated with ceramide is also associated with tyrosine levels in an Asian population (51). Interestingly, ceramide has been proposed as a mediator of the interaction between saturated fat and insulin resistance and has been associated with T2D and cardiovascular disease (53). To the best of our knowledge tyrosine levels have not previously been associated with either *FOXN SLC13A2*. The remaining 2 of the 15 SNPs identified as being associated with tyrosine in SAs, are currently not in any known genes. All 15 SNPs identified as being associated with tyrosine in SAs are in LD with each other in 1000G SA populations (all $R^2 \geq 0.38$). In agreement with a recent GWAS of GDM and T2D (54), almost all SNPs included in an instrument significantly associated with either outcome (fasting glucose or 2-hour post glucose) have previously been associated with a diabetes related disease outcome, providing additional evidence for their validity as instrumental variables. However, since most evidence of genomics-diabetes associations are sourced from non-SA populations, these results may not accurately reflect genetic associations in SAs.

## 4.1 Identified associations in white Europeans (WEs)

### 4.1.1 Leucine

Branched chain amino acids (BCAAs), including leucine, are predominantly metabolised in skeletal muscles where they regulate protein synthesis and mitochondrial functions (55). In addition, BCAAs are hormonal signalling regulators and are expected to module insulin resistance (IR) through increasing insulin secretion in human pancreatic β-cells (56, 57). Our study found leucine to be negatively associated with both fasting glucose and 2-hour post glucose levels during pregnancy in WEs; with 1 mmol/L of leucine associated with a decrease of 0.193 mmol/L in fasting glucose and 0.327 mmol/L 2-hour post glucose respectively. Although few studies have investigated the role of leucine in glucose regulation during pregnancy, interestingly the ratio of leucine/isoleucine was similarly found associated with reductions in fasting glucose in the HAPO study, a multi-ethnic cohort of pregnant women of Afro-Caribbean, Mexican American, Northern European, and Thai ancestry (58). Common dietary sources of leucine include meat products and cheese, with smaller amounts also being present in other dairy products (such as dairy and yoghurt), fish and in certain legumes and nuts, such as dried raw broad beans and pine nuts (55). Hence, dietary interventions aimed at increasing leucine levels during pregnancy, possibly through a dietary intervention promoting the consumption of lean animal protein, low-fat dairy and nuts, may help improve pregnancy hyperglycaemia in WEs.

### 4.1.2 Cholesterols

HDL cholesterol is colloquially described as 'good cholesterol' due to its role in the removal of cholesterols from atherosclerotic plaque, thereby reducing an individual's risk of CVD (59). Furthermore, low HDL levels have commonly been associated with diabetes in humans, with HDL shown to increase insulin secretion and β-cell survival (60, 61). We identified four associations between HDL cholesterol and postprandial glucose

measures in WEs. Herein, 1 mmol/L increase in S-HDL-CE confers a 1.78 mmol/L increase (95% CI 0.49 – 3.11) in 2-hour post glucose. This is consistent with previous evidence from a Finnish sample of overweight and obese women where cholesterol esters in S-HDL were higher in the serum samples of GDM cases at ~14 weeks gestation (62). Discrepancies in the direct effect of HDL cholesterol on dysglycemia have also been identified in the genetic literature (61), with a recent review highlighting that while a genetic study utilising linear relation analysis did find HDLs to have a protective effects against T2D (n cases = 2,447) (63), the same effect was not been replicated in an MR setting (n cases= 47,627) (64). When considering LDL cholesterols, only S-LDLs was found to be significantly associated with a postprandial glucose measure (fasting glucose) in WEs. Additionally, in our composite analysis of metabolite classes, S-LDLs were associated with fasting glucose and 2-hour post glucose in WEs, whereas the M-LDL and all LDL (a combined measure of S-LDLs, M-LDLs and HDLs) classes were associated with fasting glucose. Unfortunately, because composite scores were comprised of PC1 coordinates the direction of effect of these associations could not be evaluated. To our knowledge, no previous study has conducted an MR of metabolites on dysglycemic predictors of GDM.

### 4.1.3 Triglycerides

Triglycerides are an abundant class of lipid particles found in the blood, originating from either from the consumption of dietary fats or as a result of hepatic metabolism (65, 66). Once in the blood, triglycerides can be incorporated into HDL and LDL cholesterol particles. In addition to dietary triglyceride consumption, dietary fatty acids can be converted into triglycerides before they enter circulation, highlighting the complex relationship between triglyceride, cholesterol, and fatty acid levels (65).

Our results suggest triglycerides in XSVLDL (XS-VLDL-TG) associate with increased 2-hour post glucose (0.841 mmol/L) in WEs. In agreement with these findings, increased triglycerides in XSVLDL levels have also previously been associated with increased likelihood of GDM in a Finnish population (62). No other triglyceride was found to be associated with in WEs. One explanation no additional associations were detected could be due to the average BMI of the WEs in BiB. For example, an analysis of a prospective Irish cohort (~94% WE) found that triglyceride levels were only associated with GDM in obese individuals, a higher average BMI than that observed in the BiB cohort (67). Further confirmation of these findings of increased triglycerides in XS-VLDLs would suggest that this association is, at least in part, responsible for the identified associations between diets high in fats and increased prevalence of GDM in WEs (10).

## 4.2 Identified associations in South Asians (SAs)

### 4.2.1 Fatty acids

Polyunsaturated acids (PUFAs) are consumed in the diet and can be converted to long-chain PUFAs (LC-PUFAs) through a process of desaturation and elongation reactions that predominately occur in the liver (68). Changes in dietary patterns can have a large impact on fatty acid composition in the body and with-it disease risk. For example, a western dietary pattern, which has high levels of n-6 fatty acids, has been associated with GDM risk (10, 69). In a cohort of Chinese adults, total n-6 fatty acids and 18:2 n-6 levels at baseline in venous blood samples were both found to associate with an increased risk of T2D after ~8 years of follow up, while increased n-3 fatty acid levels were protective (70). However, a recent two-sample MR suggested only a negligible effect of n-6 PUFA synthesis on T2D in a predominantly WE cohort (71). Moreover, the relationship between n-3, n-6, n-9 fatty acids and GDM remains inconclusive, and in a recent (2021) systematic review none of the identified studies (n=15) was conducted in a SA population (69), highlighting the need for more studies exploring the role of fatty acids in GDM development in Asians (72).

This study provides evidence of an association between LA and total FAw6 levels and an increase in 2-hour post glucose levels during pregnancy in SAs. In addition, the fatty acid class associated with 2-hour post glucose in SAs. Through a leave-out-one sensitivity analyses for the FAw6 and LA instruments, the removal of the SNP rs12720820 (found within the *APOB* gene) resulted in non-significant associations for both exposures, indicating that this rs12720820 was the largest contributor to the identified associations and has previously been associated with cholesterol levels and the use of cholesterol lowering drugs (73). Interestingly, FAw3 also associated with increased 2-hour post glucose in SAs; however, with only one SNP potential pleiotropy could not be explored.

It is well established that fatty acid profiles can impact blood cholesterol levels (74–76). In addition, increased dietary cholesterol has previously been associated with an increased risk of GDM in a systematic review of observational studies (77). Taken together, our data confirm that fatty acids and cholesterol metabolites are in vertically pleiotropy and are likely impacting gestational dysglycemia *via* the same causal pathway. Unlike horizontal pleiotropy, vertical pleiotropy does not result in a violation of the 2$^{nd}$ MR assumption as cholesterol is not acting as a confounder, meaning MR estimates are still valid (Figure 1). Furthermore, it is also possible that this interaction between fatty acids and cholesterols may be ethnic-specific due to the absence of associations identified between fatty acids and postprandial glucose measures in WEs. In addition to possible variations in cholesterol metabolism, it is plausible that variations in fatty acid synthesis are also partially responsible for the increased GDM risk experienced by SAs. For example, variants within the *FADS* genes impact LC-PUFA conversion (78, 79). Current evidence suggests that SAs are likely to synthesise LC-PUFA more quickly than WEs, which could contribute to elevated risk of prolonged exposure to elevated LC-PUFA levels (namely, w6) and risk of dysglycemia (78, 79). If these ethnic differences in fatty acid metabolism are confirmed to be linked to disease risk, it would aid in the development of tailored GDM prevention strategies that focus on modifying fatty acid profiles in an ethnic-specific manner.

The analyses found no association between triglycerides and dysglycemia in SAs. This agrees with a recent meta-analysis that concluded, although triglyceride levels associated with likelihood of

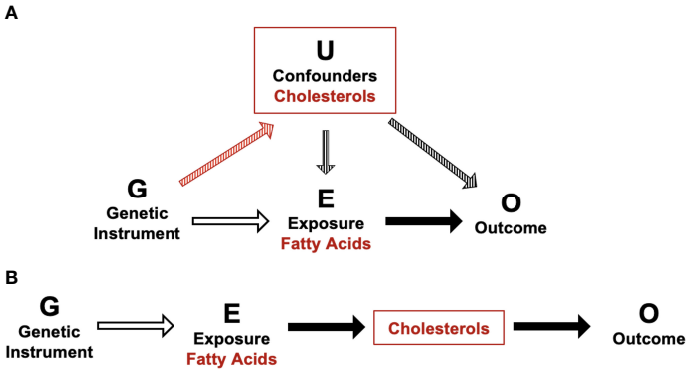



FIGURE 1
Schematic of potential horizontal and vertical pleiotropy in relation to fatty acid and cholesterol metabolites and postprandial glucose measures. **(A)** Illustration of horizontal pleiotropy. **(B)** Illustration of vertical pleiotropy. Vertical pleiotropy does not result in a violation of the 2nd MR assumption because the metabolites progress along a single linear causal pathway.

GDM ($I^2 \geq 84\%$) (80), after stratification by culture/geographical location, they found no association between triglyceride levels and likelihood of GDM. The reasons for this are unclear but it has also been shown that SAs have a higher prevalence of hypertriglyceridemia than WEs and at lower BMI levels, meaning it is possible that the difference in triglyceride levels and in SA GDM cases and controls is less pronounced than in WEs (81).

## 4.3 Strengths and limitations

This analysis has several strengths. Firstly, this study involved a large and comprehensive panel of metabolites allowing for the relationships between metabolites and postprandial glucose to be thoroughly investigated. Secondly, this is the first MR study to investigate dysglycemia during pregnancy while also being one of the few MR studies to be conducted in a SA population. Finally, through leave-one-out analyses and the searching of both Phenoscanner and GWAS Catalog databases, violations of the 2$^{nd}$ and 3$^{rd}$ MR assumptions were thoroughly investigated meaning that it was possible to conclude that identified robust associations between genetic variants and outcomes (1$^{st}$ assumption of MR) may not be subjected to horizontal pleiotropy and that identified causal associations are valid due to the absence of detectable violations of the MR assumptions.

Nonetheless, this study has some limitations. Firstly, metabolites are highly correlated meaning it is not possible to confidently interpret that an individual metabolite is independently associated with a postprandial outcome measure. To account for this limitation MR analyses were performed on composite measures of each metabolite class (when PC1 explained ≥70% of the variation in the metabolite class) to assess the overall impact of each metabolite class on pregnancy dysglycemia. Secondly, MR also assumes the level of genetically conferred exposure from conception to the time of measurement is constant, which is unknown when studying metabolites – therefore, we cannot presume that these associations would be observed outside of pregnancy. Thirdly, limited sample size may have led to some underpowered analyses and combined with high consanguinity persuaded us to use statistical modelling to account for 'relatedness' (rather than participant pruning) to preserve analytical power and study integrity. We acknowledge that this strategy has limitations, and tested for their effect (i.e., LOO analysis) (25), and look forward to the future when larger SA prospective cohort are available, and we can validate these results with increased confidence. In addition, the limited sample size meant that further adjustment could not be made at the GWAS stage since missing data in certain variables would further reduce sample size (e.g., age) and some associations may have been underpowered to detect an effect. However, a *post hoc* power analyses found that for some metabolite values significant effects only identified in one ethnicity were possible to detect in the alternate ethnicity. Fourthly, some genetic instruments included only one SNP meaning it was not possible to evaluate the impact of pleiotropy for any identified associations involving these instruments. Fifthly, it was not possible to fully assess the presence of associations between SNPs included in significant instruments, potential confounders and T2D traits in SAs due to the limited number of GWAS conducted in SAs. Although it is likely that many of the associations in WEs are also present in SAs, it is possible that not all associations identified in WE are present in SAs and that some SA-specific pleiotropic associations are unknown. Lastly, due to limitations in data availability in SAs a two-sample MR could not be conducted meaning it was not possible to assess the generalisability of these findings.

# 5 Conclusions

The presence of causal relationships between a comprehensive set of distinct metabolites and metabolite families with postprandial glucose measures (fasting glucose and 2-hour post glucose) in mid-pregnancy has been established in a UK SA and WE population. This study has found a range of metabolite values to be associated with postprandial glucose measures in WEs and high-risk SA women, although more associations were identified in WEs despite these individuals being at lower risk of GDM. In high-risk SA women, total n-6 fatty acids and the n-6 fatty acid, LA appear to

increase postprandial glucose levels suggesting that fatty acids may be responsible for a large proportion of metabolically driven risk for GDM experienced by this population. Future work in a larger sample (potentially using a two-sample MR) and a larger panel of metabolites is needed to investigate our findings and hypotheses more closely, ideally over the course of a pregnancy in order to aid in GDM prevention in this high-risk population.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: Born in Bradford [https://borninbradford.nhs.uk].

## Author contributions

HF, MI, and MZ contributed to conception and design of the study. HF organized the database and performed the statistical analysis. JM, MI, and MZ supervised the study. HF and MZ wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

MZ is currently funded by the Wellcome Trust [217446/Z/19/Z]. The Born in Bradford data reported here were supported by a Wellcome Trust infrastructure grant (101597) and the National Institute for Health Research ARC Yorkshire and Humber [NIHR200166]. Funding for the metabolomics analyses has been provided by the US National Institutes of Health [R01 DK10324], the European Research Council (ERC) under the European Union's Seventh Framework Programme [FP7/2007-2013]/ERC grant agreement no 669545 and the MRC *via* the MRC Integrative Epidemiology Unit Programme [MC_UU_00011/6].

## Acknowledgments

Born in Bradford is only possible because of the enthusiasm and commitment of the children and parents in BiB. We are grateful to all the participants, health professionals, schools and researchers who have made Born in Bradford happen. This work was undertaken on ARC4, part of the High Performance Computing (HPC) facilities at the University of Leeds, UK.

## 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.2023.1157416/full#supplementary-material

## References

1. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes mellitus. *Int J Mol Sci* (2018) 19(11):3342. doi: 10.3390/ijms19113342

2. Taylor K, Ferreira DLS, West J, Yang T, Caputo M, Lawlor DA. Differences in pregnancy metabolic profiles and their determinants between white European and south Asian women: findings from the born in Bradford cohort. *Metabolites* (2019) 9 (9):190. doi: 10.3390/metabo9090190

3. Mills HL, Patel N, White SL, Pasupathy D, Briley AL, Santos Ferreira DL, et al. The effect of a lifestyle intervention in obese pregnant women on gestational metabolic profiles: findings from the UK pregnancies better eating and activity trial (UPBEAT) randomised controlled trial. *BMC Med* (2019) 17(1):15. doi: 10.1186/s12916-018-1248-7

4. Kampmann U, Knorr S, Fuglsang J, Ovesen P. Determinants of maternal insulin resistance during pregnancy: an updated overview. *J Diabetes Res* (2019) 2019:5320156. doi: 10.1155/2019/5320156

5. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. *Nat Rev Dis Primers* (2019) 5(1):47. doi: 10.1038/s41572-019-0098-8

6. Rodrigo N, Glastras SJ. The emerging role of biomarkers in the diagnosis of gestational diabetes mellitus. *J Clin Med* (2018) 7(6):120. doi: 10.3390/jcm7060120

7. Menard V, Sotunde OF, Weiler HA. Ethnicity and immigration status as risk factors for gestational diabetes mellitus, anemia and pregnancy outcomes among food insecure women attending the Montreal diet dispensary program. *Can J Diabetes.* (2020) 44(2):139–45.e1. doi: 10.1016/j.jcjd.2019.05.004

8. Gadve SS, Chavanda S, Mukherjee AD, Aziz S, Joshi A, Patwardhan M. Risk of developing type 2 diabetes mellitus in south Asian women with history of gestational diabetes mellitus: a systematic review and meta-analysis. *Indian J Endocrinol Metab* (2021) 25(3):176–81. doi: 10.4103/ijem.IJEM_57_21

9. Fuller H, Iles M, Moore JB, Zulyniak MA. Unique metabolic profiles associate with gestational diabetes and ethnicity in low and high-risk women living in the UK. *J Nutr* 152(10):2186–97. doi: 10.1093/jn/nxac163

10. Fuller H, Moore JB, Iles MM, Zulyniak MA. Ethnic-specific associations between dietary consumption and gestational diabetes mellitus incidence: a meta-analysis. *PloS Global Public Health* (2022) 2(5):e0000250. doi: 10.1371/journal.pgph.0000250

11. Griffith RJ, Alsweiler J, Moore AE, Brown S, Middleton P, Shepherd E, et al. Interventions to prevent women from developing gestational diabetes mellitus: an overview of cochrane reviews. *Cochrane Database Syst Rev* (2020) 6:CD012394. doi: 10.1002/14651858.CD012394.pub3

12. de Souza RJ, Shanmuganathan M, Lamri A, Atkinson SA, Becker A, Desai D, et al. Maternal diet and the serum metabolome in pregnancy: robust dietary biomarkers generalizable to a multiethnic birth cohort. *Curr Dev Nutr* (2020) 4(10):nzaa144–nzaa. doi: 10.1093/cdn/nzaa144

13. Wang QY, You LH, Xiang LL, Zhu YT, Zeng Y. Current progress in metabolomics of gestational diabetes mellitus. *World J Diabetes* (2021) 12(8):1164–86. doi: 10.4239/wjd.v12.i8.1164

14. Sheehan NA, Meng S, Didelez V. Mendelian randomisation: a tool for assessing causality in observational epidemiology. In: Teare MD, editor. *Genetic epidemiology* Totowa, NJ: Humana Press (2011). p. 153–66.

15. Liu J, van Klinken JB, Semiz S, van Dijk KW, Verhoeven A, Hankemeier T, et al. A mendelian randomization study of metabolite profiles, fasting glucose, and type 2 diabetes. *Diabetes* (2017) 66(11):2915–26. doi: 10.2337/db17-0199

16. Carreras-Torres R, Johansson M, Gaborieau V, Haycock PC, Wade KH, Relton CL, et al. The role of obesity, type 2 diabetes, and metabolic factors in pancreatic cancer: a mendelian randomization study. *J Natl Cancer Inst* (2017) 109(9):djx012. doi: 10.1093/jnci/djx012

17. Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, et al. Cohort profile: the born in Bradford multi-ethnic family cohort study. *Int J Epidemiol* (2013) 42 (4):978–91. doi: 10.1093/ije/dys112

18. Bird PK, McEachan RRC, Mon-Williams M, Small N, West J, Whincup P, et al. Growing up in Bradford: protocol for the age 7-11 follow up of the born in Bradford birth cohort. *BMC Public Health* (2019) 19(1):939. doi: 10.1186/s12889-019-7222-2

19. Lawlor DA, West J, Fairley L, Nelson SM, Bhopal RS, Tuffnell D, et al. Pregnancy glycaemia and cord-blood levels of insulin and leptin in Pakistani and white British mother-offspring pairs: findings from a prospective pregnancy cohort. *Diabetologia* (2014) 57(12):2492–500. doi: 10.1007/s00125-014-3386-6

20. Josse J, Husson F. missMDA: a package for handling missing values in multivariate data analysis. *J Stat Software* (2016) 70(1):1–31. doi: 10.18637/jss.v070.i01

21. Arciero E, Dogra SA, Malawsky DS, Mezzavilla M, Tsismentzoglou T, Huang QQ, et al. Fine-scale population structure and demographic history of British pakistanis. *Nat Commun* (2021) 12(1):7189. doi: 10.1038/s41467-021-27394-2

22. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *Int J Methods Psychiatr Res* (2018) 27(2):e1608. doi: 10.1002/mpr.1608

23. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. *Nat Rev Methods Primers* (2021) 1(1). doi: 10.1038/s43586-021-00056-9

24. Sheridan E, Wright J, Small N, Corry PC, Oddie S, Whibley C, et al. Risk factors for congenital anomaly in a multiethnic birth cohort: an analysis of the born in Bradford study. *Lancet* (2013) 382(9901):1350–9. doi: 10.1016/S0140-6736(13)61132-0

25. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet* (2014) 46 (2):100–6. doi: 10.1038/ng.2876

26. 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

27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* (2007) 81

28. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* (2011) 88(1):76–82. doi: 10.1016/j.ajhg.2010.11.011

29. van den Berg S, Vandenplas J, van Eeuwijk FA, Lopes MS, Veerkamp RF. Significance testing and genomic inflation factor using high-density genotypes or whole-genome sequence data. *J Anim Breed Genet* (2019) 136(6):418–29. doi: 10.1111/jbg.12419

30. Rivadeneira F, Uitterlinden AG. Chapter 18 - genetics of osteoporosis. In: Dempster DW, Cauley JA, Bouxsein ML, Cosman F, editors. *Marcus And feldman's osteoporosis, Fifth Edition*. (London EC2Y 5AS, United Kingdom: Academic Press (2021). p. 405–51.

31. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* (2010) 26(17):2190–1. doi: 10.1093/bioinformatics/btq340

32. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* (2017) 8(1):1826. doi: 10.1038/s41467-017-01261-5

33. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. *Elife* (2018) 7:e34408. doi: 10.7554/eLife.34408

34. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* (2015) 31(21):3555–7. doi: 10.1093/bioinformatics/btv402

35. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: an r package for rapidly calculating linkage disequilibrium statistics in diverse populations. *Front Genet* (2020) 11:157. doi: 10.3389/fgene.2020.00157

36. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* (2015) 526(7571):68–74. doi: 10.1038/nature15393

37. R Development Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (2010).

38. Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag (2016).

39. 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(1). doi: 10.1186/s13742-015-0047-8

40. Budu-Aggrey A, Brumpton B, Tyrrell J, Watkins S, Modalsli EH, Celis-Morales C, et al. Evidence of a causal relationship between body mass index and psoriasis: a mendelian randomization study. *PloS Med* (2019) 16(1):e1002739. doi: 10.1371/journal.pmed.1002739

41. Fox J, Kleiber C, Zeileis A. *_ivreg: instrumental-variables regression by '2SLS', '2SM', or '2SMM', with diagnostics*. R package version 0.6-1 ed2021. (2020)

42. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. *BMJ* (2021) 375:n2233. doi: 10.1136/bmj.n2233

43. 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:D1005-D1012. doi: 10.1093/nar/gky1120

44. 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

45. 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

46. Davies NM, Holmes MV, Davey Smith G. Reading mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Bmj* (2018) 362:k601. doi: 10.1136/bmj.k601

47. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in mendelian randomization studies. *Int J Epidemiol* (2012) 42(5):1497–501. doi: 10.1093/ije/dyt179

48. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, et al. Ensembl 2021. *Nucleic Acids Res* (2020) 49(D1):D884–D91. doi: 10.1093/nar/gkaa942

49. *Ensembl. gene: SORCS1 e! ensembl: ensembl release 105; 2021* . Available at: https://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000108018;r=10:106573663-107164706.

50. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* (2008) 27(8):1133–63. doi: 10.1002/sim.3034

51. Gulbins E, Szabo I, Baltzer K, Lang F. Ceramide-induced inhibition of T lymphocyte voltage-gated potassium channel is mediated by tyrosine kinases. *Proc Natl Acad Sci USA* (1997) 94(14):7661–6. doi: 10.1073/pnas.94.14.7661

52. Chai JF, Raichur S, Khor IW, Torta F, Chew WS, Herr DR, et al. Associations with metabolites in Chinese suggest new metabolic roles in alzheimer's and parkinson's diseases. *Hum Mol Genet* (2020) 29(2):189–201. doi: 10.1093/hmg/ddz246

53. Yki-Jarvinen H, Luukkonen PK, Hodson L, Moore JB. Dietary carbohydrates and fats in nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* (2021) 18 (11):770–86. doi: 10.1038/s41575-021-00472-y

54. Pervjakova N, Moen G-H, Borges M-C, Ferreira T, Cook JP, Allard C, et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. *Hum Mol Genet* (2022) 31(19):3377–91. doi: 10.1093/hmg/ddac050

55. Rondanelli M, Nichetti M, Peroni G, Faliva MA, Naso M, Gasparri C, et al. Where to find leucine in food and how to feed elderly with sarcopenia in order to counteract loss of muscle mass: practical advice. *Front Nutr* (2021) 7:622391. doi: 10.3389/fnut.2020.622391

56. Nie C, He T, Zhang W, Zhang G, Ma X. Branched chain amino acids: beyond nutrition metabolism. *Int J Mol Sci* (2018) 19(4):954. doi: 10.3390/ijms19040954

57. Leenders M, van Loon LJ. Leucine as a pharmaconutrient to prevent and treat sarcopenia and type 2 diabetes. *Nutr Rev* (2011) 69(11):675–89. doi: 10.1111/j.1753-4887.2011.00443.x

58. Liu Y, Kuang A, Talbot O, Bain JR, Muehlbauer MJ, Hayes MG, et al. Metabolomic and genetic associations with insulin resistance in pregnancy. *Diabetologia* (2020) 63(9):1783–95. doi: 10.1007/s00125-020-05198-1

59. *CDC. LDL and HDL cholesterol: "Bad" and "Good" cholesterol: US.gov* (2020). Available at: https://www.cdc.gov/cholesterol/ldl_hdl.htm.

60. Ryckman KK, Spracklen CN, Smith CJ, Robinson JG, Saftlas AF. Maternal lipid levels during pregnancy and gestational diabetes: a systematic review and meta-analysis. *BJOG* (2015) 122(5):643–51. doi: 10.1111/1471-0528.13261

61. Wong NKP, Nicholls SJ, Tan JTM, Bursill CA. The role of high-density lipoproteins in diabetes and its vascular complications. *Int J Mol Sci* (2018) 19 (6):1680. doi: 10.3390/ijms19061680

62. Mokkala K, Vahlberg T, Pellonperä O, Houttu N, Koivuniemi E, Laitinen K. Distinct metabolic profile in early pregnancy of overweight and obese women developing gestational diabetes. *J Nutr* (2020) 150(1):31–7. doi: 10.1093/jn/nxz220

63. Qi Q, Liang L, Doria A, Hu FB, Qi L. Genetic predisposition to dyslipidemia and type 2 diabetes risk in two prospective cohorts. *Diabetes* (2012) 61(3):745–52. doi: 10.2337/db11-1254

64. 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(9):3328–33. doi: 10.2337/db14-1603

65. Alexopoulos A-S, Qamar A, Hutchins K, Crowley MJ, Batch BC, Guyton JR. Triglycerides: emerging targets in diabetes care? review of moderate hypertriglyceridemia in diabetes. *Curr Diabetes Rep* (2019) 19(4):13–. doi: 10.1007/s11892-019-1136-3

66. Laufs U, Parhofer KG, Ginsberg HN, Hegele RA. Clinical review on triglycerides. *Eur Heart J* (2020) 41(1):99–109c. doi: 10.1093/eurheartj/ehz785

67. O'Malley EG, Reynolds CME, Killalea A, O'Kelly R, Sheehan SR, Turner MJ. Maternal obesity and dyslipidemia associated with gestational diabetes mellitus (GDM). *Eur J Obstet Gynecol Reprod Biol* (2020) 246:67–71. doi: 10.1016/j.ejogrb.2020.01.007

68. Gonzalez-Soto M, Mutch DM. Diet regulation of long-chain PUFA synthesis: role of macronutrients, micronutrients, and polyphenols on Δ-5/Δ-6 desaturases and elongases 2/5. *Adv Nutr* (2021) 12(3):980–94. doi: 10.1093/advances/nmaa142

69. Hosseinkhani S, Dehghanbanadaki H, Aazami H, Pasalar P, Asadi M, Razi F. Association of circulating omega 3, 6 and 9 fatty acids with gestational diabetes mellitus: a systematic review. *BMC Endocrine Disord* (2021) 21(1):120. doi: 10.1186/s12902-021-00783-w

70. Bragg F, Kartsonaki C, Guo Y, Holmes M, Du H, Yu C, et al. Circulating metabolites and the development of type 2 diabetes in Chinese adults. *Diabetes Care* (2021) 45(2):477–80. doi: 10.2337/dc21-1415

71. Zulyniak MA, Fuller H, Iles MM. Investigation of the causal association between long-chain n-6 polyunsaturated fatty acid synthesis and the risk of type 2 diabetes: a mendelian randomization analysis. *Lifestyle Genomics* (2020) 13(5):146–53. doi: 10.1159/000509663

72. Pan X-F, Huang Y, Li X, Wang Y, Ye Y, Chen H, et al. Circulating fatty acids and risk of gestational diabetes mellitus: prospective analyses in China. *Eur J Endocrinol* (2021) 185(1):87–97. doi: 10.1530/eje-21-0118

73. *MedlinePlus. APOB gene national library of medicine (US): NIH* (2021). Available at: https://medlineplus.gov/genetics/gene/apob/.

74. Mensink RP, Zock PL, Kester AD, Katan MB. Effects of dietary fatty acids and carbohydrates on the ratio of serum total to HDL cholesterol and on serum lipids and apolipoproteins: a meta-analysis of 60 controlled trials. *Am J Clin Nutr* (2003) 77(5):1146–55. doi: 10.1093/ajcn/77.5.1146

75. Jacobson TA, Glickstein SB, Rowe JD, Soni PN. Effects of eicosapentaenoic acid and docosahexaenoic acid on low-density lipoprotein cholesterol and other lipids: a review. *J Clin Lipidol.* (2012) 6(1):5–18. doi: 10.1016/j.jacl.2011.10.018

76. Ooi EM, Ng TW, Watts GF, Barrett PH. Dietary fatty acids and lipoprotein metabolism: new insights and updates. *Curr Opin Lipidol.* (2013) 24(3):192–7. doi: 10.1097/MOL.0b013e3283613ba2

77. Schoenaker DA, Mishra GD, Callaway LK, Soedamah-Muthu SS. The role of energy, nutrients, foods, and dietary patterns in the development of gestational diabetes mellitus: a systematic review of observational studies. *Diabetes Care* (2016) 39(1):16–23. doi: 10.2337/dc15-0540

78. Koletzko B, Reischl E, Tanjung C, Gonzalez-Casanova I, Ramakrishnan U, Meldrum S, et al. FADS1 and FADS2 polymorphisms modulate fatty acid metabolism and dietary impact on health. *Annu Rev Nutr* (2019) 39:21–44. doi: 10.1146/annurev-nutr-082018-124250

79. Ameur A, Enroth S, Johansson A, Zaboli G, Igl W, Johansson AC, et al. Genetic adaptation of fatty-acid metabolism: a human-specific haplotype increasing the biosynthesis of long-chain omega-3 and omega-6 fatty acids. *Am J Hum Genet* (2012) 90(5):809–20. doi: 10.1016/j.ajhg.2012.03.014

80. Hu J, Gillies CL, Lin S, Stewart ZA, Melford SE, Abrams KR, et al. Association of maternal lipid profile and gestational diabetes mellitus: a systematic review and meta-analysis of 292 studies and 97,880 women. *EClinicalMedicine* (2021) 34:100830. doi: 10.1016/j.eclinm.2021.100830

81. Makshood M, Post WS, Kanaya AM. Lipids in south asians: epidemiology and management. *Curr Cardiovasc Risk Rep* (2019) 13(8):24. doi: 10.1007/s12170-019-0618-9

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 18 May 2023
DOI 10.3389/fimmu.2023.1148107



OPEN ACCESS

EDITED BY
Youxin Wang,
Capital Medical University, China

REVIEWED BY
Marta Kaczmarek-Ryś,
Polish Academy of Sciences, Poland
Zhipeng Zheng,
Zhejiang Chinese Medical University, China

*CORRESPONDENCE
Guang Yang
zyeygzr@163.com
Binbin Xie
xieBB@zju.edu.cn

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

RECEIVED 19 January 2023
ACCEPTED 08 May 2023
PUBLISHED 18 May 2023

CITATION
Xu D, Chen Y, Gao X, Xie W, Wang Y, Shen J, Yang G and Xie B (2023) The genetically predicted causal relationship of inflammatory bowel disease with bone mineral density and osteoporosis: evidence from two-sample Mendelian randomization.
*Front. Immunol.* 14:1148107.
doi: 10.3389/fimmu.2023.1148107



# The genetically predicted causal relationship of inflammatory bowel disease with bone mineral density and osteoporosis: evidence from two-sample Mendelian randomization

Dengyong Xu[1†], Yao Chen[2†], Xing Gao[3†], Weidong Xie[4], Ya Wang[5,6,7], Jiaying Shen[2], Guang Yang[8*] and Binbin Xie[2*]

[1]Department of Colorectal Surgery, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [2]Department of Medical Oncology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [3]Department of Oncology, The Second Affiliated Hospital of Soochow University, Suzhou, China, [4]Department of Gastrointestinal Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [5]Institute of Cancer and Basic Medicine (ICBM), Chinese Academy of Sciences, Hangzhou, China, [6]Department of Hospital Infection-Control, Cancer Hospital of the University of Chinese Academy of Sciences, Hangzhou, China, [7]Department of Hospital Infection-Control, Zhejiang Cancer Hospital, Hangzhou, China, [8]Department of Orthopedic Surgery, The Second Affiliated Hospital of Zhejiang University, Hangzhou, China

**Background:** Many existing studies indicated that patients with inflammatory bowel disease (IBD), including ulcerative colitis (UC) and Crohn's disease (CD), tend to have the risk of low total body bone mineral density (BMD), and are more likely to have osteoporosis (OS). To determine the causal relationship between IBD and bone metabolic disorders, we herein performed a two-sample Mendelian randomization analysis (TSMR) using publicly available summary statistics.

**Methods:** Summary statistics of total body BMD, OS and IBD were downloaded from the Open Genome-Wide Association Study (GWAS), FinnGen consortium and International Inflammatory Bowel Disease Genetics Consortium (IIBDGC). The European and East Asian populations have consisted in this Mendelian Randomization (MR) work. A range of quality control procedures were taken to select eligible instrument SNPs closely associated with total body BMD, OS and IBD. To make the conclusions more reliable, we applied five robust analytical methods, among which the inverse variance weighting (IVW) method acted as the major method. Besides, heterogeneity, pleiotropy and sensitivity were evaluated.

**Results:** In the European population, the genetic association of UC on total body BMD (OR=0.97, 95%CI=0.96,0.99, P<0.001) and overall IBD on total body BMD (OR=0.98, 95%CI=0.97,1.00, P=0.013) were significant, while the effect of CD on total body BMD was not significant enough (OR=0.99, 95%CI=0.98,1.00, P=0.085). All of UC, CD and overall IBD can be the genetic risk factor of having OS with pathological fracture (UC: OR=1.13, 95%CI=1.02,1.26, P=0.024,

CD: OR=1.14, 95%CI=1.05,1.25, P=0.003, overall IBD: OR=1.13, 95% CI=1.02,1.24, P=0.015). In East Asian groups, only CD had a causal relationship with OS (OR=1.04, 95% CI=1.01,1.07, P=0.019).

**Conclusion:** Our study revealed genetically predicted associations between IBD on total body BMD and OS in European and East Asian populations. This work supplemented the results of previous retrospective studies and demonstrated the necessity of BMD monitoring in patients with IBD.

KEYWORDS

ulcerative colitis, Crohn's disease, inflammatory bowel disease, bone mineral density, osteoporosis, Mendelian randomization

# Introduction

Ulcerative colitis (UC) and Crohn's disease (CD) are major forms of inflammatory bowel diseases (IBD), that mainly occurred in children and young adults (1). There is also some evidence showing that the incidence and prevalence of UC and CD in the older population are increasing in recent years (2). The characteristic of chronic and continuous inflammation of the intestines contributes to the susceptibility of colitis-associated colorectal cancer, posing a threat to the life quality and life span of IBD patients (3, 4). Besides, IBD patients are always accompanied by lower total body bone mineral density (BMD), which indicates that they are more likely to suffer from osteoporosis (OS) with or without pathological fracture or even other skeletal diseases (5–7).

The most typical features of OS are both bone quality and quantities being impaired. Meanwhile, osteoporotic fractures lead to a great disease burden in the USA with more than 25 billion dollars in predicted annual cost in 2025 (8). Fractures in some common vulnerable but important sites, represented by the hip, vertebral column and ankle, dramatically raise the risk of mortality among OS patients (9). Notably, BMD serves as a significant criterion for diagnosing and measuring the severity of OS (10). Although quantities of clues linked IBD with low total body BMD and OS, the bidirectional causal relationship between them is still ambiguous. Thus, the exploration of the genetically predicted causal relationship of IBD with BMD and OS using a novel method is urgently required.

Mendelian randomization (MR) analysis, a method of epidemiological analysis, can bolster causal inferences by employing genetic variation as an instrumental variable (IV) of the exposure factor (11). In addition to avoiding irrelevant confounders like environmental exposures, MR analysis can also lessen the impact of reverse causality, enhancing the plausibility of the causal inference (12).. In this two-sample MR work, we tried to figure out the genetically predicted causal association between IBD and skeletal diseases. Limited by the available datasets, we conducted the impact of IBD on total body BMD and OS with pathological fracture (FG) in the European population. And the impact of IBD on OS was performed in the East Asian population.

# Methods

## Study design

There are three main assumptions to be satisfied in the MR analysis: (1) instrumental variables (IVs) are closely related to exposure; (2) IVs are independent of any possible confounders; (3) IVs only affect the outcome through exposure (Figure 1). The above principles are the core of MR analysis. Two-sample MR was performed in this work, which requires two different genetic datasets to be consistent in one certain MR analysis. Of note, the data used in this work are publicly available and free to global researchers, so there was no need to provide further ethical approval and informed consent here.

## Data sources

Genetic association with the total body BMD was obtained from the meta-analysis of 30 genome-wide association studies (GWASs) of total body BMD, consisting of 56,284 European cases and 16,162,733 SNPs (13). The related summary statistics were available in the Open GWAS database (https://gwas.mrcieu.ac.uk/). Summary data of association with UC were derived from International Inflammatory Bowel Disease Genetics Consortium (IIBDGC). IIBDGC is the biggest global inflammatory bowel disease genetics database, in which European UC-associated SNPs were obtained from 13,768 cases and 33,977 controls, European CD-associated SNPs were obtained from 5,956 cases and 14,927 controls, and European overall IBD-associated SNPs were obtained from 31,665 cases and 33,977 controls (14, 15). East Asian datasets of IBD were also downloaded from IIBDGC, with 1,134 cases/3,719 controls in UC, 1,690 cases/3,719 controls in CD, and 2,824 cases/3,719 controls in overall IBD. The summary statistic of European OS with FG was gotten from the FinnGen database including 785 cases, 172,834 controls and 16,380,281 SNPs. And the data of East Asian OS were downloaded from Biobank Japan (GWAS ID: bbj-a-137; 7,788 cases/204,665 controls). To avoid genetic bias derived from ethnic differences, all the TSMR results are for the corresponding ethnicity only.

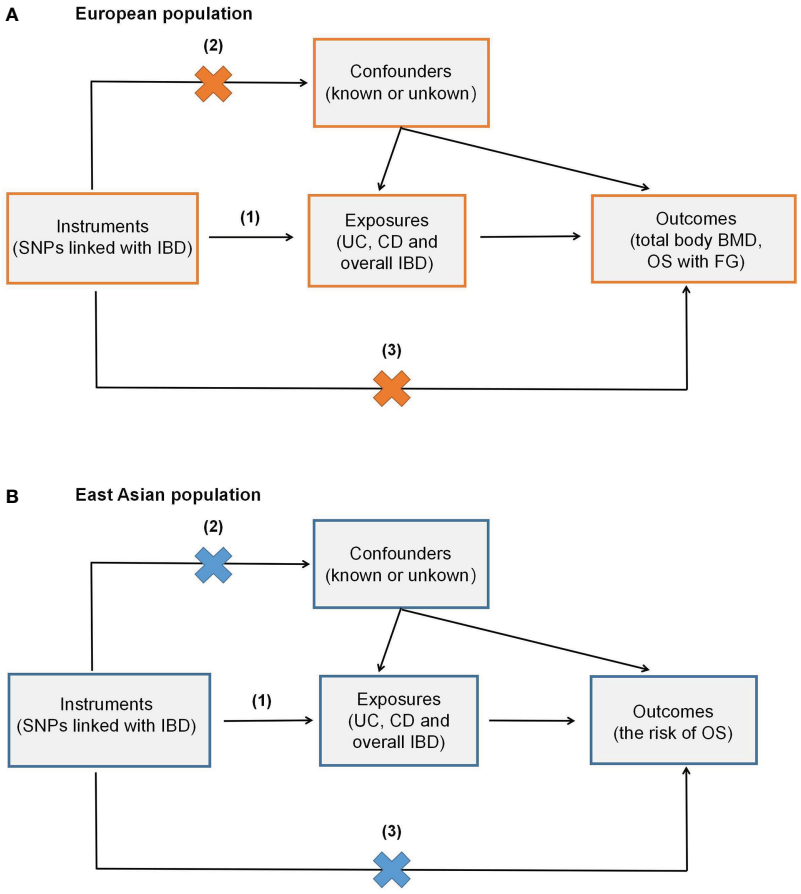



FIGURE 1
Overview of the study design. **(A)** shows the study investigating the genetic effect of IBD on total body BMD and OS with FG in the European population, while **(B)** denotes the genetic impact of IBD on OS in the East Asian population. (1) Assumption one: instrumental variables (IVs) are closely related to exposure; (2) Assumption two: IVs are independent of known or unknown confounders; (3) Assumption three: IVs only affect the outcome via exposure.

## Selection of genetic instruments

We selected the genetic instruments (IVs) according to the following inclusion criteria: (1) Single-nucleotide polymorphisms (SNPs) should be strongly associated with exposure, with the genome-wide significance level ($P < 5\times10^{-8}$); (2) Genetic variants with LD ($r^2 > 0.001$) were eliminated. To clump the independence of SNPs, the linkage disequilibrium (LD) between selected SNPs was evaluated; (3) The F-statistics ($beta^2/se^2$) > 10. The F-statistics were calculated to assess the intensity of IVs. If the F-statistics of certain SNPs were less than 10 may designate less statistical power (16). All the IVs applied in this study were summarized in Supplementary Tables 1–6.

## Statistical analysis

In our study, two-sample Mendelian randomization (TSMR) analyses were conducted through the TwoSampleMR package (version 0.5.6) with R software (version 4.2.1) (17). After IVs being selected, the IVs' data associated with both exposure and outcome was harmonized. The primary approach of MR analysis is the inverse-variance-weighted (IVW) method, which uses weighted regression of SNP-specific Wald ratios to evaluate the causal effects of genetically predicted exposure on outcome. Moreover, four other sensitivity assessment approaches were conducted simultaneously, namely Weighted median, MR Egger, Simple mode and Weighted mode, to test the consistency and heterogeneity of our results (18, 19). We also applied MR-PRESSO to evaluate pleiotropy and identify the outliers (20). The leave-one-out method was used to analyze the sensitivity of MR studies.

# Results

## Selected genetic instruments

We selected the IVs strictly in accordance with the criteria described above. As a result, 88, 53 and 134 independent SNPs were selected to be the IVs of European UC, CD and overall IBD respectively. And the amounts of IVs representing East Asian UC, CD and overall IBD were 10, 14 and 11 (details in Supplementary Tables 1–6). There was no evidence of weak instrumental bias shown by F statistics, which were also listed in the supplementary data.

## The causal effect of IBD on total body BMD, OS in the European population

In the current MR study, outlier SNPs were excluded by the MR-PRESSO method. Therefore, the final SNPs used to conduct TSMR were less than the amounts of IVs above. The TSMR results of the IVW method and SNPs number were displayed in Table 1. Scatter plots were presented in Figure 2. All three exposures (UC, CD and IBD) demonstrated a negative genetic impact on total body BMD. However, the UC (OR=0.97, 95%CI=0.96,0.99, P<0.001) and overall IBD (OR=0.98, 95%CI=0.97,1.00, P=0.013) showed a significant association with total body BMD, while the genetic predicted effect of CD on total body BMD was not significant enough (OR=0.99, 95%CI=0.98,1.00, P=0.085). Hence, in the European population, UC and overall IBD were the causal risk factors for total body BMD.

As for the genetic effect on OS with FG, 6,5 and 5 outlier SNPs in UC, CD and IBD were excluded in the following TSMR. All three exposures indicated a significant positive association with OS with FG: UC (OR=1.13, 95%CI=1.02,1.26, P=0.024), CD (OR=1.14, 95% CI=1.05,1.25, P=0.003), overall IBD (OR=1.13, 95% CI=1.02,1.24, P=0.015).

## The causal effect of IBD on OS in the East Asian population

To make clear the impact of IBD on OS in different ethnic populations, we conducted TSMR in the East Asian population. After performing the MR-PRESSO test, 7,14 and 10 SNPs were included to uncover the genetically predicted relationship of UC, CD and IBD on OS in East Asian groups. Interestingly, the results showed a significant causal effect of CD on OS (OR=1.04, 95% CI=1.01,1.07, P=0.019). While the genetic association was not significant in UC and overall IBD groups: UC on OS (OR=0.99, 95% CI=0.93,1.06, P=0.80), IBD on OS (OR=1.04, 95% CI=0.99,1.10, P=0.14). Scatter plots of IBD on OS in East Asian people were presented in Figure 3. Thus, we concluded that CD could serve as a genetically predicted causal risk factor of OS in the East Asian population, while UC and overall IBD could not.

## Sensitivity analysis

To approve the weak IV bias in the selected IVs of UC, CD and overall IBD, the F-statistic was calculated and ensured to be larger than 10. To validate the robustness of the above MR analyses, we performed multiplicity, heterogeneity and sensitivity analyses (details in Supplementary Table 7). The MR-Egger intercept and MR-PRESSO tests demonstrated no horizontal multiplicity in any of the above analyses (all P>0.05). In addition, heterogeneity was found in the heterogeneity tests of MR-Egger and IVW methods of some analyses, which are acceptable in the MR study. The results of leave-one-out sensitivity analyses indicated that the estimates of the causal effects of genetically predicted UC, CD and IBD on BMD and OS were robust (details in Supplementary Figures 1–9).

# Discussion

In adults with no underlying chronic disease, OS is an age-related abnormality of bone metabolism, which is most frequent in postmenopausal women (21). Nevertheless, previous retrospective studies have found that patients with IBD, consisting of children, adolescents and adults, tend to have lower BMD and are more likely to develop osteoporosis with or without pathological fractures (7). A population-based matched cohort study revealed that IBD patients have a 40% higher incidence to get fractures than people

TABLE 1 The inverse variance weighting (IVW) method results of the TSMR analyses.

| Exposure | Outcome | nSNP | Beta | SE | OR | OR_lCI95 | OR_uCI95 | P_value |
|---|---|---|---|---|---|---|---|---|
| European | | | | | | | | |
| UC | BMD | 85 | -0.026 | 0.008 | 0.974 | 0.959 | 0.990 | <0.001* |
| CD | BMD | 51 | -0.010 | 0.006 | 0.990 | 0.979 | 1.001 | 0.085 |
| IBD | BMD | 131 | -0.018 | 0.007 | 0.982 | 0.968 | 0.996 | 0.013* |
| UC | OS with FG | 82 | 0.122 | 0.054 | 1.130 | 1.016 | 1.257 | 0.024* |
| CD | OS with FG | 48 | 0.133 | 0.044 | 1.142 | 1.047 | 1.246 | 0.003* |
| IBD | OS with FG | 129 | 0.120 | 0.049 | 1.128 | 1.024 | 1.242 | 0.015* |
| East Asian | | | | | | | | |
| UC | OS | 7 | -0.008 | 0.032 | 0.992 | 0.932 | 1.056 | 0.801 |
| CD | OS | 14 | 0.038 | 0.016 | 1.038 | 1.006 | 1.072 | 0.019* |
| IBD | OS | 10 | 0.039 | 0.026 | 1.040 | 0.988 | 1.095 | 0.135 |

An asterisk was placed if the p value of TSMR analysis was less than 0.05, which considered as significant. The BMD here refers to total body bone mineral density. SE, standard error; OR, odds ratio; CI, confidence interval; OR_lCI95, the lower 95%CI of OR; OR_uCI95, the upper 95%CI of OR; UC, ulcerative colitis; CD, Crohn's disease; IBD, inflammatory bowel disease; OS, osteoporosis; FG, pathological fracture.

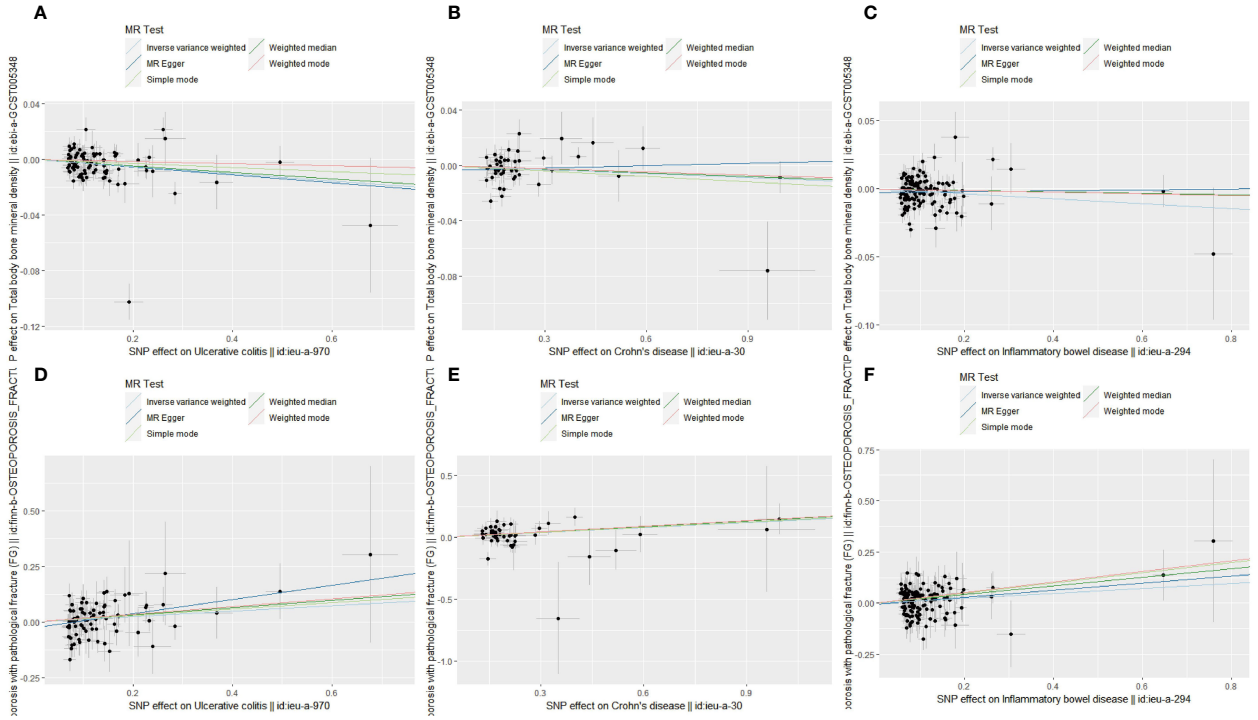



FIGURE 2
Scatter plots of MR analyses show the statistical relationship between IBD and total body BMD and the risk of OS with FG in the European population. Each dot represents an instrumental SNP. The x-axis indicates the genetic relationship with exposures (UC, CD and IBD), while the y-axis reflects the genetic association with outcomes (total body BMD and OS with FG). **(A)** UC on total body BMD: the genetically predicted UC is associated with a lower level of total body BMD; **(B)** CD on total body BMD: the genetically predicted CD is associated with a lower level of total body BMD; **(C)** overall IBD on total body BMD: the genetic predicted overall IBD is associated with a lower level of total body BMD; **(D)** UC on OS with FG: the genetic predicted UC is associated with a higher risk of OS with FG; **(E)** CD on OS with FG: the genetic predicted CD is associated with a higher risk of OS with FG; **(F)** overall IBD on OS with FG: the genetic predicted overall IBD is associated with a higher risk of OS with FG. The slope of each line shows the estimated causal effect of IBD on the bone metabolic status for each approach.

who don't have IBD (22). Therefore, the mechanisms of how IBD affects bone metabolism is remaining a heated research topic in these decades.

In the present work, we applied TSMR analyses to find out the genetic predicted association between overall IBD (including UC and CD) and total body BMD or OS. Two different samples of each exposure and outcome which have different data structures were included to avoid possible confounding factors (23). According to our TSMR results, causal relationships differed in various ethnic groups. In the European population, UC and overall IBD acted as

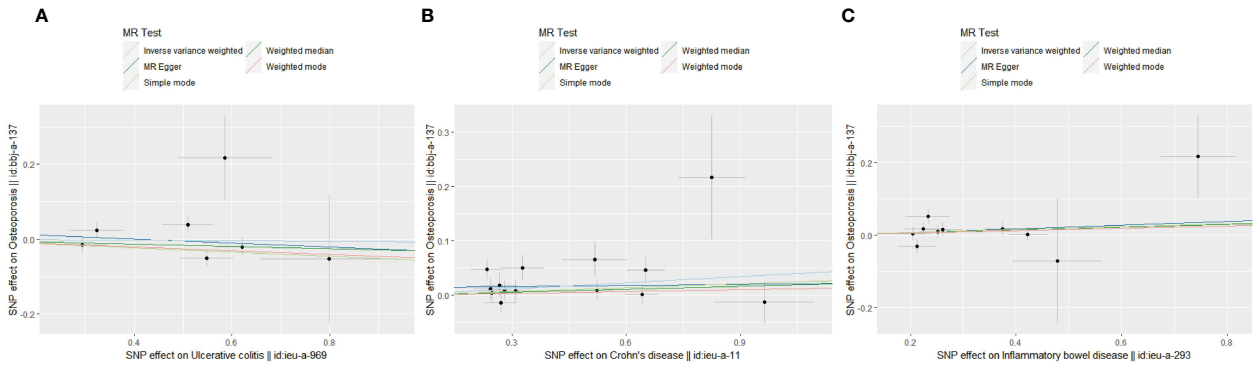



FIGURE 3
Scatter plots of MR analyses show the statistical relationship between IBD and OS in the East Asian population. Each dot represents an instrumental SNP. The x-axis indicates the genetic relationship with exposures (UC, CD and IBD), while the y-axis reflects the genetic association with outcomes (OS). **(A)** UC on OS: the genetically predicted UC is associated with a higher risk of OS; **(B)** CD on OS: the genetically predicted CD is associated with a higher risk of OS; **(C)** overall IBD on OS: the genetically predicted overall IBD is associated with a higher risk of OS. The slope of each line shows the estimated causal effect of IBD on OS for each approach.

risk factors for total body BMD, and all three factors (UC, CD and overall IBD) had a positive causal correlation with the risk of OS accompanied by FG. As for the East Asian group, CD was the only exposure that had a positive causal relationship with OS.

There were varieties of acquired mechanisms leading to low BMD in IBD populations. Firstly, inflammations in IBD patients can be one of the most important contributors to induce bone metabolic disorders. Pro-inflammatory cytokines are engaged in the inflammation of IBD, leading to persistent gastrointestinal inflammation and tissue destruction, as well as modulating bone defects (24, 25). Secondly, nutritional deficiencies are quite common in IBD populations, which play an essential role in bone loss (26, 27). Furthermore, glucocorticoids have long been considered first-line therapeutic agents for patients with IBD (28). In addition to their conventional anti-inflammatory function, glucocorticoids have been reported to enhance intestinal epithelial barrier function (29). Hence, strategies attempting to maximize the gastrointestinal local benefits of steroids while minimizing the systemic negative effects are urgently required.

Research was increasingly focusing on the relationship between the gut microbiota and OS. It appeared that there was a bidirectional interaction between bone metabolism and gut microbiota. In rats, OS has been demonstrated to cause intestinal dysfunction. Additionally, reprogramming intestinal function through the gut-bone axis, drugs potentially modulate bone metabolism (30, 31). By producing metabolites, which enter the systemic circulation from the gastrointestinal tracts, microorganisms were known to have an impact on bone (32). Given the mystery and complexity of the gut microbiota and gut-bone axis, there are still many unknown fields to be investigated.

Previous studies reported several genetic factors which may determine bone metabolism in IBD populations. The variation of *IL-6* and *IL1-ra* genes acted as independent determinants of bone loss in patients with IBD (33). As revealed, the extent of bone loss is positively correlated with the amount of variation in specific genes mentioned above. Another gene ties to the secretion level of IL-1β, called *IL1B-511*2*, whose polymorphism could also be a predictor for the risk of osteopenia or osteoporosis in IBD patients (34). Low bone mass is more likely to be present in IBD among carriers of the *IL1B-511*2* gene, who are also along with IL-1beta hypersecretion. Besides, a GWAS study performed in the Japanese population pointed out that the *SLC22A23* gene and the *MECOM* gene determined distinct genetic risk factors for bone loss in IBD patients from those in the general population (35). The genetic-related research above suggested that genetic variability plays an important role in bone metabolism in IBD.

According to the harmonizing results of the present MR study, we identified the SNPs with the strongest associations with total body BMD and OS, rs780094 (OR=1.03, 95%CI=1.02,1.04, P=2.54E-07) and rs34779708 (OR=8.40, 95%CI=7.56,9.34, P=1.24E-03), respectively (Supplementary Tables 8 and 9). Based on the search results of GWAS database, rs780094 was located on *GCKR* gene (Location: 2:27518370; Cytogenetic region:2p23.3), and rs34779708 was located on *CREM* gene (Location: 10:35177257; Cytogenetic region:10p11.21). Therefore, we targeted the known functional and disease associations of these two genes to explore whether IBD has an overlapping genetic background with BMD and OS.

cAMP responsive element modulator (*CREM*) is a gene encoding the bZIP transcription factor that binds to cAMP response elements in many viral and cellular promoters. Associated with immune system regulation, *CREM* has been implicated to play a role in various immune-mediated inflammatory processes. Aberrant expression of *CREM* was detected in patients with SLE. The expression product of *CREM* could promote the expression of inflammatory factors represented by IL-2 and IL-17 (36, 37), which in turn mediate the associated inflammatory processes. Also, *CREM* impacts T-cell activity in homeostasis and participates in the regulation of NF/κB signaling pathway (38). In addition to its role in inflammation and autoimmune diseases, Taiwanese scholars found that *CREM* expression was increased in patients with rheumatoid arthritis and its variants were involved in the progression of the disease (39). Single-cell sequencing studies showed that *CREM* regulons were most active in preosteoblast-S1 (40), implying that *CREM* variants are engaged in osteogenesis and bone homeostasis.

Glucokinase regulator (*GCKR*) gene encodes the glucokinase regulator, which is involved in metabolic modifications and is closely associated with skeletal and inflammatory phenotypes. Kasher et al. identified SNPs of the *GCKR* gene were consistently appeared in the osteoporosis phenotypes and C-reactive protein by co-localization analysis (41). The known coverage of *CREM* and *GCKR* genes reveals, to some extent, a common genetic background between IBD and bone metabolism. The effects of these loci might be diverse in UC and CD, which leads to different genetic effects observed in UC and CD subtypes.

This work is not perfect with some limitations to be mentioned here. First, based on the TSMR analysis design, this study can only provide genetic evidence for the causal relationship between IBD and OS. As a consequence, this finding can supplement earlier retrospective investigations and provide guidance for upcoming observational studies. Besides, the study was limited by the specific ethnicity. The derivation of analyses of the different ethnic populations will not be valid. It was therefore hard to generalize the result to other ethnic groups.

Some strengths also need to be highlighted. Large sample sizes in the datasets helped to reduce the bias that population stratification may cause. All of the information was gathered from reliable organizations or websites. Moreover, OS often affects older people, whereas IBD affects teenagers more frequently. Consequently, individuals with OS in IBD must be excluded from typical baseline investigations, which force applied a considerable amount of time and effort. In contrast, the current study avoided the aforementioned issues by assessing the genetic causation of the two using strongly correlated SNPs as IVs.

## Conclusion

This MR work revealed that in European population, UC and overall IBD had negative causal relationships on total body BMD, and UC, CD and overall IBD had positive causal relationships on

OS with FG. While only CD showed a positive causal association with OS in East Asian population. Our results complemented former retrospective investigations and served as a reference for future animal research and clinical treatments.

## 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

DX, XG and WX: Conceptualization and writing of the manuscript. YC and YW: Digging the data and making the tables and figures. BX, JS and GY: Review, editing and providing critical discussion. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by grants from National Natural Science Foundation of China (No. 82100542 to BX), Zhejiang Provincial Natural Science Foundation of China (No. LQ21H160027 to BX), the China Postdoctoral Science Foundation (No. 2020M681893 to BX), National Natural Science Foundation of China (No.81903160 to JS), Zhejiang Provincial Natural Science Foundation of China (No. LQ20H260002 to YW).

## 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/fimmu.2023.1148107/full#supplementary-material

## References

1. Ye Y, Manne S, Treem WR, Bennett D. Prevalence of inflammatory bowel disease in pediatric and adult populations: recent estimates from Large national databases in the united states, 2007-2016. *Inflammation Bowel Dis* (2020) 26(4):619–25. doi: 10.1093/ibd/izz182

2. Agrawal M, Christensen HS, Bøgsted M, Colombel JF, Jess T, Allin KH. The rising burden of inflammatory bowel disease in Denmark over two decades: a nationwide cohort study. *Gastroenterology* (2022) 163(6):1547–54. doi: 10.1053/j.gastro.2022.07.062

3. Wewer MD, Langholz E, Munkholm P, Bendtsen F, Seidelin JB, Burisch J. Disease activity patterns of inflammatory bowel disease - a Danish nationwide cohort study 1995-2018. *J Crohns Colitis* (2022) 17(3):329–37. doi: 10.1093/ecco-jcc/jjac140

4. Birch RJ, Burr N, Subramanian V, Tiernan JP, Hull MA, Finan P, et al. Inflammatory bowel disease-associated colorectal cancer epidemiology and outcomes: an English population-based study. *Am J Gastroenterol* (2022) 117 (11):1858–70. doi: 10.14309/ajg.0000000000001941

5. Lee JS, Lee HS, Jang BI, Kim ES, Kim SK, Kim KO, et al. Low bone mineral density in young patients newly diagnosed with inflammatory bowel disease. *Dig Dis Sci* (2021) 66(2):605–11. doi: 10.1007/s10620-020-06220-7

6. Shirazi KM, Somi MH, Rezaeifar P, Fattahi I, Khoshbaten M, Ahmadzadeh M. Bone density and bone metabolism in patients with inflammatory bowel disease. *Saudi J Gastroenterol* (2012) 18(4):241–7. doi: 10.4103/1319-3767.98428

7. Zhou T, Pan J, Lai B, Cen L, Jiang W, Yu C, et al. Bone mineral density is negatively correlated with ulcerative colitis: a systematic review and meta-analysis. *Clin Transl Med* (2020) 9(1):18. doi: 10.1186/s40169-020-00270-0

8. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the united states, 2005-2025. *J Bone Miner Res* (2007) 22(3):465–75. doi: 10.1359/jbmr.061113

9. Office of the Surgeon, G. Reports of the surgeon general. In: *Bone health and osteoporosis: a report of the surgeon general*. Rockville (MD: Office of the Surgeon General (US (2004).

10. Snodgrass P, Zou A, Gruntmanis U, Gitajn IL. Osteoporosis diagnosis, management, and referral practice after fragility fractures. *Curr Osteoporos Rep* (2022) 20(3):163–9. doi: 10.1007/s11914-022-00730-1

11. Ebrahim S, Davey Smith G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum Genet* (2008) 123(1):15–33. doi: 10.1007/s00439-007-0448-6

12. Zoccali C, Testa A, Spoto B, Tripepi G, Mallamaci F. Mendelian randomization: a new approach to studying epidemiology in ESRD. *Am J Kidney Dis* (2006) 47(2):332–41. doi: 10.1053/j.ajkd.2005.10.027

13. Medina-Gomez C, Kemp JP, Trajanoska K, Luan J, Chesi A, Ahluwalia TS, et al. Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects. *Am J Hum Genet* (2018) 102(1):88–102. doi: 10.1016/j.ajhg.2017.12.005

14. Liu JZ, Sommeren van S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* (2015) 47(9):979–86. doi: 10.1038/ng.3359

15. Goyette P, Boucher G, Mallon D, Ellinghaus E, Jostins L, Huang H, et al. High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nat Genet* (2015) 47(2):172–9. doi: 10.1038/ng.3176

16. Wang B, Liu D, Song M, Wang W, Guo B, Wang Y. Immunoglobulin G n-glycan, inflammation and type 2 diabetes in East Asian and European populations: a mendelian randomization study. *Mol Med* (2022) 28(1):114. doi: 10.1186/s10020-022-00543-z

17. Yavorska OO, Burgess S. MendelianRandomization: an r package for performing mendelian randomization analyses using summarized data. *Int J Epidemiol* (2017) 46(6):1734–9. doi: 10.1093/ije/dyx034

18. 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

19. 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

20. 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

21. Compston JE, McClung MR, Leslie WD. Osteoporosis. *Lancet* (2019) 393 (10169):364–76. doi: 10.1016/S0140-6736(18)32112-3

22. Bernstein CN, Blanchard JF, Leslie W, Wajda A, Yu BN. The incidence of fracture among patients with inflammatory bowel disease. a population-based cohort study. *Ann Intern Med* (2000) 133(10):795–9. doi: 10.7326/0003-4819-133-10-200011210-00012

23. Inoue A, Solon G. Two-sample instrumental variables estimators. *Rev Economics Stat* (2010) 92(3):557–61. doi: 10.1162/REST_a_00011

24. Tilg H, Moschen AR, Kaser A, Pines A, Dotan I. Gut, inflammation and osteoporosis: basic and clinical concepts. *Gut* (2008) 57(5):684–94. doi: 10.1136/gut.2006.117382

25. Neurath MF. Cytokines in inflammatory bowel disease. *Nat Rev Immunol* (2014) 14(5):329–42. doi: 10.1038/nri3661

26. Aghdassi E, Wendland BE, Stapleton M, Raman M, Allard JP. Adequacy of nutritional intake in a Canadian population of patients with crohn's disease. *J Am Diet Assoc* (2007) 107(9):1575–80. doi: 10.1016/j.jada.2007.06.011

27. Tucker KL. Osteoporosis prevention and nutrition. *Curr Osteoporos Rep* (2009) 7(4):111–7. doi: 10.1007/s11914-009-0020-5

28. Mowat C, Cole A, Windsor A, Ahmad T, Arnott I, Driscoll R, et al. Guidelines for the management of inflammatory bowel disease in adults. *Gut* (2011) 60(5):571–607. doi: 10.1136/gut.2010.224154

29. Greenhill C. IBD: glucocorticoids revealed to augment intestinal epithelial barrier function. *Nat Rev Gastroenterol Hepatol* (2014) 11(2):75. doi: 10.1038/nrgastro.2013.255

30. Liu J, Liu J, Liu L, Zhang G, Peng X. Reprogrammed intestinal functions in astragalus polysaccharide-alleviated osteoporosis: combined analysis of transcriptomics and DNA methylomics demonstrates the significance of the gut-bone axis in treating osteoporosis. *Food Funct* (2021) 12(10):4458–70. doi: 10.1039/D1FO00113B

31. Fernández-Murga ML, Olivares M, Sanz Y. Bifidobacterium pseudocatenulatum CECT 7765 reverses the adverse effects of diet-induced obesity through the gut-bone axis. *Bone* (2020) 141:115580. doi: 10.1016/j.bone.2020.115580

32. Zaiss MM, Jones RM, Schett G, Pacifici R. The gut-bone axis: how bacterial metabolites bridge the distance. *J Clin Invest* (2019) 129(8):3018–28. doi: 10.1172/JCI128521

33. Schulte CM, Dignass AU, Goebell H, Röher HD, Schulte KM. Genetic factors determine extent of bone loss in inflammatory bowel disease. *Gastroenterology* (2000) 119(4):909–20. doi: 10.1053/gast.2000.18158

34. Nemetz A, Tóth M, García-González MA, Zágoni T, Fehér J, Peña AS, et al. Allelic variation at the interleukin 1beta gene is associated with decreased bone mass in patients with inflammatory bowel diseases. *Gut* (2001) 49(5):644–9. doi: 10.1136/gut.49.5.644

35. Naito T, Yokoyama N, Kakuta Y, Ueno K, Kawai Y, Onodera M, et al. Clinical and genetic risk factors for decreased bone mineral density in Japanese patients with inflammatory bowel disease. *J Gastroenterol Hepatol* (2018) 33(11):1873–81. doi: 10.1111/jgh.14149

36. Rauen T, Hedrich CM, Juang YT, Tenbrock K, Tsokos GC. cAMP-responsive element modulator (CREM)alpha protein induces interleukin 17A expression and mediates epigenetic alterations at the interleukin-17A gene locus in patients with systemic lupus erythematosus. *J Biol Chem* (2011) 286(50):43437–46. doi: 10.1074/jbc.M111.299313

37. Xu WD, Zhang YJ, Wang W, Li R, Pan HF, Ye DQ. Role of CREM in systemic lupus erythematosus. *Cell Immunol* (2012) 276(1-2):10–5. doi: 10.1016/j.cellimm.2012.04.008

38. Lv S, Li J, Qiu X, Li W, Zhang C, Zhang ZN, et al. A negative feedback loop of ICER and NF-kappaB regulates TLR signaling in innate immune responses. *Cell Death Differ* (2017) 24(3):492–9. doi: 10.1038/cdd.2016.148

39. Tseng CC, Lin YZ, Lin CH, Hwang DY, Li RN, Tsai WC, et al. Genetic and epigenetic alterations of cyclic AMP response element modulator in rheumatoid arthritis. *Eur J Clin Invest* (2022) 52(5):e13715. doi: 10.1111/eci.13715

40. Wang S, Gong Y, Wang Z, Meng X, Luo Z, Papasian CJ, et al. Regulon active landscape reveals cell development and functional state changes of human primary osteoblasts in vivo. *Hum Genomics* (2023) 17(1):11. doi: 10.1186/s40246-022-00448-2

41. Kasher M, Williams FMK, Freidin MB, Malkin I, Cherny SSCIW Group, et al. Understanding the complex genetic architecture connecting rheumatoid arthritis, osteoporosis and inflammation: discovering causal pathways. *Hum Mol Genet* (2022) 31(16):2810–9. doi: 10.1093/hmg/ddac061

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 25 May 2023
DOI 10.3389/fendo.2023.1180338



OPEN ACCESS

EDITED BY
Haifeng Hou,
Shandong First Medical University, China

REVIEWED BY
Qiuxia Han,
Nankai University, China
Wei Liang,
Renmin Hospital of Wuhan University, China
Zhangzhe Peng,
Central South University, China
Rui Zeng,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Fengmin Shao
fengminshao@126.com

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

RECEIVED 06 March 2023
ACCEPTED 12 May 2023
PUBLISHED 25 May 2023

CITATION
Ren X, Wang W, Cao H and Shao F (2023) Diagnostic value of serum cathepsin S in type 2 diabetic kidney disease. *Front. Endocrinol.* 14:1180338. doi: 10.3389/fendo.2023.1180338



# Diagnostic value of serum cathepsin S in type 2 diabetic kidney disease

Xuejing Ren[1,2†], Wanqing Wang[3†], Huixia Cao[1] and Fengmin Shao[1,2*]

[1]Henan Key Laboratory of Kidney Disease and Immunology, Henan Provincial Clinical Research Center for Kidney Disease, Zhengzhou University People's Hospital, Henan Provincial People's Hospital, Zhengzhou, Henan, China, [2]Academy of Medical Sciences, Zhengzhou, Henan, China, [3]Health Management Centre, People's Hospital of Zhengzhou University, Central China Fuwai Hospital, Zhengzhou, Henan, China

**Background:** Identification of risk factors that have causal effects on the occurrence of diabetic kidney disease (DKD), is of great significance in early screening and intervening for DKD, and in delaying the progression of DKD to end-stage renal disease. Cathepsin S (Cat-S), a novel non-invasive diagnostic marker, mediates vascular endothelial dysfunction. The diagnostic value of Cat-S for DKD has rarely been reported in clinical studies.

**Objective:** To analyze whether Cat-S is a risk factor for DKD and evaluate the diagnostic value of serum Cat-S for DKD.

**Methods:** Forty-three healthy subjects and 200 type 2 diabetes mellitus (T2DM) patients were enrolled. T2DM patients were divided into subgroups according to various criteria. Enzyme-linked immunosorbent assay was used to detect serum Cat-S levels among different subgroups. Spearman correlation analysis was used to analyze correlations between serum Cat-S and clinical indicators. Multivariate logistic regression analysis was performed to analyze risk factors for the occurrence of DKD and decreased renal function in T2DM patients.

**Results:** Spearman analysis showed that serum Cat-S level was positively correlated with urine albumin creatinine ratio (r=0.76, $P<0.05$) and negatively correlated with estimated glomerular filtration rate (r=−0.54, $P<0.01$). Logistic regression analysis showed that increased serum Cat-S and cystatin C(CysC) were independent risk factors for DKD and decreased renal function in T2DM patients ($P<0.05$). The area under the receiver operating characteristic (ROC) curve was 0.900 of serum Cat-S for diagnosing DKD, and when the best cut-off value was 827.42 pg/mL the sensitivity and specificity were 71.6% and 98.8%, respectively. Thus, serum Cat-S was better than CysC for diagnosing DKD (for CysC, the area under the ROC curve was 0.791, and when the cut-off value was 1.16 mg/L the sensitivity and specificity of CysC were 47.4% and 98.8%, respectively).

**Conclusion:** Increased serum Cat-S were associated with the progression of albuminuria and decreased renal function in T2DM patients. The diagnostic value of serum Cat-S was better than that of CysC for DKD. Monitoring of serum Cat-S

levels could be helpful for early screening DKD and assessing the severity of DKD and could provide a new strategy for diagnosing DKD.

KEYWORDS

type 2 diabetes mellitus, diabetic kidney disease, cathepsin S, risk factor, diagnostic value

# 1 Introduction

Diabetic kidney disease (DKD) is one of the most common and serious microvascular complications of diabetes, with the incidence as high as 20%-40% in patients with type 2 diabetes mellitus (T2DM) (1). There are usually no obvious symptoms and signs in the early stage of DKD, which leads to delayed treatment, losing the best opportunity for early intervention and progressing to end-stage renal disease (ESRD) gradually. Therefore, the identification of risk factors with causal effects on the incidence of DKD is crucial for the prevention of the onset of DKD and also lays a foundation for achieving the early screening of DKD, reducing the morbidity rate and delaying progression of DKD to ESRD.

Renal biopsy pathology is the gold standard for the diagnosis of DKD, but it is not suitable for diagnosing early DKD because of the risk of bleeding. Therefore, non-invasive screening for diagnosing early DKD is more valuable for clinical studies. Some studies have suggested serum cystatin C (CysC) as an indicator for diagnosing early DKD, but this is controversial (2–4). Therefore, it is of great importance to explore novel non-invasive biomarkers for early diagnosing DKD.

Cathepsin S (Cat-S) was a secreted cysteine proteolytic enzyme that is mainly expressed in macrophages (5). Macrophages undergoing chemotaxis adhere to the basement membrane of blood vessels and secrete Cat-S, and the secreted Cat-S was involved in hydrolysis of elastin, laminin, collagen and other extracellular matrix components, causing vascular damage (6). The deficiency of Cat-S gene or the activity of Cat-S was inhibited were strongly associated with neovascularization (7). In recent years, studies have shown that upregulation of Cat-S was associated with the development of IgA nephropathy, lupus nephritis, insulin resistance, diabetes and other renal diseases (8–10). Monocyte/macrophage-derived Cat-S has been found to activate protease-activated receptor-2 on glomerular endothelial cells, causing endothelial damage, albumin leakage, inflammation and glomerulosclerosis (11). Inhibition of Cat-S has been shown to reduce infiltration of renal inflammatory cells and downregulate the expression of inflammatory cytokines in kidney tissues in a DKD mouse model, as well as delaying the progression of DKD (11). However, few clinical studies have evaluated the correlation between serum Cat-S and the diagnostic value for DKD.

Vascular injury was among the main mechanisms in the pathogenesis of DKD (12). Therefore, detecting the expression of Cat-S which is a vascular injury marker and determining whether serum Cat-S in T2DM patients is a risk factor for the incidence of DKD are of great significance for preventing the incidence of DKD, strengthening early screening for DKD, and searching for new therapeutic targets for the treatment of DKD in T2DM patients. Therefore, in this study, we detected the expression of serum Cat-S in DKD patients, analyzed the correlation between serum Cat-S and the severity of DKD and analyzed whether serum Cat-S was a risk factor for DKD in T2DM patients, and investigated the diagnostic value for DKD, to provide a theoretical basis for early screening of DKD.

# 2 Materials and methods

## 2.1 Participants

A total of 200 patients with T2DM aged between 18 and 80 years were selected as subjects. A further 43 healthy subjects who underwent physical examination in Henan Provincial People's Hospital during the same period were selected as the healthy control group, which was matched with the enrolled T2DM participants in gender and age (18 females and 25 males, with an average age of 53.72 ± 9.95 years). The healthy controls had no history of diabetes, kidney disease or any major diseases, and did not meet any of the diagnostic criteria for diabetes. This study was approved by the Ethics Committee of Henan Provincial People's Hospital [approval number: (2021) No. (153)], and all research subjects signed informed consents.

## 2.2 Inclusion criteria

Patients were eligible for inclusion if they: (1) aged between 18 and 80 years old; (2) met the diagnostic criteria for diabetes mellitus (DM) formulated by the ADA in 2020 (13) and/or the Chinese Guidelines for Clinical Diagnosis and Treatment of Diabetic Kidney Disease in 2021 (14); (3) had complete clinical data available; (4) signed the informed consent form voluntarily.

## 2.3 Exclusion criteria

Participants were excluded if they (1) had other types of diabetes; (2) had diabetes combined with acute or chronic infection; (3) had diabetes combined with acute cardiovascular or cerebrovascular diseases; (4) had diabetes combined with chronic obstructive pulmonary disease and aspartate aminotransferase or

alanine transaminase ≥2 times the upper limit of normal simultaneously; (5) had complications involving other renal diseases; (6) were women who were pregnant or lactating; (7) were affected by diabetic ketoacidosis, hypoglycemia coma, lactic acidosis, hyperosmolar coma, surgery, trauma or other stress states; (8) had undergone dialysis or renal transplantation.

## 2.4 Grouping scheme

In grouping mode 1, T2DM patients were divided into three groups according to urine albumin creatinine ratio (UACR) which reflects the severity of albuminuria: T2DM group (UACR <30 mg/g, n=84); early DKD group (UACR 30–300 mg/g, n=41) and clinical DKD group (UACR >300 mg/g, n=75). In grouping mode 2, T2DM patients were divided into two groups according to estimated glomerular filtration rate (eGFR): normal renal function group (eGFR ≥90 mL/min/1.73 $m^2$, n=141) and decreased renal function group (eGFR <90mL/min/1.73 $m^2$, n=59).

## 2.5 Clinical data collection

The patients' gender, age, body mass index (BMI), admission systolic blood pressure (SBP), diastolic blood pressure (DBP), fasting blood glucose (FBG), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), retinol-binding protein (RBP), blood urea nitrogen (BUN), serum creatinine (Scr), serum uric acid (SUA), CysC and other clinical data were collected. The eGFR was calculated using the Modification of Diet in Kidney Disease equation: eGFR = 186 × age (Scr/88.4) $^{-1.154}$ × age$^{-0.203}$ (women × 0.742) (15).

## 2.6 Sample collection and Cat-S detection

All patients were required to fast for more than 8 h. In the early morning of the following day, 5 mL of non-anticoagulant venous blood was taken on an empty stomach and allowed to coagulate at room temperature for 2 h. After centrifugation (1000×g for 15 min at 4°C), the serum was collected and stored in a freezer at −80 °C until use. Serum Cat-S concentration (pg/mL) was detected by enzyme-linked immunosorbent assay (ELISA; Wuhan Huamei Biological Engineering, Co., Ltd.) and the operation was carried out strictly according to the instructions. Finally, 5 mL of morning urine was collected for the determination of UACR by immunoturbidimetry.

## 2.7 Statistical analysis

Measurement data conforming to a normal distribution were expressed as mean ± standard deviation ($\bar{X} \pm \sigma$), and measurement data with non-normal distribution were expressed as M (Q1, Q3). As this study involved comparisons of multiple measurement data, a homogeneity of variance test was performed. The homogeneity of variance was analyzed by one-way analysis of variance, and the multiple comparison LSD was used for the comparison between samples. Non-parametric test was used for uneven variance, Kruskal–Wallis test was used for multiple group comparisons, and Bonferroni correction was used for pairwise comparisons. Qualitative data were compared between groups with chi-square test. Spearman rank correlation analysis was used to analyze the correlations between serum Cat-S level and clinical indicators. Multivariate logistic regression was used to analyze the risk factors for DKD in T2DM patients. The AUC (area under the receiver operating characteristic curve (ROC)) was used to analyze the diagnostic value of serum Cat-S and CysC in early DKD. All tests were two-tailed and $P<0.05$ was considered to indicate statistical significance. All statistical analysis were performed by SPSS 25.0.

# 3 Results

## 3.1 Comparison of clinical parameters and serum Cat-S under grouping mode 1

Forty-three healthy subjects (healthy control group), 84 T2DM patients (T2DM group), 41 early DKD patients (early DKD group), and 75 clinical DKD patients (clinical DKD group) were enrolled in this study. There were significant differences in BMI, FBG, HbAlC, SUA, Scr, eGFR, BUN, UACR, TG, TC, HDL-C and LDL-C among the four groups ($P<0.05$). There were no significant difference in the ratio of male to female and age among the four groups ($P>0.05$). There were significant differences in SBP, DBP, CysC and RBP among the T2DM group, early DKD group and clinical DKD group (Table 1). Serum Cat-S concentrations in each group were as follows: healthy control group: 312.51 (216.58,476.86) pg/mL; T2DM group: 476.45 (257.98,616.72) pg/mL; early DKD group: 684.61 (468.21,834.20) pg/mL; clinical DKD group: 1688.71 (1236.28, 2227.65) pg/mL. There were significant differences in serum Cat-S concentration among the T2DM group, early DKD group and clinical DKD group ($P<0.05$). The serum Cat-S level in the T2DM group was higher than that in the healthy control group, but the difference was not statistically significant ($P>0.05$) (Table 1).

## 3.2 Correlations between serum Cat-S and clinical parameters

The serum Cat-S level was positively correlated with UACR (r=0.76), RBP, BUN, Scr, TC, SUA, CysC (r=0.548) ($P<0.05$) and was negatively correlated with eGFR (r=−0.54) ($P<0.001$). The serum CysC level was positively correlated with UACR (r=0.604), RBP, BUN, Scr, TG, SUA, Cat-S (r=0.548) ($P<0.05$) and was negatively correlated with eGFR (r=−0.779) ($P<0.001$) (Table 2).

## 3.3 Univariate logistic regression analysis of all parameters under grouping mode 1

Taking the occurrence of DKD or not as the dependent variable (assignment: T2DM=1, early DKD=2, clinical DKD=3) and SBP,

TABLE 1 Comparison of clinical indicators and serum Cat-S under grouping mode 1.

| Projects | Healthy control group | T2DM group | Early DKD group | Clinical DKD group | *P Value* |
|---|---|---|---|---|---|
| Gender (female/male)[a] | 25/18 | 32/52 | 14/27 | 26/49 | 0.960 |
| Age(year)[b] | 53.72 ± 9.95 | 56.26 ± 10.73 | 57.78 ± 10.72 | 53.69 ± 9.71 | 0.113 |
| BMI(Kg/m$^2$)[b] | 24.20 (23.00, 25.20) | 25.09 (23.44, 27.61)* | 26.11 (23.84, 28.58)* | 26.41 (24.22, 29.06)* | < 0.001 |
| Duration of DM (years)[c] | 0 | 9.5(4.0~16.75) | 9.0(4.0~18.0) | 10.0(3.0~16.0) | 0.793 |
| SBP (mmHg)[b] | 124.6 ± 7.85 | 131.57 ± 16.67 | 134.80 ± 14.75 | 147.15 ± 18.70$^{\$\&}$ | < 0.001 |
| DBP (mmHg)[b] | 74.77 ± 8.42 | 75.31 ± 10.98 | 77.05 ± 12.43 | 81.90 ± 3.55$^{\$}$ | 0.001 |
| FBG (mmol/L)[c] | 4.91(4.60, 5.15) | 7.37(5.89, 9.07)* | 8.18(6.27, 9.83)* | 7.79(5.29, 9.17)* | < 0.001 |
| HbAlC(mmol/L)[c] | 5.40(4.90, 5.60) | 8.00(7.23, 9.80)* | 8.90(7.50,10.65)* | 8.00(7.18, 9.00) * | < 0.001 |
| SUA (μmol/L)[c] | 295(241, 340) | 275(227, 357) | 280(215, 346) | 358 (289, 425) $^{*\$\&}$ | < 0.001 |
| Scr (μmol/L)[c] | 60.0(55.0, 67.0) | 57.5(44.0, 64.0) | 55.0(48.0, 62.5) | 95.0(68.0,156.0)$^{*\$\&}$ | < 0.001 |
| eGFR mL/min/(1.73 m$^2$)[c] | 111.62 (96.12, 126.74) | 132.45 (109.63, 158.73)* | 135.56 (108.91, 152.09)* | 66.28 (36.39, 101.24)$^{*\$\&}$ | < 0.001 |
| BUN (mmol/L)[c] | 4.86(4.13, 5.82) | 5.83(4.87, 7.00)* | 5.52(4.83, 6.39) | 8.97(6.11,12.20)$^{*\$\&}$ | < 0.001 |
| UACR (mg/g)[c] | 10.2(6.4, 11.0) | 10.0(6.4, 17.1) | 92.4(53.9,188.1)$^{*\$}$ | 3271.0(1378.0, 6012.5)$^{*\$\&}$ | < 0.001 |
| TG (mmol/L)[c] | 1.20(0.90, 1.58) | 1.42(1.08, 2.28) | 1.64(1.19, 2.59)* | 1.81(1.46, 2.36)$^{*\$}$ | < 0.001 |
| TC (mmol/L)[c] | 4.38(3.69, 5.12) | 4.58(3.85, 5.46) | 4.19(3.49, 5.10) | 5.18(4.32, 6.59)$^{*\$\&}$ | < 0.001 |
| HDL-C (mmol/L)[c] | 1.37(1.17, 1.52) | 1.10(0.98, 1.29)* | 1.06(0.90, 1.36)* | 1.08(0.88, 1.31)* | < 0.001 |
| LDL-C (mmol/L)[c] | 2.92(2.14, 3.59) | 2.73(2.10, 3.40) | 2.43(1.93, 3.08) | 2.88(2.41, 3.96)$^{\&}$ | 0.011 |
| Cat-S (pg/mL)[c] | 312.51 (216.58,476.86) | 476.45 (257.98,616.72) | 684.61 (468.21,834.20)$^{*\$}$ | 1688.71 (1236.28, 2227.65)$^{*\$\&}$ | < 0.001 |
| CysC (mg/L)[c] | – | 0.84(0.75, 0.93) | 0.88(0.84, 0.96) | 1.57(1.04, 2.30)$^{\$\&}$ | < 0.001 |
| RBP (mg/L)[c] | – | 38.0(33.9, 46.2) | 39.2(33.8, 46.7) | 53.9(46.0, 66.3) $^{\$\&}$ | < 0.001 |

BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; FBG, fasting blood glucose; HbAlC, glycosylated hemoglobin; SUA, blood uric acid; Scr, serum creatinine; eGFR, estimated glomerular filtration rate; BUN, blood urea nitrogen; UACR, urinary albumin/creatinine ratio; TG, triglyceride; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; Cat-S, Cathepsin S; CysC, cystatin C; RBP, retinol-binding protein; Data are presented as( x ± s)or M (1/4, 3/4) unless indicated; *: P < 0.05 compared with healthy controls; $: Comparison with T2DM group P < 0.05; &: comparison with early DKD group P < 0.05; [a] $\chi 2$ test; [b] one-way analysis of variance; [c] rank sum test.

TABLE 2 Spearman analysis between serum Cat-S, CysC and clinical parameters (n=200).

| Projects | Cat-S | | CysC | |
|---|---|---|---|---|
| | r | *P Value* | r | *P Value* |
| UACR (mg/g) | 0.76 | < 0.001 | 0.604 | < 0.001 |
| eGFR mL/min/(1.73 m$^2$) | -0.54 | < 0.001 | -0.779 | < 0.001 |
| RBP (mg/L) | 0.38 | < 0.001 | 0.610 | < 0.001 |
| BUN (mmol/L) | 0.40 | < 0.001 | 0.609 | < 0.001 |
| Scr (μmol/L) | 0.53 | < 0.001 | 0.749 | < 0.001 |
| FBG (mmol/L) | -0.08 | 0.28 | -0.127 | 0.074 |
| TC (mmol/L) | 0.14 | 0.04 | 0.108 | 0.126 |
| TG (mmol/L) | 0.12 | 0.09 | 0.216 | 0.002 |
| LDL-C (mmol/L) | 0.07 | 0.35 | 0.076 | 0.282 |

*(Continued)*

TABLE 2 Continued

| Projects | Cat-S | | CysC | |
|---|---|---|---|---|
| | r | *P Value* | r | *P Value* |
| HDL-C (mmol/L) | -0.08 | 0.27 | -0.159 | 0.025 |
| SUA (μmol/L) | 0.25 | < 0.001 | 0.451 | < 0.001 |
| Cat-S (pg/mL) | – | – | 0.548 | < 0.001 |
| CysC (mg/L) | 0.548 | < 0.001 | – | – |

DBP, SUA, TC, TG, LDL-C, eGFR, BUN, RBP, Cat-S and CysC as the independent variables. Univariate logistic regression analysis showed that: Cat-S may be an influential factor on the occurrence of early DKD ($P$<0.001); and SBP, DBP, SUA, TC, LDL-C, eGFR, BUN, RBP, Cat-S and CysC may be influential factors on the occurrence of clinical DKD ($P$<0.01) (Table 3).

## 3.4 Multivariate logistic regression analysis of serum Cat-S and CysC under grouping mode 1

The data of this study did not meet the requirements of the ordered logistic regression parallelism test, so multivariate logistic regression analysis was used to analyze the risk factors for early DKD and clinical DKD. Taking whether DKD occurs or not as the dependent variable (assignment: T2DM=1, early DKD=2, clinical DKD=3), the results showed that after correcting SBP, DBP, SUA, TC, eGFR, BUN, RBP and other factors, we found that Cat-S and CysC were independent risk factors for DKD and clinical DKD in T2DM patients ($P$<0.05). For every 0.1 unit increase in Cat-S, the relative risks of early DKD and clinical DKD in T2DM patients increased by 1.541 times and 5.690 times, respectively. For every 10 units increase in CysC, the relative risks of early DKD and clinical DKD in T2DM increased by 1.611 and 2.880 times, respectively (Table 4).

## 3.5 Logistic regression analysis of serum Cat-S on DKD

When UACR ≥ 30 mg/g was defined as the standard for the diagnosis of DKD, UACR < 30 mg/g was defined as DKD (-) group, while UACR ≥ 30 mg/g was defined as DKD (+) group. Taking whether DKD occurs or not as the dependent variable (assignment: DKD (-) =1, DKD (+) =2) and Age, Weight, Hight, SBP, DSP, Cat-S, BMI, HbAlC, eGFR, BUN, SUA, FBG, CysC, RBP, TC, TG, LDL-C, LDL-C and other factors as independent variables, the univariate logistic regression analysis showed that SBP, DSP, eGFR, BUN, SUA, RBP, TC, Cat-S, CysC may be the influencing factors of DKD (+) (P < 0.05) (Table 2). Taking whether DKD occurs or not as the dependent variable (assignment: DKD (-) =1, DKD (+) =2), the multivariate logistic regression analysis was performed and the results showed: after adjusting for SBP, DSP, eGFR, BUN, SUA, RBP, TC and other factors, we found the increased Cat-S and CysC were risk factors for the onset of DKD (P < 0.05). The relative risk of DKD in T2DM patients increased by 1.626 times for every 0.1 unit increase in Cat-S. For every 10 units of CysC, the relative risk of DKD in T2DM increased by 1.657 times (Table 5).

TABLE 3 Univariate Logistic regression analysis of all parameters under grouping mode 1.

| Independent variable | Eraly DKD | | Clinical DKD | |
|---|---|---|---|---|
| | *OR (95%CI)* | *P Value* | *OR (95%CI)* | *P Value* |
| SBP (mmHg) | 1.012(0.989~1.036) | 0.298 | 1.055(1.033~1.078) | < 0.001 |
| DPB (mmHg) | 1.01(0.981~1.049) | 0.409 | 1.05(1.023~1.086) | 0.001 |
| SUA (μmol/L) | 0.99(0.993~1.003) | 0.377 | 1.00(1.004~1.012) | < 0.001 |
| TC (mmol/L) | 0.86(0.633~1.190) | 0.378 | 1.59(1.243~2.051) | < 0.001 |
| TG (mmol/L) | 1.02(0.846~1.238) | 0.810 | 1.01(0.857~1.193) | 0.898 |
| LDL-C (mmol/L) | 0.76(0.506~1.160) | 0.202 | 1.58(1.157~2.161) | 0.004 |
| eGFR mL/min/(1.73 $m^2$) | 0.99(0.989~1.007) | 0.661 | 0.96(0.949~0.971) | < 0.001 |
| BUN (mmol/L) | 0.88(0.712~1.102) | 0.278 | 1.50(1.285~1.762) | < 0.001 |
| RBP (mg/L) | 0.98(0.952~1.010) | 0.194 | 1.04(1.020~1.067) | < 0.001 |
| Cat-S (0.1ng/dL) | 1.50(1.242~1.814) | < 0.001 | 2.90(2.131~3.957) | < 0.001 |
| CysC (10mg/L) | 1.28(0.997~1.648) | 0.053 | 2.09(1.619~2.722) | < 0.001 |

TABLE 4 Multivariate Logistic regression of Cat-S and CysC under grouping mode 1.

| | serum Cat-S (0.1ng/L) | | serum CysC (10mg/L) | |
|---|---|---|---|---|
| | *OR (95%CI)* | *P Value* | *OR (95%CI)* | *P Value* |
| T2DM | 1 | | 1 | |
| Eraly DKD | 1.541(1.244~1.910) | <0.001 | 1.611(1.041~2.492) | 0.032 |
| Clinical DKD | 5.69(1.856~17.506) | 0.002 | 2.880(1.175~7.062) | 0.021 |

TABLE 5 Multivariate logistic regression analysis of serum Cat-S on DKD.

| Independent variable | Univariate | | Multivariate | |
|---|---|---|---|---|
| | *OR* | *P Value* | *OR* | *P Value* |
| SBP (mmHg) | 1.038(1.019, 1.057) | <0.001 | 1.004(0.972, 1.037) | 0.824 |
| DSP (mmHg) | 1.039(1.012, 1.067) | 0.04 | 1.030(0.986, 1.076) | 0.180 |
| eGFR mL/min/(1.73 $m^2$) | 0.980(0.973, 0.987) | <0.001 | 1.004(0.990, 1.018) | 0.588 |
| BUN (mmol/L) | 1.278(1.129, 1.447) | <0.001 | 0.825(0.627, 1.084) | 0.167 |
| SUA (μmol/L) | 1.005(1.001, 1.008) | 0.006 | 0.999(0.994, 1.005) | 0.821 |
| RBP (mg/L) | 1.022(1.002, 1.042) | 0.027 | 0.995(0.975, 1.016) | 0.657 |
| TC (mmol/L) | 1.304(1.056, 1.610) | 0.014 | 1.180(0.820, 1.699) | 0.374 |
| Cat-S (0.1 ng/dL) | 1.700(1.435, 2.014) | <0.001 | 1.626(1.335, 1.981) | <0.001 |
| CysC (10 mg/L) | 1.640(1.339, 2.009) | <0.001 | 1.657(1.111, 2.470) | 0.013 |

## 3.6 Diagnostic value of serum Cat-S and CysC in clinical DKD under grouping mode 1

UACR >300 mg/g was defined as the standard for the diagnosis of clinical DKD. The ROC curve analysis gave a result of AUC which was 0.978 for serum Cat-S for diagnosing clinical DKD, and when the cut-off value of serum Cat-S was 974.14 pg/mL, the sensitivity was 96% and the specificity was 96% for diagnosing clinical DKD ($P$<0.001). While the AUC of CysC was 0.874 for diagnosing clinical DKD, and when the cut-off value of CysC was 1.16 mg/L, the sensitivity was 72% and the specificity was 98% ($P$<0.001). Results indicated that serum Cat-S was superior to CysC in the diagnosis of clinical DKD. The AUC for the combined diagnosis of clinical DKD was 0.991, the sensitivity increased to 98% and the specificity was 96% (Figure 1 and Table 6).

## 3.7 Diagnostic value of serum Cat-S and CysC in DKD under grouping mode 1

Defining UACR ≥30 mg/g as the standard for the diagnosis of early DKD, we assessed the diagnostic value of serum Cat-S and CysC based on the AUC of the ROC curve. The AUC for Cat-S in the diagnosis of DKD was 0.900, the cut-off value was 827.42 pg/mL, and the sensitivity and specificity were 71.6% and 98.8%, respectively. The AUC for CysC in the diagnosis of DKD was 0.791, and, when the cut-off value was 1.16mg/L, the sensitivity and specificity of CysC in the diagnosis of DKD were 47.4% and 98.8%, respectively (Figure 2 and Table 7).

## 3.8 Comparison of clinical parameters and serum Cat-S under grouping mode 2

There was no significant difference in gender, age, BMI, FBG or HDL-C between the normal renal function group and decreased renal function group ($P$>0.05). The SBP, DBP, SUA, Scr, BUN, UACR, TC, TG, LDL-C, Cat-S and CysC in the normal renal function group were significantly lower than those in the decreased renal function group ($P$<0.05) (Table 8).

## 3.9 Logistics regression analysis of renal function decline in T2DM patients under grouping mode 2

Taking whether eGFR declined or not as the dependent variable (assignment: eGFR ≥90 mL/min/(1.73 $m^2$)=0, eGFR <90 mL/min/(1.73 $m^2$)=1) and SBP, DBP, SUA, BUN, UACR, TC, LDL-C, Cat-S, CysC and RBP as independent variables, univariate logistic regression analysis was performed using the "enter" method. The results identified SBP, DBP, SUA, BUN, UACR, TC, LDL-C, Cat-S, CysC and RBP as potential risk factors for eGFR <90mL/min/(1.73 $m^2$) ($P$<0.05). We took the significantly meaningful variables (SBP, DBP, SUA, BUN, UACR, TC, LDL-C, Cat-S, CysC and RBP) from the univariate logistic regression analysis as independent variables

10.3389/fendo.2023.1180338

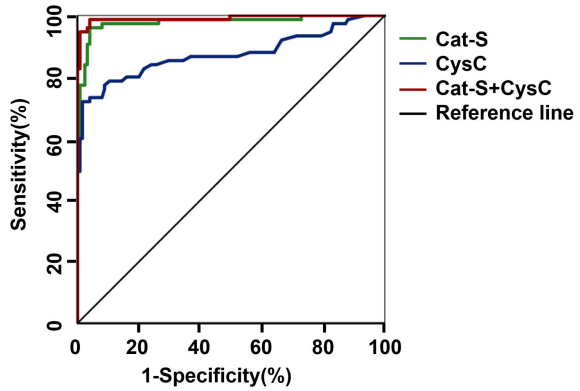



FIGURE 1
ROC curve of serum Cat-S and CysC in diagnosing clinical DKD. Cat-S, Cathepsin S; CysC, cystatin C.

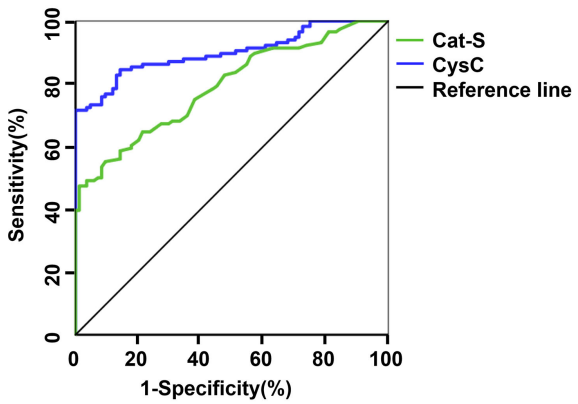



FIGURE 2
ROC curve of serum Cat-S and CysC in diagnosing DKD. Cat-S, Cathepsin S; CysC, cystatin C.

for multivariate logistic regression analysis by the "backward: LR" method. The results showed that increased SBP (odds ratio (OR) =1.044, 95% CI=1.004–1.087, *P*=0.033), increased BUN (OR=1.674, 95% CI=1.228–2.280, *P*=0.001), increased TC (OR=1.956, 95% CI=1.324–2.889, *P*=0.001), increased Cat-S (OR=2.835, 95% CI=1.260–6.381, *P*=0.012) and elevated CysC (OR=1.345, 95% CI=1.116–1.620, *P*=0.002) were independent risk factors for eGFR <90 mL/min/(1.73 $m^2$). Moreover, the relative risk of eGFR <90 mL/min/(1.73 $m^2$) increased with increasing Cat-S concentration; that is, the higher the levels of SBP, BUN, TC, Cat-S and CysC, the higher the probability of eGFR <90 mL/min/(1.73 $m^2$) (Table 9).

## 3.10 Diagnostic value of serum Cat-S and CysC in T2DM with decreased renal function

When eGFR <90 mL/min/(1.73 $m^2$) was defined as the diagnostic criterion for renal function decline. The ROC analysis gave an AUC of 0.854 for serum Cat-S in diagnosing decreased renal function of DKD, and when the cut-off value of serum Cat-S was 974.14 pg/mL, the sensitivity and specificity of serum Cat-S in diagnosing DKD were 85% and 80% respectively. While the AUC of serum CysC was 0.937 for diagnosing decreased renal function of DKD; and when the cut-off value of serum CysC was 1.08mg/L, the sensitivity and specificity of serum CysC in the diagnosis of decreased renal function were 88% and 93%, respectively. These results indicated that CysC was superior to serum Cat-S in diagnosing renal function decline (eGFR <90mL/min/(1.73 $m^2$)) in T2DM patients. The combination of serum CysC and Cat-S in the diagnosis of renal function decline showed an improved diagnostic value, with the AUC increased to 0.945, the sensitivity to 90% and the specificity to 85% compared with diagnosing by serum Cat-S only (Table 10 and Figure 3).

# 4 Discussion

DKD is characterized by persistent increased albuminuria excretion and/or progressive decline in renal function, which eventually progress to ESRD and severely affects the quality of life of patients (16). We found that elevated serum Cat-S was associated with the progression of albuminuria and decreased renal function in T2DM patients and could be used to assess the severity of DKD, and presented better diagnostic value than traditional biomarker CysC for diagnosing DKD. These results lay a foundation for the exploration of novel biomarkers for diagnosing early DKD and also indicate potential possibility for the prevention of DKD, early diagnosis, early treatment, as well as improving prognosis of patients.

The pathogenesis of DKD was complex and involved multiple pathways including injury of glomerular endothelial cells, activation of transforming growth factor-β1 and inflammatory responses (17). Cat-S was a proteolytic enzyme that remains active in both acidic and neutral environments, and injection of recombinant Cat-S has been shown to damage glomerular endothelial cells to induce proteinuria and glomerulosclerosis in DKD mice (11). Cat-S was also found to promote renal fibrosis by regulating the TGF-β1/Smad pathway in TGF-β1-stimulated renal tubular epithelial cells, which resulting in decreased renal function (18). In addition, Cat-S

TABLE 6 The diagnostic value of serum Cat-S and CysC in clinical DKD under grouping mode 1.

| Subjects | AUC | *P Value* | *95%CI* | optimal cut-off value | Sensitivity% | Specificity% | youden index |
|---|---|---|---|---|---|---|---|
| Cat-S | 0.978 | < 0.001 | 0.957~1 | 974.14pg/mL | 0.96 | 0.96 | 0.920 |
| CysC | 0.874 | < 0.001 | 0.814~0.934 | 1.16mg/L | 0.72 | 0.98 | 0.704 |
| Cat-S+CysC | 0.991 | < 0.001 | 0.978~1 | – | 0.98 | 0.96 | 0.947 |

TABLE 7 The diagnostic value of serum Cat-S and CysC in DKD under grouping mode 1.

| Subjects | AUC | *P Value* | *95%CI* | optimal cut-off value | Sensitivity% | Specificity% | youden index |
|---|---|---|---|---|---|---|---|
| Cat-S | 0.900 | < 0.001 | 0.858~0.943 | 827.42pg/mL | 0.716 | 0.988 | 0.706 |
| CysC | 0.791 | < 0.001 | 0.731~0.852 | 1.16mg/L | 0.474 | 0.988 | 0.462 |

TABLE 8 Comparison of clinical indicators and serum Cat-S under grouping mode 2.

| Projects | Normal renal function group (n=141) | Decreased renal function group (n=59) | $Z/\chi^2/t$ | *P Value* |
|---|---|---|---|---|
| Gender(female/male) [a] | 47/94 | 25/34 | 1.475 | 0.225 |
| Age (year) [b] | 55.29 ± 10.56 | 56.32 ± 10.21 | 0.636 | 0.526 |
| BMI ($Kg/m^2$) [b] | 25.94 ± 3.33 | 26.73 ± 3.92 | 1.447 | 0.149 |
| SBP (mmHg) [b] | 133.37 ± 15.75 | 149.32 ± 19.68 | 5.528 | < 0.001 |
| DBP (mmHg) [b] | 76.89 ± 10.71 | 81.12 ± 13.03 | 2.383 | 0.018 |
| FBG (mmol/L) [c] | 7.45 (5.98, 9.33) | 7.00 (5.28, 9.29) | -1.421 | 0.155 |
| HbAlC (mmol/L) [c] | 8.40 (7.25, 9.90) | 8.00 (7.35, 8.90) | -1.644 | 0.100 |
| SUA (μmol/L) [c] | 281.0 (228.5, 347.0) | 374.0 (323.0, 447.0) | -5.754 | < 0.001 |
| Scr (μmol/L) [c] | 57.0(46.0, 64.0) | 124.0 (85.0, 193.0) | -10.663 | < 0.001 |
| eGFR mL/min/(1.73 $m^2$) [c] | 131.75 (110.42, 158.49) | 48.66 (28.88, 71.84) | -11.143 | < 0.001 |
| BUN (mmol/L) [c] | 5.65 (4.87, 6.71) | 9.57 (7.65, 13.58) | -8.388 | < 0.001 |
| UACR (mg/g) [c] | 26.30(9.85, 172.00) | 3896.70(1487.40, 6345.60) | -8.503 | < 0.001 |
| TC (mmol/L) [c] | 4.57 (3.72, 5.44) | 5.20 (4.32, 6.62) | -3.682 | < 0.001 |
| TG (mmol/L) [c] | 1.52 (1.18, 2.27) | 2.00 (1.46, 2.59) | -2.644 | 0.008 |
| HDL-C (mmol/L) [c] | 1.08 (0.96, 1.31) | 1.09 (0.91, 1.31) | -0.441 | 0.659 |
| LDL-C (mmol/L) [c] | 2.67 (2.02, 3.29) | 3.13 (2.49, 4.04) | -3.338 | 0.001 |
| Cat-S (pg/mL) [c] | 594.24(374.05, 787.46) | 1642.78(1107.57, 2192.67) | -7.889 | < 0.001 |
| CysC (mg/L) [c] | 0.87 (0.78, 0.96) | 1.78 (1.25, 2.37) | -9.740 | < 0.001 |
| RBP (mg/L) [c] | 39.0 (33.8, 47.0) | 58.8 (47.3, 71.3) | -7.823 | < 0.001 |

normal renal function group (eGFR≥90mL/min/1.73 $m^2$); decreased renal function group (eGFR<; 90mL/min/1.73 $m^2$); [a] $\chi 2$ test; [b] one-way analysis of variance; [c] rank sum test.

promoted the secretion of pro-inflammatory cytokines such as tumor necrosis factor-α and interleukin-1 and activated ELR + CXC chemokines, thereby recruiting inflammatory cells such as macrophages, which secreted Cat-S to exacerbate the inflammatory response in turn (19, 20). The above results suggest that Cat-S is involved not only in dysfunction of glomerular endothelial cells and injury of glomerular filtration barrier that leading to proteinuria, but also in renal fibrosis and inflammatory response which promoting the development of DKD and leading to the decline in renal function. The relationship between Cat-S and DKD has been limited to fundamental research, and few clinical studies have evaluated the relationship between serum Cat-S and DKD in patients with T2DM. We grouped T2DM patients according to UACR and eGFR to investigate the correlation between serum Cat-S level and the severity of DKD, and to assess the diagnostic value of serum Cat-S for diagnosing DKD and the role of serum Cat-S in the assessment of renal function.

Serum Cat-S levels in T2DM patients were closely related to the severity of DKD. In this study, DKD was staged according to UACR which reflected the level of albuminuria in T2DM patients. The results showed that the level of serum Cat-S in patients with different DKD stages tended to increase with the increasing UACR; that is, the serum Cat-S level was positively correlated with the level of UACR. A study involving 103 DM patients found that serum Cat-S levels were higher in DM patients than in healthy control group (21), which is consistent with the results of our study, but they did not analyze the relationship between serum Cat-S and the severity of DKD. T2DM patients were further grouped according to eGFR. We found serum Cat-S was negatively correlated with eGFR, indicating that the increase of serum Cat-S was concurrent with the decrease of renal function. Serum Cat-S has been found to be negatively correlated with eGFR in German chronic kidney disease (CKD) patients and Swedish community CKD patients (22), which were consistent with the results of this study. These results suggest that serum Cat-S in T2DM patients increases

TABLE 9 Logistics regression analysis of renal function decline in T2DM under grouping mode 2.

| Independent variable | Univariate | | Multivariate | |
|---|---|---|---|---|
| | *OR* | *P Value* | *OR* | *P Value* |
| SBP (mmHg) | 1.055 (1.034~1.077) | < 0.001 | 1.044 (1.004~1.087) | 0.033 |
| DBP (mmHg) | 1.033 (1.005~1.061) | 0.02 | 0.943 (0.885~1.003) | 0.064 |
| SUA (μmol/L) | 1.011 (1.007~1.015) | < 0.001 | 1.006 (0.999~1.014) | 0.093 |
| BUN (mmol/L) | 1.949 (1.589~2.390) | < 0.001 | 1.674 (1.228~2.280) | 0.001 |
| UACR (mg/g) | 1.001 (1.001~1.001) | < 0.001 | – | – |
| TC (mmol/L) | 1.630 (1.288, 2.063) | < 0.001 | 1.956 (1.324~2.889) | 0.001 |
| LDL-C (mmol/L) | 1.855 (1.359~2.531) | < 0.001 | – | – |
| Cat-S (0.1ng/mL) | 5.501 (3.264~9.272) | < 0.001 | 2.835 (1.260~6.381) | 0.012 |
| CysC (10mg/L) | 1.799 (1.489~2.174) | < 0.001 | 1.345 (1.116~1.620) | 0.002 |
| RBP (mg/L) | 1.058 (1.034~1.082) | < 0.001 | – | – |

TABLE 10 The diagnostic value of serum Cat-S and CysC in T2DM with renal function decline.

| Subjects | AUC | *P Value* | 95%CI | optimal cut-off value | Sensitivity | Specificity | youden index |
|---|---|---|---|---|---|---|---|
| Cat-S | 0.854 | < 0.001 | 0.797~0.911 | 974.14pg/mL | 0.85 | 0.80 | 0.632 |
| CysC | 0.937 | < 0.001 | 0.893~0.981 | 1.08mg/L | 0.88 | 0.93 | 0.810 |
| Cat-S+CysC | 0.945 | < 0.001 | 0.909~0.982 | – | 0.90 | 0.85 | 0.813 |

with the progression of DKD. Thus, serum Cat-S is expected to applicate as a novel biomarker for the early diagnosis of clinical DKD.

We grouped T2DM patients according to UACR which reflected the severity of albuminuria and according to eGFR, and found that the elevated serum Cat-S and CysC were independent risk factors for early DKD and decreased renal function in T2DM patients. Multivariate logistic regression analysis under grouping mode 1 showed that with an increase of 0.1 units of Cat-S, the relative risks of early DKD and clinical DKD increased 1.541 times and 5.690 times, respectively; whereas with an increase of 10 units of CysC, the relative risks of early DKD and clinical DKD increased by 1.611 times and 2.880 times, respectively. Multivariate logistic regression analysis under grouping mode 2 showed that elevated serum Cat-S and CysC were independent risk factors for eGFR < 90mL/min/(1.73 $m^2$); that is, the higher the serum Cat-S and CysC levels, the higher the probability of eGFR < 90mL/min/(1.73 $m^2$). In db/db mice with type 2 diabetes, Cat-S inhibitors or PAR2 inhibitors have been reported to reduce albuminuria and glomerulosclerosis, as well as other organ complications such as diabetic retinopathy (11). Therefore, we conclude that serum Cat-S was associated with the development of increased albuminuria and decreased renal function, and could be a potential new therapeutic target for the prevention of DKD.

Serum Cat-S performed valuable diagnostic efficacy in the diagnosis of DKD. Previous studies have suggested that serum CysC in T2DM patients was positively correlated with UACR (23), which was a sensitive indicator of renal impairment and could be used as a marker for early diagnosis of DKD (24). However, in this study, serum CysC increased significantly in T2DM patients with massive albuminuria, which indicating that serum CysC could not be used to diagnose early DKD effectively. Therefore, we further analyzed and compared the diagnostic value of serum Cat-S and serum CysC in DKD. When UACR ≥30mg/g was used as the criterion for the diagnosis of early DKD, ROC analyzed the diagnostic value of serum Cat-S showed that when the optimal cut-off value for serum Cat-S was 827.42 pg/mL, the diagnostic sensitivity and specificity of serum Cat-S were higher than those of serum CysC, suggesting that serum Cat-S has better diagnostic efficacy for the diagnosis of DKD. When eGFR < 90mL/

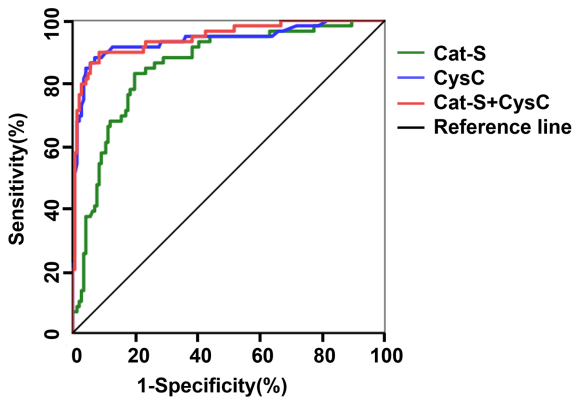



FIGURE 3
ROC curve of serum Cat-S and CysC in diagnosing renal function decline in T2DM.

min/(1.73 $m^2$) was used as the diagnostic criterion for decreased renal function, we found that serum Cat-S and serum CysC had similar diagnostic value for decreased renal function in T2DM patients. However, the combined diagnostic value of serum Cat-S and serum CysC for diagnosing decreased renal function in patients with T2DM was better than that of either of the two alone, suggesting that serum Cat-S level could reflect the decline of renal function of T2DM patients to a certain extent.

However, this study had some limitations. However, this study had some limitations. Researchers have found that strict control of blood glucose can delay the progression of DKD in T2DM patients (25). Considering that diabetes medications may be used, there were no differences in HbAlC between subgroups in this study. We did not count the application of hypoglycemic drugs, but it will be a new direction in our following study. Patney et al. demonstrated that hypertension accelerated the progression of renal disease and led to increased morbidity and mortality from cardiovascular complications in DKD patients (26). We found that elevated SBP and DBP were influential factors in the occurrence of early DKD and decreased renal function. Therefore, controlling cardiovascular complications such as blood pressure may be one of the directions to slow the progression of DKD. It will be an important direction to research DKD with other cardiovascular complications. In the follow-up study, we will explore the diagnostic value of Cat-S for diagnosing DKD with or without cardiovascular complications. It was a single-center cross-sectional study, and further verification is needed in a multi-center prospective study with a large sample size and we plan to follow up DKD patients to analyze the correlation between Cat-S and the prognosis of DKD. In addition, most DKD patients we enrolled didn't undergo renal biopsy, thus the relationship between Cat-S and renal pathology was not clarified in our study. It was a single-center cross-sectional study, and further verification is needed in a multi-center prospective study with a large sample size. In addition, most DKD patients were clinically confirmed but not pathologically confirmed because clinically confirmed DKD patients usually did not undergo renal biopsy, thus the relationship between Cat-S and renal pathology was not clarified in our study.

## 5 Conclusion

In conclusion, we concluded that the elevated serum Cat-S were associated with the progression of albuminuria and decreased renal function in T2DM patients. Serum Cat-S increased with the severity of albuminuria in patients with DKD, and the diagnostic value of serum Cat-S was better than that of CysC for diagnosing DKD. The level of serum Cat-S could reflect the decline of renal function in T2DM patients to a certain extent. As a risk factor that affects the incidence of DKD, serum Cat-S is expected to be a new biomarker for the early diagnosis and severity assessment of DKD.

## 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 the Ethics Committee of Henan Provincial People's Hospital [Approval number (2021): No (153)]. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

FS designed the study, WW collected the data, XR and WW analyzed the data. XR wrote the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was funded by the research grants the Key Science and Technology Project of Henan Province (No.201300310700). There have no conflicts or financial involvement in this study with any organization or entity with a financial interest.

## Acknowledgments

We thank the members of Ethics Committee of Henan Provincial People's Hospital and the Center for Clinical Data Research for their permission to conduct the experiment and helpful data for the research.

## 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

# References


1. Conserva F, Gesualdo L, Papale M. A systems biology overview on human diabetic nephropathy: from genetic susceptibility to post-transcr iptional and post-translational modifications. *J Diabetes Res* (2016) 2016:7934504. doi: 10.1155/2016/7934504

2. Zhang D, Ye S, Pan T. The role of serum and urinary biomarkers in the diagnosis of early diabetic nephropathy in patients with type 2 diabetes. *PeerJ* (2019) 7:e7079. doi: 10.7717/peerj.7079

3. Li WN, Wang JL, Ge LN, Shan JH, Zhang CB, Liu JP. Growth arrest-specific protein 6 (Gas6) as a noninvasive biomarker for early detection of diabetic nephropathy. *Clin Exp Hypertension* (2017) 39(4):382–7. doi: 10.1080/10641963.2017.1288739

4. Arceo ES, Dizon GA, Tiongco REG. Serum cystatin c as an early marker of nephropathy among type 2 diabetics: a meta-analysis. *Diabetes Metab Syndr* (2019) 13 (6):3093–7. doi: 10.1016/j.dsx.2019.11.007

5. Nikolic-Paterson DJ. Cathepsin s-dependent protease-activated receptor-2 activation: a new mechanism of endothelial dysfunction. *J Am Soc Nephrol* (2016) 27 (6):1577–9. doi: 10.1681/ASN.2015101162

6. Figueiredo JL, Aikawa M, Zheng C, Aaron J, Lax L, Libby P, et al. Selective cathepsin s inhibition attenuates atherosclerosis in apolipoprotein e-deficient mice with chronic renal disease. *Am J Pathol* (2015) 185(4):1156–66. doi: 10.1016/j.ajpath.2014.11.026

7. Wilkinson RD, Williams R, Scott CJ, Burden RE. Cathepsin s: therapeutic, diagnostic, and prognostic potential. *Biol Chem* (2015) 396(8):867–82. doi: 10.1515/hsz-2015-0114

8. Zhao JY, Yang YC, Wu YB. The clinical significance and potential role of cathepsin s in IgA nephropathy. *Front Pediatrics* (2021) 9:10. doi: 10.3389/fped.2021.631473

9. Rupanagudi KV, Kulkarni OP, Lichtnekert J, Darisipudi MN, Mulay SR, Schott B, et al. Cathepsin s inhibition suppresses systemic lupus erythematosus and lupus nephritis because cathepsin s is essential for MHC class II-mediated CD4 T cell and b cell priming. *Ann Rheumatic Diseases* (2015) 74(2):452–63. doi: 10.1136/annrheumdis-2013-203717

10. Jobs E, Risérus U, Ingelsson E, Sundström J, Jobs M, Nerpin E, et al. Serum cathepsin s is associated with decreased insulin sensitivity and the development of type 2 diabetes in a community-based cohort of elderly men. *Diabetes Care* (2013) 36 (1):163–5. doi: 10.2337/dc12-0494

11. Kumar Vr S, Darisipudi MN, Steiger S, Devarapu SK, Tato M, Kukarni OP, et al. Cathepsin s cleavage of protease-activated receptor-2 on endothelial cells promotes microvascular diabetes complications. *J Am Soc Nephrol* (2016) 27(6):1635–49. doi: 10.1681/ASN.2015020208

12. Thomas MC, Brownlee M, Susztak K, Sharma K, Jandeleit-Dahm KA, Zoungas S, et al. Diabetic kidney disease. *Nat Rev Dis Primers* (2015) 1:15018. doi: 10.1038/nrdp.2015.18

13. American Diabetes A. 2. classification and diagnosis of diabetes: standards of medical care in diabetes-2020. *Diabetes Care* (2020) 43(Suppl 1):S14–31. doi: 10.2337/dc20-S002

14. Expert group of Chinese Society of Nephrology. Chinese Guidelines for clinical diagnosis and treatment of diabetic kidney disease. *Chin J Nephrol* (2021) 37(3):255–304. doi: 10.3760/cma.j.cn441217-20201125-00041

15. Ma YC, Zuo L, Chen JH, Luo Q, Yu XQ, Li Y, et al. Modified glomerular filtration rate estimating equation for Chinese patients with chronic kidney disease. *J Am Soc Nephrol JASN* (2006) 17(10):2937–44. doi: 10.1681/ASN.2006040368

16. Papadopoulou-Marketou N, Kanaka-Gantenbein C, Marketos N, P. Chrousos G, Papassotiriou. I. Biomarkers of diabetic nephropathy: a 2017 update. *Crit Rev Clin Lab Sci* (2017) 54(5):326–42. doi: 10.1080/10408363.2017.1377682

17. Gembillo G, Ingrasciotta Y, Crisafulli S, Luxi N, Siligato R, Santoro D, et al. Kidney disease in diabetic patients: from pathophysiology to pharmacological aspects with a focus on therapeutic inertia. *Int J Mol Sci* (2021) 22(9):4824. doi: 10.3390/ijms22094824

18. Zhang X, Zhou Y, Yu X, Huang Q, Fang W, Li J, et al. Differential roles of cysteinyl cathepsins in TGF-beta signaling and tissue fibrosis. *iScience* (2019) 19:607–22. doi: 10.1016/j.isci.2019.08.014

19. Repnik U, Starr AE, Overall CM, Turk B. Cysteine cathepsins activate ELR chemokines and inactivate non-ELR chemokines. *J Biol Chem* (2015) 290(22):13800–11. doi: 10.1074/jbc.M115.638395

20. Klinngam W, Fu R, Janga SR, Edman MC, Hamm-Alvarez SF. Cathepsin s alters the expression of pro-inflammatory cytokines and MMP-9, partially through protease -activated receptor-2, in human corneal epithelial cells. *Int J Mol Sci* (2018) 19 (11):3530. doi: 10.3390/ijms19113530

21. Liu J, Ma L, Yang J, Ren A, Sun Z, Yan G, et al. Increased serum cathepsin s in patients with atherosclerosis and diabetes. *Atherosclerosis* (2006) 186(2):411–9. doi: 10.1016/j.atherosclerosis.2005.08.001

22. Steubl D, Kumar SV, Tato M, Mulay SR, Larsson A, Lind L, et al. Circulating cathepsin-s levels correlate with GFR decline and sTNFR1 and sTNFR2 levels in mice and humans. *Sci Rep* (2017) 7:43538. doi: 10.1038/srep43538

23. Siddiqi Z, Karoli R, Kaul A, Fatima J, Varshney S, Beg MS. Evaluation of neutrophil gelatinase-associated lipocalin and cystatin c as early markers of diabetic nephropathy. *Ann Afr Med* (2017) 16(3):101–6. doi: 10.4103/aam.aam_12_17

24. Stankute I, Radzeviciene L, Monstaviciene A, Dobrovolskiene R, Danyte E, Verkauskiene R. Serum cystatin c as a biomarker for early diabetic kidney disease and dyslipidemia in young type 1 diabetes patients. *Medicina (Kaunas)* (2022) 58(2):218. doi: 10.3390/medicina58020218

25. Agrawal L, Azad N, Bahn GD, Ge L, Reaven PD, Hayward RA, et al. Long-term follow-up of intensive glycaemic control on renal outcomes in the veterans affairs diabetes trial (VADT). *Diabetologia* (2018) 61(2):295–9. doi: 10.1007/s00125-017-4473-2

26. Patney V, Whaley-Connell A, Bakris G. Hypertension management in diabetic kidney disease. *Diabetes Spectr* (2015) 28(3):175–80. doi: 10.2337/diaspect.28.3.175

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 18 July 2023
DOI 10.3389/fendo.2023.1215521



OPEN ACCESS

EDITED BY
Youxin Wang,
Capital Medical University, China

REVIEWED BY
Aleksandra Klisic,
Primary Health Care Center
Podgorica, Montenegro
Giovanni Tarantino,
University of Naples Federico II, Italy

*CORRESPONDENCE
Qian Zhao
zhaoqian6628@163.com

RECEIVED 02 May 2023
ACCEPTED 03 July 2023
PUBLISHED 18 July 2023

CITATION
Tan M-Y, Mo C-Y, Li F and Zhao Q (2023) The association between serum uric acid and hypertriglyceridemia: evidence from the national health and nutrition examination survey (2007–2018). *Front. Endocrinol.* 14:1215521. doi: 10.3389/fendo.2023.1215521



# The association between serum uric acid and hypertriglyceridemia: evidence from the national health and nutrition examination survey (2007–2018)

Mo-Yao Tan[1], Chao-Yue Mo[2], Fang Li[3] and Qian Zhao[1]*

[1]Chengdu Integrated TCM and Western Medicine Hospital, Chengdu, China, [2]College of Life and Science, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [3]Clinical Medical School, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

**Background:** Accumulating evidence suggests that elevated serum uric acid (SUA) may be a risk factor for hypertriglyceridemia (HTG). However, the epidemiological evidence for the association between SUA and HTG is limited. This article aimed to use the data from National Health and Nutrition Examination Survey (NHANES) (2007–2018) database to bridge the research gap.

**Methods:** This cross-sectional study used data from 10027 adults involved in NHANES from 2007-2018. We designed the exposure variable as SUA and the outcome variable as HTG. The covariates included demographics, questionnaires, laboratory, and examination information. Weighted logistic regression and subgroup analysis were used to explore the independent association between SUA and HTG. Furthermore, interaction tests were also carried out to evaluate the strata differences. Generalized additive models (GAM), smooth curve fittings, and threshold effect analysis were applied to examine the non-linear relationship.

**Results:** A total of 10027 participants were included, of which 3864 were HTG participants and 6163 were non-HTG participants. After fully adjusting for confounders, weighted multiple logistic regression models revealed a 77% increase in the risk of HTG when each unit of log2-SUA increased. There was also a positive association between elevated log2-SUA and developed risk of HTG in the quartile (Q) groups (Q1 OR: 1.00; Q2 OR: 1.17 [95%CI: 0.95,1.45]; Q3 OR: 1.43 [95%CI: 1.16,1.78]; Q4 OR: 1.68 [95%CI: 1.36,2.08]. The subgroup analysis results remained consistent across strata, with a strong positive correlation between SUA and HTG. Interaction tests showed no dependence on physical activity (PA), gender, BMI, smoking status, alcohol intake, hypertension, and diabetes for this positive association between log2-SUA and HTG (all $p$ for interaction >0.05). The participants' age may impact the strength of the association between SUA and HTG ($p$ for interaction <0.05).

**Conclusion:** There is a positive association between SUA and HTG in US adults. Considering that SUA may be a risk factor for HTG, individuals diagnosed with HTG should prioritize the daily management of SUA as part of their comprehensive care.

KEYWORDS

serum uric acid, hypertriglyceridemia, NHANES, weighted logistic regression analysis, subgroup analysis

# Introduction

Hypertriglyceridemia (HTG), characterized by fasting serum triglyceride (TG) levels exceeding 1.7 mmol/L (1), represents a prevalent lipid metabolism disorder (2). The development of HTG involves a complex interplay of genetic and non-genetic factors (3), leading to its classification into primary HTG and secondary HTG based on the underlying etiology (4). According to NHANES data, it is estimated that there are approximately 10.84 million individuals diagnosed with HTG in the United States (5). Moreover, severe HTG has been linked to a substantial rise in healthcare costs, ranging from 33% to 38% annually (6), thereby significantly amplifying the societal healthcare burden. Studies have shown that HTG poses not only a risk for pancreatitis and cardiovascular disease but also exhibits a robust association with obesity, diabetes, and nonalcoholic fatty liver disease (NAFLD) (7–9). Notable is the potential link between NAFLD and bladder cancer development (10). In a comprehensive cohort study, researchers observed that elevated serum triglyceride (TG) levels were independently associated with a more severe progression of pancreatitis and a higher likelihood of complications (11). Furthermore, a previous study revealed that serum TG levels reaching 10.20 mmol/L or higher were linked to a 10% to 20% increased risk of pancreatitis (12). The implementation of lifestyle interventions, such as dietary modifications, weight management, and increased physical activity, is considered the primary and most valuable approach for treating HTG. By significantly lowering TG levels through these interventions, it is possible to not only prevent pancreatitis but also reduce the risk of cardiovascular disease (2, 13, 14).

SUA is the final oxidation product of exogenous and endogenous purine metabolism and is produced in the intestine, liver, and muscle (15–17). Endogenous purines are the main component of purines in the body, accounting for 80% of total purines, which are mainly derived from the oxidative breakdown of the body's nucleic acids (16). Exogenous purines are derived from dietary intake, including seafood, fatty and organ meats (e.g., liver and kidney), fructose, and alcohol (18). In recent years, there has been an observed increase in people's SUA levels, leading to the emergence of hyperuricemia as a significant public health concern. Previous data have indicated that the prevalence of hyperuricemia in the United States is as high as approximately 20% (19). Notably, research has demonstrated an association between hyperuricemia and urological cancer (20). This may be related to dietary habits, lifestyle changes, and medication use (21, 22). Over the years, numerous studies have consistently demonstrated a strong association between high SUA levels and the development of cardiovascular disease. Furthermore, elevated SUA has been linked to increased risks of all-cause mortality and cardiovascular mortality (23, 24). High SUA is also associated with the development of many other diseases, including diabetes, hypertension, NAFLD and kidney disease (25–27). Above all, elevated SUA levels can alter the body's physiopathology and heighten the susceptibility to diseases.

One hypothesis is that SUA levels may be related to HTG. It is thought-provoking that both HTG and high levels of SUA can lead to the development of NAFLD (9, 27). Gout patients have been shown to have an increased risk of urologic cancers (20), and HTG-induced NAFLD has also been shown to indirectly raise this risk (10). Research investigating the relationship between SUA and metabolic syndrome (MetS) found that hyperuricemia showed the strongest association with high TG (PR = 2.32, 95% CI: 1.84-2.91) (28). A 5-year cohort study conducted in Japan revealed an increased risk of low-density lipoprotein (LDL) and HTG with elevated SUA levels (29). Additionally, previous evidence indicates that elevated SUA levels can induce mitochondrial abnormalities, contributing to the progression of HTG (8). Notably, an animal experiment demonstrated lower lipase activity in the high SUA group compared to the low SUA group (30). Decreased lipase activity is associated with reduced TG catabolism (31). Moreover, apolipoprotein E (ApoE) has been implicated in SUA-induced HTG (32).

Drawing from the existing evidence, it is suggested that SUA could serve as a risk factor for HTG. Accordingly, we collected NHANES data (2007–2018) to provide epidemiological evidence, and conducted an investigation into the relationship between SUA and HTG by employing weighted multivariate logistic regression and performing subgroup analysis.

# Materials and methods

## Data source

The NHANES database is a nationally conducted cross-sectional study aiming to evaluate the health and nutritional status of non-institutionalized residents in the United States.

Administered by the National Center for Health Statistics (NCHS), NHANES collects data through interviews and examinations. The study design employs a stratified multistage probability sampling method, ensuring a highly representative sample. The NHANES protocol has been reviewed and approved by the Research Ethics Review Committee of the National Center for Health Statistics, and all participants have provided written informed consent. The publicly available data used in our analysis can be accessed at https://www.cdc.gov/nchs/nhanes/. This study adhered strictly to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) principle for cross-sectional studies (Supplementary material 1) (33).

## Study population

The study incorporated data from six cross-sectional cycles (2007–2018) of NHANES. The initial inclusion of 59842 participants in the study was based on the inclusion criteria that participants were at least 20 years old and that SUA and HTG data was available. To ensure data integrity, participants with missing data on prescription for cholesterol (PFC), alcohol intake, the ratio of family income to poverty (PIR), education level, body mass index (BMI), and sedentary time were excluded. Ultimately, a total of 10,027 participants who satisfied the aforementioned inclusion and exclusion criteria were included in the data analysis (Figure 1).

## Measurements and definition of variables

### Exposure variable and outcome variable

SUA was used as the exposure variable in this study. To account for its right-skewed distribution, log2 transformation was applied to SUA during subgroup and regression analyses. HTG was defined as an outcome variable, with HTG being classified as serum TG content greater than or equal to 1.7 mmol/L, according to endocrine clinical guidelines (4).

### Covariates

Based on prior research and clinical experience, we have incorporated the following summary of covariates that might influence the relationship between SUA and HTG (34, 35). The

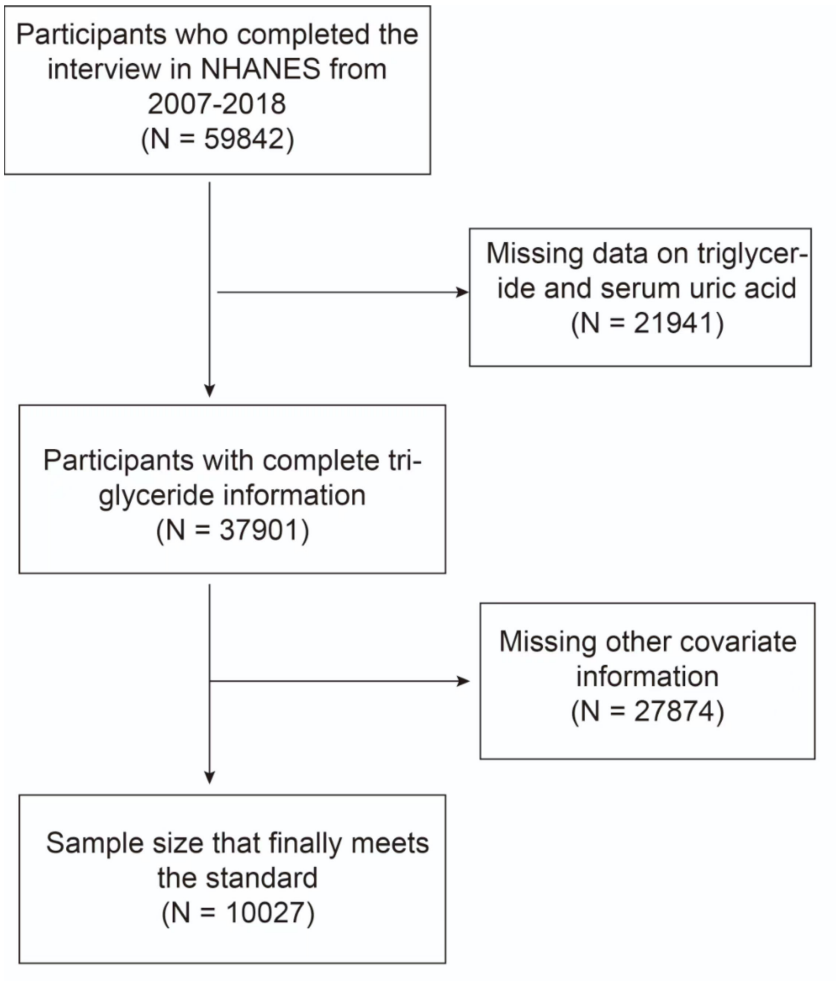



FIGURE 1
Flowchart of the sample selection from the 2007–2018 National Health and Nutrition Examination Survey (NHANES).

study considered the following continuous covariates: age (years), sedentary time (minutes), Alanine aminotransferase (ALT, U/L), Aspartate aminotransferase (AST, U/L), creatinine (μmol/L), blood urea nitrogen (mmol/L), and total cholesterol (mmol/L). Categorical variables included: gender (Male/Female), race (Mexican American/other Hispanic/Non-Hispanic White/Non-Hispanic Black/Other Race), educational level (High school or above high school/less than High school), smoking status (Yes: smoking at least 100 cigarettes; No: smoking less than 100 cigarettes), PA (yes was defined as engaging in any moderate recreational activities for at least 10 continuous minutes), hypertension (yes/no), diabetes (yes/no), prescription for cholesterol (PFC) (yes/no), and ratio of family income to poverty (PIR). Hypertension was defined as an average systolic blood pressure ≥130mmHg or diastolic blood pressure ≤ 80mmHg, or taking medication for hypertension (36). Diabetes was diagnosed based on three criteria (1): self-reported diagnosis by a physician or healthcare professional (2), HbA1c (glycated hemoglobin) level over 6.5%, and (3) fasting blood glucose (FPG) level over 126 mg/dL (37). PIR was classified as low-income (PIR ≤ 1.3), middle-income (PIR > 1.3–3.5), and high-income (PIR> 3.5) (38). Marital status was divided into living alone, married, or living with a partner (34). Alcohol intake was classified as mild, moderate, and heavy. Heavy alcohol use was defined as ≥3 drinks per day for females or ≥ four drinks per day for males (39). Moderate alcohol use was defined as 2-3 drinks per day for females and 3-4 drinks per day for males. Mild alcohol use was regarded as others (40). The BMI was categorized as underweight or normal (<25 kg/m$^2$), overweight (≥25 to <30 kg/m$^2$), and obese (≥30 kg/m$^2$) (38).

## Statistical analysis

All statistical analyses were conducted according to CDC guidelines. Considering the complicated multistage cluster survey design, NHANES-generated sampling statistics strata, clusters, and weights were used to ensure the results were generalizable to the U.S. population (41).

Continuous variables were presented as mean with standard deviation (SD), while categorical variables were expressed as percentages. To address the right-skewed distribution of SUA data, log2 transformation was applied for regression and subgroup analysis. The statistical analysis comprised four main steps, aiming to examine the association between SUA levels and HTG among the selected participants. Firstly, participants' TG levels were categorized into HTG and non-HTG groups based on clinical guidelines. Differences between these groups were assessed using the chi-square test for categorical variables and the weighted Student's t-test for continuous variables. In the second step, weighted multivariate logistic regression models were employed to examine the independent association between SUA and HTG in three models. Model 1 did not include any covariate adjustments. Model 2 was adjusted for gender, age, and race. Model 3 included adjustments for all covariates, including age, gender, race, education level, sedentary time, AST, ALT, creatinine, blood urea nitrogen, total cholesterol, PIR, body mass index (BMI), smoking status, alcohol intake, physical activity (PA), hypertension, diabetes, and PFC. Furthermore, SUA was transformed from a continuous variable to a categorical variable (Q) for further analysis. In the third step, a subgroup analysis was conducted to examine the impact of different subgroups on the results. Interaction tests were employed to explore potential heterogeneity between these subgroups. Additionally, GAM, smooth curve fittings, and threshold effect analysis were utilized to investigate the non-linear relationship between SUA and HTG in greater detail.

If the two-sided value $P < 0.05$, the null hypothesis was rejected. All analysis was performed using Empower software (www.empowerstats.com; X&Y solutions, Inc., Boston MA) and R software (version 4.1.2; http://www.R-project.org, R Foundation for Statistical Computing, Vienna, Austria).

# Results

## Basic characteristics of the included participants

Table 1 presents the weighted baseline characteristics of participants selected from NHANES 2007 to 2018, stratified by the presence of HTG. The analysis included 3,864 participants with HTG. The average age of the HTG group was 50.95 ± 22.04 years, with 58.47% being male and 41.53% female. In comparison, the non-HTG group consisted of 6,163 participants with a mean age of 49.15 ± 30.04 years, with 44.56% being male and 55.44% female. Significant differences between the HTG and non-HTG groups were observed in terms of age, gender, race, education level, PIR, marital status, BMI, drinking, smoking status, PA, hypertension, cholesterol prescription, diabetes, PFC, ALT, AST, SUA, creatinine, blood urea nitrogen, and total cholesterol (all $p < 0.05$).

## Association between SUA and the HTG

Table 2 presents the association between SUA and the risk of HTG. A significant positive correlation was observed between SUA and HTG. In Model 1, the odds ratio (OR) was 3.40 (95% CI: 2.84-4.06), indicating a significant association. Model 2, which adjusted for age, gender, race, BMI, education level, PIR, PA, sedentary time, hypertension, diabetes, creatinine, blood urea nitrogen, total cholesterol, ALT, AST, smoking status, drinking, and PFC, also showed a positive association (OR = 3.17, 95% CI: 2.62-3.84). Furthermore, in Model 3, after full adjustment, a positive association between SUA and HTG was still observed (OR = 1.77, 95% CI: 1.44-2.18).

To gain further insights into the relationship between SUA and HTG, log-transformed SUA was categorized into quartiles. In the fully adjusted Model 3, when comparing the highest quartile (Q4) with the lowest quartile (Q1), the OR was 1.68 (95% CI: 1.36-2.08), indicating a stable positive association between higher SUA levels and HTG.

## Subgroup analysis

While informative, subgroup analysis was conducted to further assess the robustness of the association between SUA

10.3389/fendo.2023.1215521

TABLE 1 Weighted baseline characteristics of participants.

| | Non-Hypertriglyceridemia (n = 6163) | Hypertriglyceridemia (n = 3864) | P-value |
|---|---|---|---|
| **Age (year)** | 49.15 ± 30.04 | 50.95 ± 22.04 | <0.001 |
| **Sedentary time (minute)** | 394.38 ± 348.66 | 405.25 ± 297.01 | 0.0831 |
| **AST (U/L)** | 24.39 ± 14.01 | 26.74 ± 23.15 | <0.001 |
| **ALT (U/L)** | 23.09 ± 17.62 | 29.20 ± 27.27 | <0.001 |
| **Creatinine (μmol/L)** | 77.27 ± 30.84 | 80.28 ± 28.22 | <0.001 |
| **Blood Urea Nitrogen (mmol/L)** | 4.98 ± 3.20 | 5.20 ± 3.17 | 0.001 |
| **Serum uric acid (μmol/L)** | 310.67 ± 112.35 | 348.06 ± 111.47 | <0.001 |
| **Total Cholesterol (mmol/L)** | 4.89 ± 1.80 | 5.45 ± 2.06 | <0.001 |
| **Gender (%)** | | | <0.001 |
| Male | 44.56 | 58.47 | |
| Female | 55.44 | 41.53 | |
| **Race (%)** | | | <0.001 |
| Mexican American | 4.83 | 7.52 | |
| other Hispanic | 4.36 | 5.17 | |
| Non-Hispanic White | 73.27 | 75.12 | |
| Non-Hispanic Black | 11.37 | 4.93 | |
| Other Race | 6.16 | 7.26 | |
| **Education Level (%)** | | | <0.001 |
| High school or above high school | 92.20 | 90.44 | |
| Less than high school | 7.80 | 9.56 | |
| **PIR (%)** | | | 0.0099 |
| low-income | 12.95 | 15.15 | |
| middle-income | 31.83 | 32.97 | |
| high-income | 55.22 | 51.88 | |
| **Marital status (%)** | | | <0.001 |
| **Married** | 59.51 | 63.51 | |
| **living alone** | 33.84 | 5.01 | |
| **living with a partner** | 6.65 | 6.63 | |
| **BMI ($kg/m^2$) (%)** | | | <0.001 |
| underweight or normal | 33.89 | 14.09 | |
| overweight | 34.04 | 33.58 | |
| obese | 32.07 | 52.34 | |
| **Smoking status (%)** | | | <0.001 |
| Yes | 43.17 | 51.89 | |
| No | 56.83 | 48.11 | |
| **Alcohol intake (%)** | | | <0.001 |
| Mild | 53.94 | 54.84 | |
| Moderate | 35.17 | 30.57 | |
| Heavy | 10.89 | 14.59 | |

*(Continued)*

TABLE 1 Continued

| | Non-Hypertriglyceridemia (n = 6163) | Hypertriglyceridemia (n = 3864) | P-value |
|---|---|---|---|
| **Physical activity (%)** | | | <0.001 |
| Yes | 54.42 | 48.80 | |
| No | 45.58 | 51.20 | |
| **Hypertension (%)** | | | <0.001 |
| Yes | 32.24 | 44.45 | |
| No | 67.76 | 55.55 | |
| **Cholesterol prescription (%)** | | | <0.001 |
| Yes | 27.05 | 40.88 | |
| No | 72.95 | 59.12 | |
| **Diabetes (%)** | | | <0.001 |
| Yes | 7.41 | 15.73 | |
| No | 92.59 | 84.27 | |

ALT, alanine transaminase; AST, aspartate transaminase; PIR, Ratio of family income to poverty; BMI, body mass index.
All values are presented as proportion (%), or mean ± standard deviation.

and HTG. Interaction tests were also performed to assess the influence of different variables (Supplementary material 2). The results of the subgroup analysis revealed a consistent positive correlation between SUA and HTG across different subgroups, indicating the robustness of the association. Notably, no significant interactions were observed for gender, BMI, smoking status, drinking, hypertension, diabetes, and PA, suggesting that the association was not dependent on these variables (all $p$ for interaction > 0.05).

However, age was found to significantly impact the strength of the SUA-HTG association (all $p$ for interaction < 0.05). The results indicated that participants under the age of 60 were at a higher risk compared to those aged 60 and older, with an odds ratio of 2.47 (95% CI: 1.96-3.12). This suggests that age plays a role in modifying the association between SUA and HTG, with younger individuals exhibiting a stronger association.

## Identification of non-linear relationship

This study revealed a non-linear relationship between SUA and HTG, as demonstrated by the results of GAM and smoothed curve fitting presented in Figure 2. The log-likelihood ratio test showed a $p$-value of less than 0.001 when comparing the linear regression model to a two-piecewise linear regression model, indicating that the two-piecewise linear regression model provided a better fit for the data. Using the two-piecewise linear regression model and recursive algorithm, Table 3 presents the findings. The point of inflection in the U-shaped association between SUA and HTG was identified as 7.86 umol/L for log2-SUA. To the left of the inflection point, the effect size (log2 transformed) was 0.76 (95% CI: 0.45, 1.28) with a $p$-value of 0.30, suggesting a non-significant association. However, to the right of the inflection point, SUA showed a significant positive correlation with HTG. The effect size

TABLE 2 Weighted Multivariate logistic regression models of SUA with hypertriglyceridemia.

| log-SUA (umol/L) | OR[a](95% CI), P-value | | |
|---|---|---|---|
| | Model 1[b] | Model 2[c] | Model 3[d] |
| **Continuous** | 3.40 (2.84, 4.06) <0.001 | 3.17 (2.62, 3.84) <0.001 | 1.77 (1.44, 2.18) <0.001 |
| **Categories** | | | |
| **Quartile 1 (≤8.06)** | **Reference** | **Reference** | **Reference** |
| **Quartile 2 (8.06-8.33)** | 1.48 (1.22,1.80) <0.001 | 1.41 (1.15,1.73) <0.001 | 1.17 (0.95,1.45) 0.14 |
| **Quartile 3 (8.34-8.57)** | 2.20 (1.81,2.70) <0.001 | 2.01 (1.63,2.48) <0.001 | 1.43 (1.16,1.78) 0.0018 |
| **Quartile 4 (>8.57)** | 3.29 (2.73,3.97) <0.001 | 3.02 (2.46,3.69) <0.001 | 1.68 (1.36,2.08) <0.001 |

SUA, serum uric acid; 95% CI, 95% confidence interval; OR, odds ratio.
In sensitivity analysis, SUA is converted from a continuous variable to a categorical variable (quartile); OR[a], effect size; Model 1[b], no covariates were adjusted; Model 2[c], adjusted for gender, age, and race; Model 3[d], adjusted for gender, age, race, PA, sedentary time, ALT, AST, creatinine, blood urea nitrogen, education level, the ratio of family income to poverty, marital status, body mass index, alcohol intake, smoking status, hypertension, cholesterol prescription, total cholesterol, and diabetes.

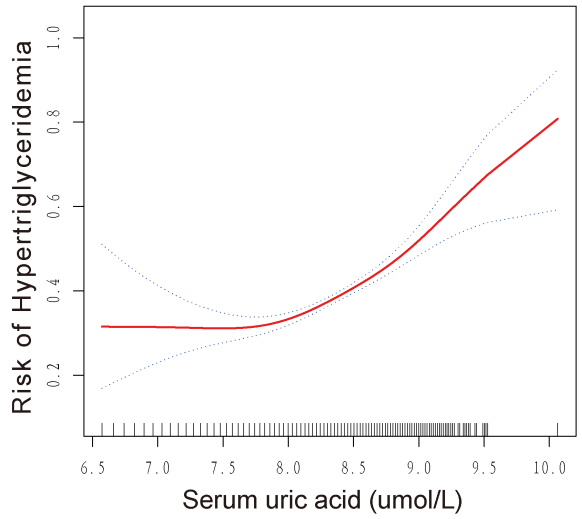



FIGURE 2
smoothed curve fitting: Dose-response relationship between SUA and hypertriglyceridemia.

(log2 transformed) was 2.16 (95% CI: 1.82, 2.56) with a *p*-value of <0.001, indicating a strong and significant association between higher SUA levels and HTG.

## Discussion

Our study revealed a significant positive association between SUA and HTG. Subgroup analysis indicated that this association was consistent across different subgroups. Interaction tests demonstrated that the association was independent of gender, BMI, smoking status, alcohol intake, physical activity, hypertension, and diabetes. Interestingly, participants under the age of 60 had a higher risk of developing HTG compared to those aged 60 and older. Moreover, we observed a U-shaped association between SUA and HTG, with an inflection point identified at 7.86 umol/L.

TABLE 3 Threshold effect analysis of SUA on HTG using two-piecewise linear regression model.

| log2-SUA (umol/L) | Adjust OR (95% CI) P value |
|---|---|
| Fitting by linear regression model | 1.84 (1.60, 2.13) <0.0001 |
| Fitting by two-piecewise linear regression model | |
| Inflection point | 7.86 |
| < 7.86 | 0.76 (0.45, 1.28) 0.30 |
| > 7.86 | 2.16 (1.82, 2.56) <0.0001 |
| Log likelihood ratio test | <0.001 |

Adjusted for gender, age, race, PA, sedentary time, ALT, AST, creatinine, blood urea nitrogen, education level, ratio of family income to poverty, marital status, body mass index, alcohol intake, smoking status, hypertension, cholesterol prescription, total cholesterol,and diabetes.

The association between SUA and Metabolic Syndrome (MetS) has been extensively studied in previous research. MetS is characterized by the abnormal accumulation of multiple metabolic components, such as obesity, HTG, low high-density lipoprotein cholesterol (HDL-C), hypertension, and insulin resistance (IR). Several studies have indicated that SUA might serve as an independent risk factor for MetS (42, 43). These studies have provided evidence supporting the potential role of SUA in the development of MetS (44–46). The result of a cohort study revealed that high SUA concentrations may increase the risk of MetS among Chinese adults (47). There is evidence that SUA may be associated with IR, which is one of the diagnostic indicators of MetS (48). Furthermore, Xanthine oxidoreductase (XO), an important enzyme involved in the production of SUA, has been suggested to play a crucial role in the development of MetS (49). Animal experiments have shown that lowering SUA levels can prevent and reverse MetS features in fructose-fed rats, including lower blood pressure, reduced serum triglycerides, decreased hyperinsulinemia, and weight gain (50). These findings, along with our study results, support the connection between SUA and the development of MetS.

The underlying mechanism linking SUA and HTG remains unclear. However, studies have suggested that high intracellular SUA levels can lead to increased oxidative stress in mitochondria. In an *in vitro* study by Yang et al. (51), hepatocytes treated with different concentrations of SUA exhibited increased apoptotic activity, accumulation of Reactive Oxygen Species (ROS), and elevated 8-hydroxydeoxyguanosine levels compared to control cells, indicating mitochondrial DNA damage. This mitochondrial dysfunction could contribute to the release of citrate into the cytosol, initiating lipogenesis and triglyceride synthesis (52). In addition, Wang et al. (53) found that the prevalence of the E2 allele of ApoE was correlated with increased SUA level. ApoE is known to play a role in regulating lipoprotein metabolism. Studies have suggested that ApoE may contribute to decreased clearance of very low-density lipoprotein (VLDL) through the interaction with the hepatic remnant receptor, leading to elevated levels of VLDL cholesterol and VLDL triglycerides (54). In addition, it has been suggested that elevated SUA may be associated with reduced lipase activity (30). In the study by Zheng et al. (31), it was found that elevated SUA levels may hinder the breakdown of triglycerides (TG) by reducing lipase activity. This inhibition of TG catabolism could contribute to a higher prevalence of HTG in individuals with high SUA levels.

In our subgroup analysis, we found that the association between SUA and HTG was more pronounced in participants under 60 years old compared to those over 60 years old. This observation is consistent with previous evidence suggesting that the health effects of SUA are stronger in younger individuals (55). For example, Kawamoto et al. (56) reported higher risk factor values for cardiovascular disease induced by SUA in women aged <55 compared to older participants. However, the underlying mechanism for this age-related effect of SUA on HTG is not yet well understood and requires further exploration in future scientific research. Moreover, our threshold effects analysis revealed that the inflection point for the association between log2-SUA and HTG was 7.86 umol/L. Beyond this threshold, there was a significantly higher risk of HTG. However, more experimental studies are needed to

determine more precise threshold effects and to elucidate the mechanisms involved.

This study has several strengths that contribute to its scientific value. Firstly, it utilized a large sample size of 10,027 participants, providing robust statistical power for accurately assessing the association between SUA and HTG. Secondly, the study utilized data from the NHANES database, which is a nationally representative population-based sample. By incorporating appropriate sampling weights, the study results can be generalized to the entire US population. Lastly, the study accounted for various confounding covariates based on previous research and clinical experience, minimizing the potential bias caused by these factors.

However, there are certain limitations that should be acknowledged. Firstly, being a cross-sectional study, it is unable to establish a causal relationship between SUA and HTG. Further longitudinal studies are needed to explore the temporal nature of this association. Secondly, despite adjusting for several potential confounding factors, it is possible that other unmeasured variables might have influenced the results. Additionally, caution should be exercised when extrapolating the findings to populations outside of the United States, as the study was restricted to American participants. Finally, due to the limited amount of relevant research published in the last five years, not all of our primary references are recent works, which may have an impact on the timeliness of this study.

## Conclusions

In conclusion, our study demonstrates a strong positive association between SUA levels and HTG in the US adult population, indicating that elevated SUA may contribute to an increased risk of HTG. These findings highlight the importance of managing and controlling SUA levels in HTG patients to prevent disease progression. Crucial for future research is to conduct large-scale and high-quality prospective studies to validate our conclusions and further explore the underlying mechanisms of this association.

## 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: https://www.cdc.gov/nchs/nhanes/.

## Ethics statement

The studies involving human participants were reviewed and approved by National Center for Health Statistics (NCHS). The patients/participants provided their written informed consent to participate in this study.

## Author contributions

Conception and design of the study: M-YT, QZ. Data collection and analysis: M-YT, C-YM, FL, QZ. Draft paper: M-YT, C-YM, QZ. Approve the final paper to be published: All authors.

## Acknowledgments

The author would like to acknowledge Qian Zhao from Chengdu Integrated TCM and Western Medicine Hospital.

## 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.2023.1215521/full#supplementary-material

## References

1. Yang AL, McNabb-Baltar J. Hypertriglyceridemia and acute pancreatitis. *Pancreatology* (2020) 20(5):795–800. doi: 10.1016/j.pan.2020.06.005

2. Simha V. Management of hypertriglyceridemia. *Bmj* (2020) 371:m3109. doi: 10.1136/bmj.m3109

3. Watts GF, Ooi EMM, Chan DC. Demystifying the management of hypertriglyceridaemia. *Nat Rev Cardiol* (2013) 10(11):648–61. doi: 10.1038/nrcardio.2013.140

4. Berglund L, Brunzell JD, Goldberg AC, Goldberg IJ, Sacks F, Murad MH, et al. Evaluation and treatment of hypertriglyceridemia: an endocrine society clinical practice guideline. *Endocrine Soc* (2012) 97(9):2969–89. doi: 10.1210/jc.2011-3213

5. Li Z, Zhu G, Chen G, Luo M, Liu X, Chen Z, et al. Distribution of lipid levels and prevalence of hyperlipidemia: data from the nhanes 2007-2018. *Lipids Health Dis* (2022) 21(1):111. doi: 10.1186/s12944-022-01721-y

6. Nichols GA, Arondekar B, Garrison LPJr. Patient characteristics and medical care costs associated with hypertriglyceridemia. *Am J Cardiol* (2011) 107(2):225–9. doi: 10.1016/j.amjcard.2010.09.010

7. Ndrepepa G. Uric acid and cardiovascular disease. *Clin Chim Acta* (2018) 484:150–63. doi: 10.1016/j.cca.2018.05.046

8. Parhofer KG, Laufs U. The diagnosis and treatment of hypertriglyceridemia. *Dtsch Arztebl Int* (2019) 116(49):825–32. doi: 10.3238/arztebl.2019.0825

9. Hou X, Guan Y, Tang Y, Song A, Zhao J, Ren L, et al. A correlation study of the relationships between nonalcoholic fatty liver disease and serum triglyceride concentration after an oral fat tolerance test. *Lipids Health Dis* (2021) 20(1):54. doi: 10.1186/s12944-021-01483-z

10. Tarantino G, Crocetto F, Di Vito C, Creta M, Martino R, Pandolfo SD, et al. Association of nafld and insulin resistance with non metastatic bladder cancer patients: a cross-sectional retrospective study. *J Clin Med* (2021) 10(2):346. doi: 10.3390/jcm10020346

11. Pascual I, Sanahuja A, García N, Vázquez P, Moreno O, Tosca J, et al. Association of elevated serum triglyceride levels with a more severe course of acute pancreatitis: cohort analysis of 1457 patients. *Pancreatology* (2019) 19(5):623–9. doi: 10.1016/j.pan.2019.06.006

12. Scherer J, Singh VP, Pitchumoni CS, Yadav D. Issues in hypertriglyceridemic pancreatitis: an update. *J Clin Gastroenterol* (2014) 48(3):195–203. doi: 10.1097/01.mcg.0000436438.60145.5a

13. Santos-Baez LS, Ginsberg HN. Hypertriglyceridemia-causes, significance, and approaches to therapy. *Front Endocrinol (Lausanne)* (2020) 11:616. doi: 10.3389/fendo.2020.00616

14. Chait A, Feingold KR. Approach to patients with hypertriglyceridemia. *Best Pract Res Clin Endocrinol Metab* (2023) 37(3):101659. doi: 10.1016/j.beem.2022.101659

15. Johnson RJ, Lanaspa MA, Gaucher EA. Uric acid: a danger signal from the rna world that may have a role in the epidemic of obesity, metabolic syndrome, and cardiorenal disease: evolutionary considerations. *Semin Nephrol* (2011) 31(5):394–9. doi: 10.1016/j.semnephrol.2011.08.002

16. Dawson J, Quinn T, Walters M. Uric acid reduction: a new paradigm in the management of cardiovascular risk? *Curr Med Chem* (2007) 14(17):1879–86. doi: 10.2174/092986707781058797

17. Hediger MA, Johnson RJ, Miyazaki H, Endou H. Molecular physiology of urate transport. *Physiology* (2005) 20(2):125–33. doi: 10.1152/physiol.00039.2004

18. Kang D-H, Chen W. Uric acid and chronic kidney disease: new understanding of an old problem. *Semin Nephrol* (2011) 31(5):447–52. doi: 10.1016/j.semnephrol.2011.08.009

19. Chen-Xu M, Yokose C, Rai SK, Pillinger MH, Choi HK. Contemporary prevalence of gout and hyperuricemia in the united states and decadal trends: the national health and nutrition examination survey, 2007-2016. *Arthritis Rheumatol* (2019) 71(6):991–9. doi: 10.1002/art.40807

20. Chen CJ, Yen JH, Chang SJ. Gout patients have an increased risk of developing most cancers, especially urological cancers. *Scand J Rheumatol* (2014) 43(5):385–90. doi: 10.3109/03009742.2013.878387

21. Yokose C, McCormick N, Choi HK. Dietary and lifestyle-centered approach in gout care and prevention. *Curr Rheumatol Rep* (2021) 23(7):51. doi: 10.1007/s11926-021-01020-y

22. Ashiq K, Bajwa MA, Tanveer S, Qayyum M, Ashiq S, Khokhar R, et al. A comprehensive review on gout: the epidemiological trends, pathophysiology, clinical presentation, diagnosis and treatment. *J Pak Med Assoc* (2021) 71(4):1234–8. doi: 10.47391/jpma.313

23. Zhang S, Wang Y, Cheng J, Huangfu N, Zhao R, Xu Z, et al. Hyperuricemia and cardiovascular disease. *Curr Pharm Des* (2019) 25(6):700–9. doi: 10.2174/1381612825666190408122557

24. Cheong E, Ryu S, Lee JY, Lee SH, Sung JW, Cho DS, et al. Association between serum uric acid and cardiovascular mortality and all-cause mortality: a cohort study. *J Hypertens* (2017) 35 Suppl 1:S3–s9. doi: 10.1097/hjh.0000000000001330

25. Katsiki N, Dimitriadis GD, Mikhailidis DP. Serum uric acid and diabetes: from pathophysiology to cardiovascular disease. *Curr Pharm Des* (2021) 27(16):1941–51. doi: 10.2174/1381612827666210104124320

26. Ejaz AA, Johnson RJ, Shimada M, Mohandas R, Alquadan KF, Beaver TM, et al. The role of uric acid in acute kidney injury. *Nephron* (2019) 142(4):275–83. doi: 10.1159/000499939

27. Wei F, Li J, Chen C, Zhang K, Cao L, Wang X, et al. Higher serum uric acid level predicts non-alcoholic fatty liver disease: a 4-year prospective cohort study. *Front Endocrinol (Lausanne)* (2020) 11:179. doi: 10.3389/fendo.2020.00179

28. Gonçalves JP, Oliveira A, Severo M, Santos AC, Lopes C. Cross-sectional and longitudinal associations between serum uric acid and metabolic syndrome. *Endocrine* (2012) 41(3):450–7. doi: 10.1007/s12020-012-9629-8

29. Kuwabara M, Borghi C, Cicero AFG, Hisatome I, Niwa K, Ohno M, et al. Elevated serum uric acid increases risks for developing high ldl cholesterol and hypertriglyceridemia: a five-year cohort study in Japan. *Int J Cardiol* (2018) 261:183–8. doi: 10.1016/j.ijcard.2018.03.045

30. Gibson T, Kilbourn K, Horner I, Simmonds HA. Mechanism and treatment of hypertriglyceridaemia in gout. *Ann Rheum Dis* (1979) 38(1):31–5. doi: 10.1136/ard.38.1.31

31. Zheng R, Ren P, Chen Q, Yang T, Chen C, Mao Y. Serum uric acid levels and risk of incident hypertriglyceridemia: a longitudinal population-based epidemiological study. *Ann Clin Lab Sci* (2017) 47(5):586–91.

32. Straat ME, Martinez-Tellez B, Nahon KJ, Janssen LGM, Verhoeven A, van der Zee L, et al. Comprehensive (Apo)Lipoprotein profiling in patients with genetic hypertriglyceridemia using lc-Ms and nmr spectroscopy. *J Clin Lipidol* (2022) 16(4):472–82. doi: 10.1016/j.jacl.2022.04.004

33. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (Strobe) statement: guidelines for reporting observational studies. *Lancet* (2007) 370(9596):1453–7. doi: 10.1016/s0140-6736(07)61602-x

34. Zhang Y, Liu W, Zhang W, Cheng R, Tan A, Shen S, et al. Association between blood lead levels and hyperlipidemiais: results from the nhanes (1999-2018). *Front Public Health* (2022) 10:981749. doi: 10.3389/fpubh.2022.981749

35. Duan H, Zhang R, Chen X, Yu G, Song C, Jiang Y, et al. Associations of uric acid with liver steatosis and fibrosis applying vibration controlled transient elastography in the united states: a nationwide cross-section study. *Front Endocrinol (Lausanne)* (2022) 13:930224. doi: 10.3389/fendo.2022.930224

36. He P, Li H, Liu C, Liu M, Zhang Z, Zhang Y, et al. U-Shaped association between dietary copper intake and new-onset hypertension. *Clin Nutr* (2022) 41(2):536–42. doi: 10.1016/j.clnu.2021.12.037

37. Casagrande SS, Lee C, Stoeckel LE, Menke A, Cowie CC. Cognitive function among older adults with diabetes and prediabetes, nhanes 2011-2014. *Diabetes Res Clin Pract* (2021) 178:108939. doi: 10.1016/j.diabres.2021.108939

38. Seidell JC, Deurenberg P, Hautvast JG. Obesity and fat distribution in relation to health–current insights and recommendations. *World Rev Nutr Diet* (1987) 50:57–91. doi: 10.1159/000414170

39. CDC/National Center for Health Statistics. *National health and nutrition examination survey, nhanes (2019-2020)* (2021). Available at: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/manuals.aspx?BeginYear=2019 (Accessed May 21, 2022).

40. Phillips JA. Dietary guidelines for americans, 2020-2025. *Workplace Health Saf* (2021) 69(8):395. doi: 10.1177/21650799211026980

41. Johnson CL, Paulose-Ram R, Ogden CL, Carroll MD, Kruszon-Moran D, Dohrmann SM, et al. National health and nutrition examination survey: analytic guidelines, 1999-2010. *Vital Health Stat 2* (2013) 161):1–24.

42. Saklayen MG. The global epidemic of the metabolic syndrome. *Curr Hypertens Rep* (2018) 20(2):12. doi: 10.1007/s11906-018-0812-z

43. Lemieux I, Després JP. Metabolic syndrome: past, present and future. *Nutrients* (2020) 12(11):3051. doi: 10.3390/nu12113501

44. Cibičková Ľ, Langová K, Vaverková H, Kubíčková V, Karásek D. Correlation of uric acid levels and parameters of metabolic syndrome. *Physiol Res* (2017) 66(3):481–7. doi: 10.33549/physiolres.933410

45. Xu J, Liu C, Fu L, Li L, Wang T. The association of serum uric acid with metabolic syndrome and its components-from a single-clinical centre in China. *Int J Clin Pract* (2021) 75(4):e13845. doi: 10.1111/ijcp.13845

46. Rubio-Guerra AF, Morales-López H, Garro-Almendaro AK, Vargas-Ayala G, Durán-Salgado MB, Huerta-Ramírez S, et al. Circulating levels of uric acid and risk for metabolic syndrome. *Curr Diabetes Rev* (2017) 13(1):87–90. doi: 10.2174/1573399812666150930122507

47. Ren P, Gao M. Association between metabolic syndrome and the serum uric acid: a cohort study. *Clin Lab* (2018) 64(5):719–26. doi: 10.7754/Clin.Lab.2017.171037

48. Hu X, Rong S, Wang Q, Sun T, Bao W, Chen L, et al. Association between plasma uric acid and insulin resistance in type 2 diabetes: a mendelian randomization analysis. *Diabetes Res Clin Pract* (2021) 171:108542. doi: 10.1016/j.diabres.2020.108542

49. Battelli MG, Bortolotti M, Polito L, Bolognesi A. The role of xanthine oxidoreductase and uric acid in metabolic syndrome. *Biochim Biophys Acta* (2018) 1864(8):2557–65. doi: 10.1016/j.bbadis.2018.05.003

50. Nakagawa T, Hu H, Zharikov S, Tuttle KR, Short RA, Glushakova O, et al. A causal role for uric acid in fructose-induced metabolic syndrome. *Am Physiol Soc* (2006) 290(3):F625–F31. doi: 10.1152/ajprenal.00140.2005

51. Yang Y, Zhou Y, Cheng S, Sun JL, Yao H, Ma L. Effect of uric acid on mitochondrial function and oxidative stress in hepatocytes. *Genet Mol Res* (2016) 15(2):1–11. doi: 10.4238/gmr.15028644

52. Ipsen DH, Lykkesfeldt J, Tveden-Nyborg P. Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease. *Cell Mol Life Sci* (2018) 75(18):3313–27. doi: 10.1007/s00018-018-2860-6

53. Wang C, Yan W, Wang H, Zhu J, Chen H. Apoe polymorphism is associated with blood lipid and serum uric acid metabolism in hypertension or coronary heart disease in a Chinese population. *Pharmacogenomics* (2019) 20(14):1021–31. doi: 10.2217/pgs-2019-0048

54. Marais AD. Apolipoprotein e and atherosclerosis. *Curr Atheroscler Rep* (2021) 23(7):34. doi: 10.1007/s11883-021-00933-4

55. Nagahama K, Inoue T, Kohagura K, Ishihara A, Kinjo K, Ohya Y. Hyperuricemia predicts future metabolic syndrome: a 4-year follow-up study of a Large screened cohort in Okinawa, Japan. *Hypertens Res* (2013) 37(3):232–8. doi: 10.1038/hr.2013.137

56. Kawamoto R, Tabara Y, Kohara K, Kusunoki T, Abe M, Miki T. Synergistic influence of age and serum uric acid on blood pressure among community-dwelling Japanese women. *Hypertens Res* (2013) 36(7):634–8. doi: 10.1038/hr.2013.5

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 19 July 2023
DOI 10.3389/fendo.2023.1172323



OPEN ACCESS

EDITED BY
Di Liu,
Shenzhen Institute of Advanced Technology (CAS), China

REVIEWED BY
Jianhua Ma,
First Affiliated Hospital, Nanjing Medical University, China
Daniel Rubin,
Temple University, United States

*CORRESPONDENCE
Yang Zou
jxyxyzy@163.com

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

RECEIVED 23 February 2023
ACCEPTED 29 June 2023
PUBLISHED 19 July 2023

CITATION
Yang R, Kuang M, Qiu J, Yu C, Sheng G and Zou Y (2023) Assessing the usefulness of a newly proposed metabolic score for visceral fat in predicting future diabetes: results from the NAGALA cohort study. *Front. Endocrinol.* 14:1172323. doi: 10.3389/fendo.2023.1172323



# Assessing the usefulness of a newly proposed metabolic score for visceral fat in predicting future diabetes: results from the NAGALA cohort study

Ruijuan Yang[1,2†], Maobin Kuang[2,3†], Jiajun Qiu[2,3†], Changhui Yu[2,3], Guotai Sheng[4] and Yang Zou[2*]

[1]Department of Endocrinology, Jiangxi Provincial People's Hospital, Medical College of Nanchang University, Nanchang, Jiangxi, China, [2]Jiangxi Cardiovascular Research Institute, Jiangxi Provincial People's Hospital, The First Affiliated Hospital of Nanchang Medical College, Nanchang, Jiangxi, China, [3]Department of Cardiology, Jiangxi Provincial People's Hospital, Medical College of Nanchang University, Nanchang, Jiangxi, China, [4]Jiangxi Provincial Geriatric Hospital, Jiangxi Provincial People's Hospital, The First Affiliated Hospital of Nanchang Medical College, Nanchang, Jiangxi, China

**Objective:** Visceral adipose tissue assessment holds significant importance in diabetes prevention. This study aimed to explore the association between the newly proposed Metabolic Score for Visceral Fat (METS-VF) and diabetes risk and to further assess the predictive power of the baseline METS-VF for the occurrence of diabetes in different future periods.

**Methods:** This longitudinal cohort study included 15,464 subjects who underwent health screenings. The METS-VF, calculated using the formula developed by Bello-Chavolla et al., served as a surrogate marker for visceral fat obesity. The primary outcome of interest was the occurrence of diabetes during the follow-up period. Established multivariate Cox regression models and restricted cubic spline (RCS) regression models to assess the association between METS-VF and diabetes risk and its shape. Receiver operating characteristic (ROC) curves were used to compare the predictive power of METS-VF with body mass index (BMI), waist circumference (WC), waist-to-height ratio (WHtR), and visceral adiposity index (VAI) for diabetes, and time-dependent ROC analysis was conducted to assess the predictive capability of METS-VF for the occurrence of diabetes in various future periods.

**Results:** During a maximum follow-up period of 13 years, with a mean of 6.13 years, we observed that the cumulative risk of developing diabetes increased with increasing METS-VF quintiles. Multivariable-adjusted Cox regression analysis showed that each unit increase in METS-VF would increase the risk of diabetes by 68% (HR 1.68, 95% CI 1.13, 2.50), and further RCS regression analysis revealed a possible non-linear association between METS-VF and diabetes risk (*P* for non-linearity=0.002). In addition, after comparison by ROC analysis, we found that METS-VF had significantly higher predictive power for diabetes than other general/visceral adiposity indicators, and in time-dependent ROC analysis, we further considered the time-dependence of diabetes status and METS-VF and

found that METS-VF had the highest predictive value for predicting medium- and long-term (6-10 years) diabetes risk.

**Conclusion:** METS-VF, a novel indicator for assessing visceral adiposity, showed a significantly positive correlation with diabetes risk. It proved to be a superior risk marker in predicting the future onset of diabetes compared to other general/visceral adiposity indicators, particularly in forecasting medium- and long-term diabetes risk.

KEYWORDS

diabetes, METS-VF, predictive power, time-dependent ROC analysis, visceral adiposity

# Introduction

Diabetes is one of the most common chronic diseases that endangers the physical health of the world population and cause disability and death (1, 2). The treatment and management of diabetic patients heavily burden the world's healthcare systems and have become an important global public health challenge (3, 4). Under the background that diabetes currently cannot be completely cured, the early identification of people at risk of developing diabetes and primary prevention of diabetes are of great public health importance (5).

Obesity is an important risk factor for the development and progression of diabetes (6), and obese people are usually at a higher risk for diabetes. Notably, compared to general adiposity due to increased subcutaneous fat, visceral adiposity is more harmful to the organism, especially fat deposits in organs such as the liver and skeletal muscle, which cause more pronounced hepatic and peripheral insulin resistance thereby leading to the development of metabolic diseases such as diabetes (7–9). The gold standard measure for clinical assessment of visceral adiposity is magnetic resonance imaging (MRI), but it is not suitable for diabetes screening and clinical prevention in large populations due to its expensive testing costs and complex procedures (10). In addition, anthropometric abdominal adiposity indicators WC, WHtR, and waist-to-hip ratio can also indicate the risk of visceral adiposity, but they cannot accurately distinguish between abdominal visceral adipose tissue and subcutaneous adipose tissue (11).

METS-VF is a newly developed surrogate for assessing visceral adiposity that integrates demographic parameters (age and sex), anthropometric obesity parameters (BMI and WHtR), and glycemic lipid parameters [fasting plasma glucose (FPG), triglyceride (TG), and high-density lipoprotein cholesterol (HDL-C)]. It was developed by Bello-Chavolla OY et al. Following validation and comparison by Bello-Chavolla OY et al., METS-VF was found to provide a significantly superior estimate of human visceral adiposity compared to other commonly used surrogate indicators for abdominal adiposity. Moreover, it exhibited high agreement with gold standard measurements (12). Several subsequent observational studies have shown that METS-VF had good risk assessment/predictive power for metabolic diseases closely related to visceral adiposities such as chronic kidney disease, hypertension, and hyperuricemia (13–16). However, the correlation between METS-VF and diabetes risk has only been explored in a rural population in China (17), and the predictive power of baseline METS-VF for the future development of diabetes in the general population and the effect of temporal progression on the predictive power of METS-VF are currently unknown. Therefore, the current study comprehensively analyzed and compared the risk assessment/predictive ability of METS-VF for diabetes based on a larger sample size general population cohort and further explored the predictive power of METS-VF for the occurrence of diabetes in different future periods using time-dependent ROC analysis.

# Methods

## Study design and ethics approval

We conducted a retrospective cohort study of subjects in the NAGALA cohort (NAfld in the Gifu Area, Longitudinal Analysis) to assess the usefulness of the newly proposed METS-VF for predicting future diabetes. Information on the NAGALA cohort study was described in detail in a previously published article (18). In brief, the NAGALA cohort was established in 1994 and included a study sample of people who underwent health screenings at Murakami Memorial Hospital. Given that the vast majority of people who underwent health screenings at the hospital will have repeat screenings in the future, with 60% of these subjects receiving one or two health screenings per year, the NAGALA research project team conducted a long-term follow-up survey for future incident diabetes and incident non-alcoholic fatty liver disease. In a previously published article, Prof. Okamura reported that the NAGALA cohort study was approved by the Murakami Memorial Hospital ethics committee and written informed consent was obtained from all study participants, and that detailed data from the study were uploaded to the Dryad public database for sharing (19). The current study is a secondary analysis of the NAGALA study, and the subjects' identifying information has been anonymized in the data set used. Therefore, the Ethics Committee of Jiangxi Provincial People's Hospital waived the process of

obtaining written informed consent for the current study, approved the protocol of the current study, and supervised the entire process of the current study. See STROBE statement (S1Text).

## Study population

The current study extracted data from the NAGALA cohort of 20,944 subjects who underwent health screenings between May 1994 and December 2016. We further excluded subjects with the following conditions according to the study objectives: (1) At baseline, 323 who had been diagnosed with diabetes, 416 with liver disease (other than fatty liver), and 808 with FPG ≥6.1mmol/L; (2) 2,321 who were taking medications at baseline, 739 with excessive alcohol consumption (20), and 863 with incomplete data; (3) 10 who withdrew from the study during follow-up for unknown reasons. Ultimately, 15,464 subjects were included in the current study for analysis, and the detailed flow chart was shown in Figure 1.

## Baseline data collection and definition of diabetes

Standardized trained medical examiners collected basic information on sex, age, smoking and drinking status, and exercise habits by means of a questionnaire. Smoking status was defined using never, past and current smoking; drinking status was defined as non/small, light, moderate, and heavy drinking based on the subject's weekly alcohol consumption in the month prior to study participation (20); and having an exercise habit was defined as the subject having at least one physical activity per week. Anthropometric indicators of height, weight, WC, and systolic and diastolic blood pressure (SBP and DBP) were measured indoors with subjects wearing light clothing and no shoes using standard methods. Fatty liver diagnosis was based on the evaluation of liver contrast and brightness in abdominal ultrasound images by gastroenterologists (18, 21). In addition, forearm venous blood samples were drawn from subjects after fasting for at least 8 hours and sent to a standard laboratory, and then using an automated biochemical analyzer measured concentrations of biochemical parameters such as aspartate aminotransferase (AST), HDL-C, alanine aminotransferase (ALT), glycosylated hemoglobin (HbA1c), FPG, gamma-glutamyl transferase (GGT), TG, and total cholesterol (TC).

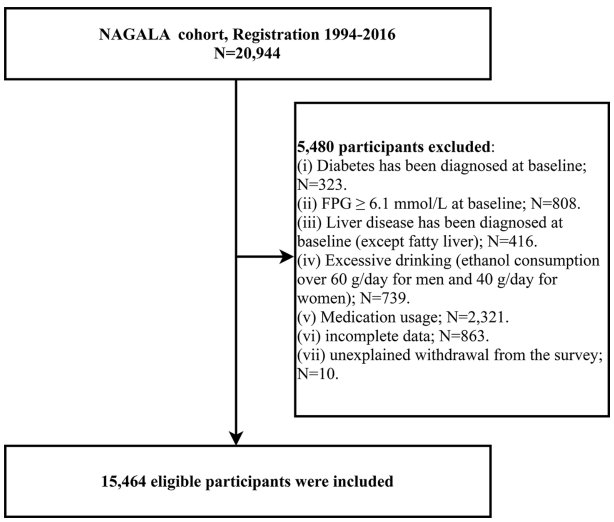



FIGURE 1
Flowchart of the selection process of study subjects.

## Primary outcome

The incidence of diabetes among the subjects during the follow-up period was considered the primary outcome in the current study. According to the American Diabetes Association criteria, diabetes was defined as HbA1c ≥6.5% or FPG ≥7.0 mmol/L measured during follow-up or self-reported diabetes (verified through the examination of subjects' medical records or blood glucose measurements) by the subject (22).

## Calculation formulas for METS-VF, BMI, WHtR, and VAI

METS-VF = 4.466 + 0.011*[(Ln((Ln((2 * FPG) + TG) * BMI)/(Ln(HDL- C))))$^3$] +3.239*[(Ln(WHtR))$^3$] + 0.319 * sex + 0.594 * (Ln(age)) (12). Note: sex in the METS-VF calculation formula was a binary response variable (men=1, women=0).

$$BMI = weight(kg)[height(m)]^2$$

$$WHtR = WC(cm)/height(cm)$$

$$VAI(men) = (WC/(39.68 + (1.88*BMI))*(TG/1.03)*(1.31/HDL - C)$$

(23)

$$VAI(women) = (WC/(36.58 + (1.89*BMI))*(TG/0.81)*(1.52/HDL - C)$$

(23)

## Statistical analysis

All statistical analyses for the current study were done on R Language 3.4.3 and Empower(R) 2.0 software and were set to be significant at two-sided $P<0.05$. METS-VF values were calculated and all subjects were grouped according to quintiles of METS-VF values [Quintile 1 (Q1)<5.03, Q2 (5.03 to 5.58), Q3 (5.58 to 6.00), Q4 (6.00 to 6.42), Q5 ≥6.42] using the quantile function. Described the baseline data of the subjects according to the quintiles of METS-VF, and chose different description methods and comparison methods between groups according to the type of data; among

them, continuous variables with normal and skewed distribution were described as mean (standard deviation) and median (interquartile range), respectively, and comparisons between groups were performed using one-way ANOVA and Kruskal-Willis H test, respectively, while categorical variables were described as frequencies (%), and comparisons between groups were made using chi-square tests. In addition, we used Kaplan-Meier curves to describe the cumulative hazard of developing diabetes in each METS-VT quintile during the follow-up period and subsequently examined the differences between the groups using log-rank tests and finally made a preliminary determination of whether the proportional hazards assumption for establishing multivariate Cox regression models was met based on the results of Kaplan-Meier analysis (24).

To clarify the association between baseline indicators and diabetes risk and to initially explore the association of METS-VF with diabetes risk, we first estimated the hazard ratio (HR) and 95% confidence interval (CI) for each baseline indicator associated with diabetes risk using univariate Cox regression analysis. Subsequently, we checked for collinearity between all covariates and METS-VF by multiple linear regression analysis and excluded collinear variables with a final variance inflation factor greater than 5 from later model adjustments (25). According to the recommendations of the STROBE guidelines (26), we established four stepwise adjusted multivariate Cox regression models; Model 1 was adjusted for age, sex, and BMI; Model 2 considered the potential effects of fatty liver and lifestyle-related factors (smoking and drinking status and exercise habits) on the basis of Model 1; Model 3 was further adjusted for liver function-related indicators (ALT, AST, and GGT); finally, Model 4 continued to adjust for SBP, TG, HDL-C, TC, and HbA1c on the basis of Model 3. We incorporated METS-VF into 4 multivariate Cox regression models as continuous variables and categorical variables in quintiles, respectively, and calculated trends associated with diabetes risk based on the median of METS-VF quintiles in the models. Furthermore, to detect any possible linear or non-linear dependence between METS-VF, BMI, WC, WHtR, and VAI and diabetes risk, we utilized a 4-knot RCS model to fit dose-response curves for these variables at the 5th, 35th, 65th, and 95th percentiles. Prior to plotting the dose-response curves, we also conducted separate collinearity screenings to examine the presence of collinearity between BMI, WC, WHtR, VAI, and other covariates. Based on the results of the collinearity screening analysis, we adjusted for covariates that showed no collinearity with the respective obesity indicators in the RCS regression models.

ROC curves were constructed and the area under the curves (AUCs) was calculated to assess the predictive power of baseline METS-VF and several traditional visceral adiposity indicators, WC, WHtR, VAI, and BMI, for diabetes, and the differences in predictive power between METS-VF and the other indicators were compared using the DeLong test (27). Additionally, to assess the effect of time factors on the ability of METS-VF to predict the future occurrence of diabetes, we also calculated the AUCs, optimal thresholds, sensitivity, and specificity of baseline METS-VF for predicting the occurrence of diabetes at each time point from 2 to 12 years in the future using time-dependent ROC analysis. Subsequently, we evaluated the calibration of the predictive model by plotting calibration curves to assess the agreement between predicted probabilities and observed probabilities; internal validation was conducted using the bootstrap algorithm with 1,000 repetitions (28).

# Results

## Baseline characterization

After screening the study population according to inclusion and exclusion criteria, a total of 15,464 subjects were eventually enrolled in the current study (Figure 1), with a mean age of 43.71 years, of which 54.51% were men. Table 1 groups all subjects according to the quintiles of METS-VF and describes and compares the baseline information of each group; we found that with the increase of METS-VF quintile, the proportion of subjects who were men, fatty liver patients, alcohol drinkers, and current and past smokers all gradually increased, while the proportion of those with an exercise habit gradually decreased (All $P<0.001$). Regarding the anthropometric indicators and biochemical parameters of the subjects, except for HDL-C levels, which decreased with the increase of METS-VF quintile, the levels of other indicators such as age, height, weight, BMI, WC, ALT, AST, GGT, TC, TG, FPG, HbA1c, SBP, and DBP increased with the increase of METS-VF quintile (All $P<0.001$).

During a follow-up period of up to 13 years with an average duration of 6.13 years, a total of 373 individuals developed diabetes, resulting in an incidence rate of 39.88/10,000 person-years. Notably, the incidence rate of diabetes demonstrated a gradual upward trend across the quintiles of METS-VF. Specifically, the incidence rates for Q1-Q5 were 0.4%, 0.9%, 1.3%, 2.5%, and 7.0%, respectively. Moreover, we used Kaplan-Meier curves to describe the cumulative hazard of developing diabetes in each METS-VF quintile during the follow-up period (Figure 2), and the results also showed a progressive increase in the risk of developing diabetes with increasing METS-VF quintiles and no significant intersection of the curves (log-rank $P<0.0001$), which also suggested that our data followed the proportional hazard assumption.

## Association of METS-VF with diabetes

Supplementary Table 1 shows the results of the univariate Cox regression analysis between baseline variables and diabetes risk, where we found that all baseline variables were significantly associated with diabetes risk ($P<0.0001$) except for exercise habits, which was borderline positive [(HR 0.76, 95% CI 0.56, 1.02), $P=0.0641$], where each unit increase in METS-VF increased the risk of diabetes by 414% (HR 5.14, 95% CI 4.27, 6.19). To further explore the independent association of METS-VF with diabetes risk, we included METS-VF as continuous and categorical variables, respectively, in four multivariate Cox regression models (Table 2), in which the non-collinear variables were adjusted stepwise while the collinear variables weight, WC, and DBP were excluded (Supplementary Table 2). When we preliminarily adjusted age,

TABLE 1 Baseline characteristics of subjects and incidence of diabetes grouped according to METS-VF quintiles.

| | METS-VF quintiles | | | | | *P*-value |
|---|---|---|---|---|---|---|
| | Quintile 1 (< 5.03) | Quintile 2 (5.03 to 5.58) | Quintile 3 (5.58 to 6.00) | Quintile 4 (6.00 to 6.42) | Quintile 5 (≥ 6.42) | |
| Subjects, n | 3091 | 3090 | 3091 | 3090 | 3091 | |
| Sex | | | | | | <0.001 |
| Women | 2405 (77.81%) | 1834 (59.35%) | 1297 (41.96%) | 910 (29.45%) | 588 (19.02%) | |
| Man | 686 (22.19%) | 1256 (40.65%) | 1794 (58.04%) | 2180 (70.55%) | 2503 (80.98%) | |
| Age, year | 38.00 (35.00-43.00) | 40.00 (36.00-47.00) | 42.00 (37.00-49.00) | 45.00 (39.00-52.00) | 49.00 (41.00-55.00) | <0.001 |
| Height, m | 1.63 (0.08) | 1.64 (0.09) | 1.66 (0.09) | 1.67 (0.08) | 1.67 (0.08) | <0.001 |
| Weight, kg | 49.66 (6.32) | 55.36 (7.41) | 60.22 (8.10) | 65.10 (8.83) | 72.82 (10.95) | <0.001 |
| BMI, $kg/m^2$ | 18.73 (1.45) | 20.51 (1.42) | 21.82 (1.52) | 23.39 (1.73) | 26.13 (2.76) | <0.001 |
| WC, cm | 65.09 (3.91) | 71.33 (3.40) | 76.32 (3.41) | 80.97 (3.77) | 88.63 (6.16) | <0.001 |
| ALT, U/L | 13.00 (11.00-17.00) | 14.00 (11.00-18.00) | 16.00 (12.00-22.00) | 19.00 (15.00-26.00) | 24.00 (18.00-34.00) | <0.001 |
| AST, U/L | 16.00 (13.00-19.00) | 16.00 (13.00-19.00) | 17.00 (14.00-21.00) | 18.00 (15.00-22.00) | 20.00 (16.00-25.00) | <0.001 |
| GGT, U/L | 12.00 (9.00-15.00) | 13.00 (10.00-17.00) | 15.00 (11.00-21.00) | 18.00 (13.00-27.00) | 23.00 (16.00-34.00) | <0.001 |
| HDL-C, mmol/L | 1.71 (0.38 | 1.59 (0.39) | 1.47 (0.37) | 1.34 (0.35) | 1.20 (0.30) | <0.001 |
| TC, mmol/L | 4.82 (0.81) | 4.95 (0.81) | 5.13 (0.84) | 5.27 (0.86) | 5.46 (0.85) | <0.001 |
| TG, mmol/L | 0.50 (0.37-0.69) | 0.59 (0.42-0.81) | 0.75 (0.53-1.05) | 0.93 (0.64-1.33) | 1.19 (0.84-1.72) | <0.001 |
| FPG, mg/dL | 88.77 (6.92) | 90.95 (6.89) | 93.06 (6.90) | 95.02 (6.88) | 97.01 (6.63) | <0.001 |
| HbA1c, % | 5.10 (0.30) | 5.13 (0.30) | 5.16 (0.31) | 5.20 (0.32) | 5.27 (0.34) | <0.001 |
| SBP, mmHg | 105.32 (12.22) | 109.63 (12.55) | 114.21 (13.06) | 118.55 (13.50) | 124.75 (15.23) | <0.001 |
| DBP, mmHg | 65.14 (8.37) | 67.99 (8.88) | 71.17 (9.47) | 74.62 (9.52) | 78.97 (10.22) | <0.001 |
| Exercise habits | 510 (16.50%) | 574 (18.58%) | 594 (19.22%) | 550 (17.80%) | 478 (15.46%) | <0.001 |
| Fatty liver | 9 (0.29%) | 77 (2.49%) | 291 (9.41%) | 761 (24.63%) | 1599 (51.73%) | <0.001 |
| Drinking status | | | | | | <0.001 |
| Non/small | 2720 (88.00%) | 2494 (80.71%) | 2305 (74.57%) | 2193 (70.97%) | 2090 (67.62%) | |
| Light | 213 (6.89%) | 330 (10.68%) | 397 (12.84%) | 405 (13.11%) | 409 (13.23%) | |
| Moderate | 131 (4.24%) | 214 (6.93%) | 279 (9.03%) | 324 (10.49%) | 409 (13.23%) | |
| Heavy | 27 (0.87%) | 52 (1.68%) | 110 (3.56%) | 168 (5.44%) | 183 (5.92%) | |
| Smoking status | | | | | | <0.001 |
| Never | 2395 (77.48%) | 2045 (66.18%) | 1769 (57.23%) | 1554 (50.29%) | 1264 (40.89%) | |
| Past | 273 (8.83%) | 455 (14.72%) | 604 (19.54%) | 722 (23.37%) | 895 (28.96%) | |
| Current | 423 (13.68%) | 590 (19.09%) | 718 (23.23%) | 814 (26.34%) | 932 (30.15%) | |
| Diabetes incidence | 11 (0.4%) | 27 (0.9%) | 41 (1.3%) | 77 (2.5%) | 217 (7.0%) | <0.001 |

Values were expressed as mean (SD) or medians (quartile interval) or n (%). BMI, body mass index; WC, Waist circumference; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase; HDL-C, high-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride; HbA1c, hemoglobin A1c; FPG, fasting plasma glucose; SBP, systolic blood pressure; DBP, Diastolic blood pressure; METS-VF, Metabolic Score for Visceral Fat.

sex, and BMI in Model 1, we found that METS-VF as a continuous variable remained significantly positively correlated with diabetes risk (HR 2.81, 95% CI 1.92, 4.12), while as a categorical variable, taking Q1 as a reference, the risk of diabetes increased gradually with the increase of METS-VF quintile and the two were linearly correlated (*P*-trend<0.001). After further adjusting the fatty liver and lifestyle indicators (Model 2), and liver function-related parameters (Model 3), the HR of METS-VF associated with diabetes risk decreased slightly, while the direction and linear trend of the association remained unchanged. Ultimately, we additionally adjusted for SBP, TG, HDL-C, TC, and HbA1c in Model 4 and found that each unit increase in METS-VF would

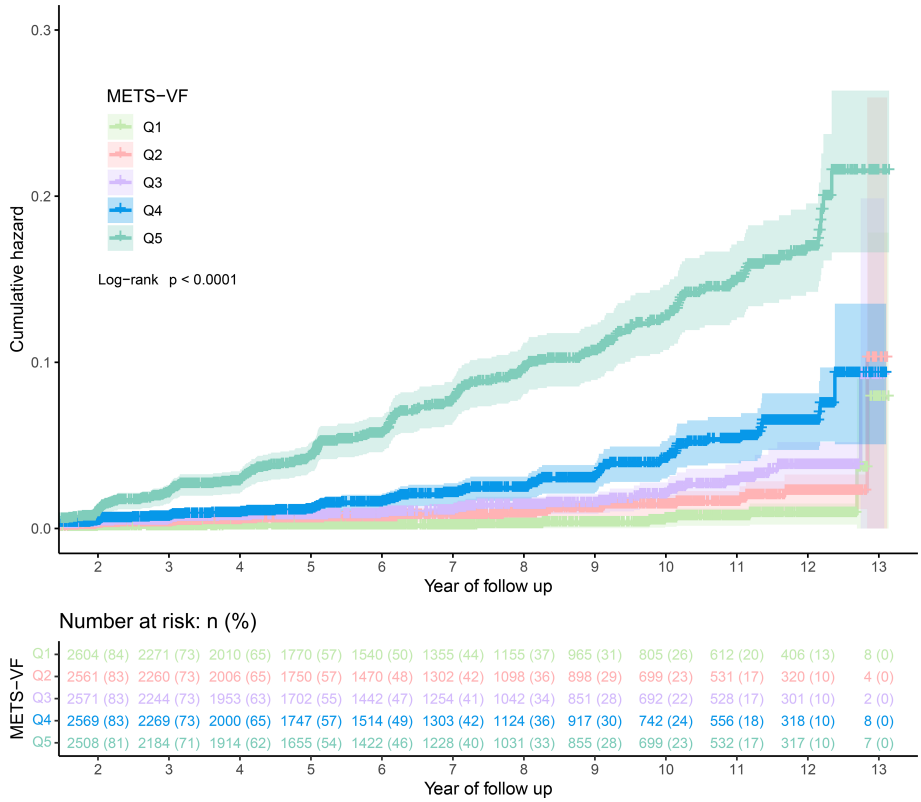



FIGURE 2
Kaplan-meier curve of METS-VF quartiles over time. METS-VF, Metabolic Score for Visceral Fat.

increase the risk of diabetes by 68% (HR 1.68, 95% CI 1.13, 2.50); in addition, Q5 still had the highest diabetes risk (HR 2.15, 95% CI 0.98, 4.70) with Q1 as a reference in Model 4, but the linear association between METS-VF quintiles and diabetes risk was not significant after trend test (*P*-trend=0.0946), which suggested that there may be a non-linear relationship between the two.

## Non-linear association between METS-VF, BMI, WC, WHtR and VAI and diabetes risk

We employed a 4-knot RCS regression model to fit the dose-response curves for METS-VF, BMI, WC, WHtR, and VAI in relation to the risk of diabetes. Adjustments for non-collinear variables were made in the corresponding RCS regression models based on the results of collinearity analysis (Supplementary Tables 2-6). The RCS analysis revealed that the association between METS-VF and diabetes risk was non-linear (*P* for non-linearity=0.002) (Figure 3); when the METS-VF value was in the Q3 (5.58-6.00) interval, the slope of the curve increased significantly with the increase of METS-VF, implying that METS-VF had a stronger correlation with diabetes risk in the Q4 and Q5 intervals compared to the Q1 and Q2 intervals. Moreover, BMI, WC, WHtR, and VAI demonstrated similar shapes of association with the risk of diabetes, with evident threshold points on the curves, and all exhibiting non-linear correlations (Supplementary Figures 1-4; All *P* for non-linearity<0.05).

## Comparison of METS-VF with BMI, WC, WHtR, and VAI in predicting diabetes and time-dependent ROC analysis

Table 3 shows the AUC, Sensitivity, Specificity, Positive predictive value and Negative predictive value (NPV) of METS-VF, BMI, WC, WHtR, and VAI for predicting diabetes. Overall, BMI, VAI, WC, WHtR, and METS-VF all had a good predictive performance for diabetes with AUC values of 0.73 (0.71, 0.76), 0.74 (0.71, 0.77), 0.74 (0.72, 0.77), 0.74 (0.72, 0.77), 0.77 (0.75, 0.80), respectively. After comparison, it was found that METS-VF had a significantly higher AUC value (0.77) than other indicators, showing the highest predictive accuracy for future diabetes risk (All *P*<0.05, DeLong test). In addition, all the aforementioned indicators of visceral obesity exhibited high NPV, with METS-VF having the highest NPV of 99.10%.

This study also used time-dependent ROC analysis to further explore the predictive power of METS-VF for each time point over the next 2-12 years regarding the occurrence of diabetes (Table 4). Additionally, the calibration ability of METS-VF in predicting long-term diabetes risk (7-12 years) was evaluated using calibration curves (Figure 4). The results of the analysis showed that the predictive power of METS-VF for future diabetes risk gradually increased from the 2nd year of follow-up, until the AUC reached the highest value of 0.79 at the 7th and 8th year of follow-up, and then gradually decreased from the 9th year; specifically, METS-VF had higher AUC values (>0.77) and more stable prediction thresholds

TABLE 2 Multivariable Cox regression analyses for the association between METS-VF and the incidence of diabetes.

| | HR (95%CI) | | | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 |
| METS-VF (continuous) | 2.81 (1.92, 4.12) | 2.75 (1.87, 4.03) | 2.45 (1.67, 3.58) | 1.68 (1.13, 2.50) |
| Quintile 1 | Ref | Ref | Ref | Ref |
| Quintile 2 | 1.70 (0.83, 3.45) | 1.70 (0.84, 3.46) | 1.77 (0.87, 3.60) | 1.63 (0.80, 3.33) |
| Quintile 3 | 2.01 (1.00, 4.02) | 2.03 (1.01, 4.05) | 2.05 (1.03, 4.10) | 1.60 (0.79, 3.24) |
| Quintile 4 | 2.51 (1.26, 5.02) | 2.52 (1.26, 5.03) | 2.48 (1.24, 4.95) | 1.67 (0.82, 3.40) |
| Quintile 5 | 4.08 (1.94, 8.60) | 4.13 (1.96, 8.71) | 3.81 (1.80, 8.06) | 2.15 (0.98, 4.70) |
| *P*-trend | <0.001 | <0.001 | 0.0001 | 0.0946 |

HR, Hazard ratio; CI, confidence interval; other abbreviations as in Table 1.
Model 1 adjusted for age, sex, and BMI.
Model 2 adjusted for age, sex, BMI, fatty liver, habits of exercise, smoking status, and drinking status.
Model 3 adjusted for age, sex, BMI, fatty liver, habits of exercise, smoking status, drinking status, ALT, AST, and GGT.
Model 4 adjusted for age, sex, BMI, fatty liver, habits of exercise, smoking status, drinking status, ALT, AST GGT, SBP, TG, HDL-C, TC, and HbA1c.

(6.03-6.37) for predicting diabetes over the next 6-10 years, which was an ideal risk marker for predicting the occurrence of diabetes in the future medium- and long-term. Furthermore, the calibration curves in Figure 4 demonstrated that the predicted diabetes risk by METS-VF aligned well with the observed diabetes risk in the year-7 to year-12 period. This indicated that METS-VF had a reliable predictive accuracy for diabetes.

## Discussion

In this longitudinal cohort study conducted on a large general population, we had the following important findings: (1) There was a significant and positive correlation between METS-VF, a novel indicator for assessing visceral adiposity, and diabetes risk, but this correlation may be non-linear, and when METS-VF exceeded the Q3 (5.58-6.00) interval, its correlation with diabetes risk was further enhanced. (2) METS-VF demonstrated significantly better performance compared to several other commonly used surrogate indicators of visceral adiposity, VAI, WC, WHtR, and BMI, in predicting future diabetes risk. (3) For the first time, we discovered that METS-VF exhibited higher AUC values and more stable predictive thresholds for predicting diabetes risk over the next 6-10 years, and was an ideal risk marker for future medium- to long-term diabetes risk.

The global prevalence of diabetes and obesity has shown an almost parallel increase in recent years, particularly in Asian populations, primarily in East Asia (6, 29). Importantly,

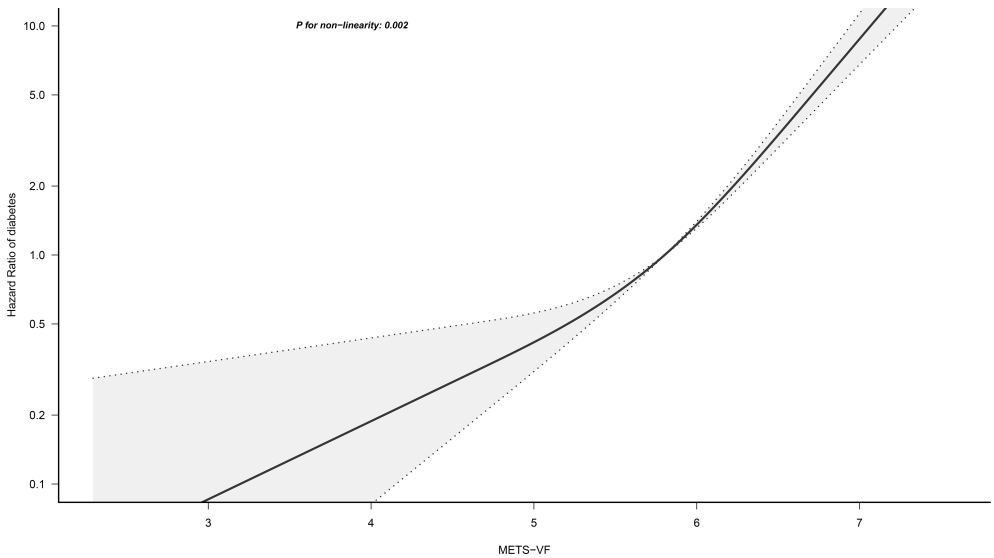



FIGURE 3
Restricted cubic spline analysis of METS-VF for the estimation of the risk of diabetes. METS-VF, Metabolic Score for Visceral Fat. Restricted cubic spline model adjusted for sex, age, fatty liver, height, BMI, exercise habits, ALT, AST, GGT, HDL-C, TC, TG, HbA1c, drinking status, smoking status, FPG, and SBP.

TABLE 3 Area under the ROC curve, Sensitivity, Specificity, PPV, and NPV of METS-VF, BMI, WC, VAI, and WHtR to predict diabetes.

| | AUC | 95%CI low | 95%CI up | Specificity | Sensitivity | PPV | NPV |
|---|---|---|---|---|---|---|---|
| BMI* | 0.7327 | 0.7068 | 0.7585 | 71.82% | 62.73% | 5.22% | 98.73% |
| VAI* | 0.7410 | 0.7145 | 0.7674 | 68.18% | 71.58% | 5.27% | 98.98% |
| WC* | 0.7424 | 0.7164 | 0.7685 | 71.63% | 65.42% | 5.39% | 98.82% |
| WHtR* | 0.7424 | 0.7167 | 0.7682 | 77.01% | 60.32% | 6.09% | 98.74% |
| METS-VF | 0.7731 | 0.7493 | 0.7969 | 65.07% | 76.14% | 5.12% | 99.10% |

ROC, receiver-operating characteristic curve; AUC, area under the ROC curve; PPV, positive predictive value; NPV, negative predictive value; VAI, visceral adiposity index; WHtR, waist-to-height index; CI, confidence interval; Other abbreviations as in Table 1; *, P<0.05 compared with METS-VF.

epidemiological evidence indicates that the overall body fat content in Asian populations is typically lower compared to Western populations. However, abdominal obesity, characterized by the accumulation of fat around the abdomen, is a prominent feature of obesity in Asian populations (30–32). It is also a significant risk factor for metabolic disorders such as diabetes and cardiovascular disease (33, 34). In previous studies related to the mechanism of abdominal obesity leading to metabolic complications, most of them emphasized the importance of increased visceral fat rather than subcutaneous fat (35, 36), because subcutaneous adipose tissue is considered to be the largest and least metabolically harmful storage site for excess fat in the body (37), while deposition of excess adipose tissue such as ceramide or diacylglycerol in organs such as the liver and skeletal muscle will cause endocrine dysfunction, dysfunction of pro-inflammatory factors and mitochondrial dysfunction in visceral adipose tissue and also lead to increased levels of free fatty acids thereby antagonizing hepatic insulin (38–40). Given that abdominal subcutaneous adipose tissue and visceral adipose tissue may have opposite biological functions on the body's glucose metabolism, accurate differentiation and measurement of visceral adipose tissue will help to assess and predict the occurrence and progression of diabetes.

Although MRI techniques and dual-energy X-ray absorptiometry (DXA) techniques can currently be used in clinical practice to accurately measure visceral fat content, the high economic and technical costs of these ancillary techniques make them unsuitable for use in primary health care and large-scale epidemiological investigations of diabetes (10). To address this issue, a large number of researchers are working to develop simple parameters that can more accurately identify and assess visceral adiposity. METS-VF is a new parameter for assessing visceral adiposity tissue developed and validated by Bello-Chavolla OY et al. in July 2019, and the detailed steps of its development and validation have been described elsewhere (12). Briefly, Bello-Chavolla OY et al. prospectively recruited a discovery cohort of 366 subjects with DXA measurements from healthcare institutions and used the visceral fat content of subjects obtained from DXA measurements as the dependent variable, and used several simple indicators (metabolic score for insulin resistance, age, sex, and WHtR), which are considered to be closely related to visceral fat content, as independent variables (41, 42), and then used non-linear regression analysis to fit the prediction model with the highest agreement with DXA measurements, namely METS-VF. METS-VF was subsequently validated by applying it to two validation cohorts of subjects with DXA+MRI measurements and subjects with bio-electrical impedance measurements, respectively. Their results showed that METS-VF was more accurate in predicting visceral adiposity than other obesity indicators such as WC, WHtR, VAI, and BMI both in the discovery cohort and in the validation cohort. In several subsequently published observational studies, cross-sectional data from Yu P et al. and longitudinal cohort data from Feng L et al. showed that METS-VF was significantly and independently positively associated with chronic kidney disease and had stronger risk assessment/predictive power for chronic kidney disease compared to other obesity indicators (13, 14). In addition, METS-VF has also been shown to be an independent predictor of hypertension and hyperuricemia (15, 16).

In the current study, we found a significant association between METS-VF and diabetes risk after adjusting for a large number of confounding factors associated with diabetes risk, with each unit increase in METS-VF increasing the risk of diabetes by 68%. Additionally, by observing Model 4 in Table 2 and the dose-response curve in Figure 3, we found that there may be a non-linear correlation between METS-VF and diabetes risk, with a change in the correlation around the Q3 (5.58 to 6.00) interval of METS-VF, and a significantly stronger correlation with diabetes

TABLE 4 Areas under the time-dependent ROC curves, Best thresholds, Sensitivity, and Specificity for METS-VF predicting future diabetes risk.

| | 2-years | 3-years | 4-years | 5-years | 6-years | 7-years | 8-years | 9-years | 10-years | 11-years | 12-years |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AUC | 0.70 | 0.75 | 0.74 | 0.76 | 0.77 | 0.79 | 0.79 | 0.77 | 0.77 | 0.76 | 0.75 |
| Best threshold | 6.14 | 6.34 | 5.93 | 6.36 | 6.37 | 6.13 | 6.37 | 6.06 | 6.03 | 6.03 | 6.08 |
| Sensitivity | 64.77% | 59.63% | 77.01% | 61.13% | 61.71% | 76.21% | 63.92% | 76.76% | 77.15% | 76.58% | 71.06% |
| Specificity | 67.05% | 76.62% | 56.58% | 78.25% | 78.79% | 67.00% | 79.02% | 63.84% | 62.76% | 63.25% | 65.98% |

AUC, area under the time-dependent ROC curves.

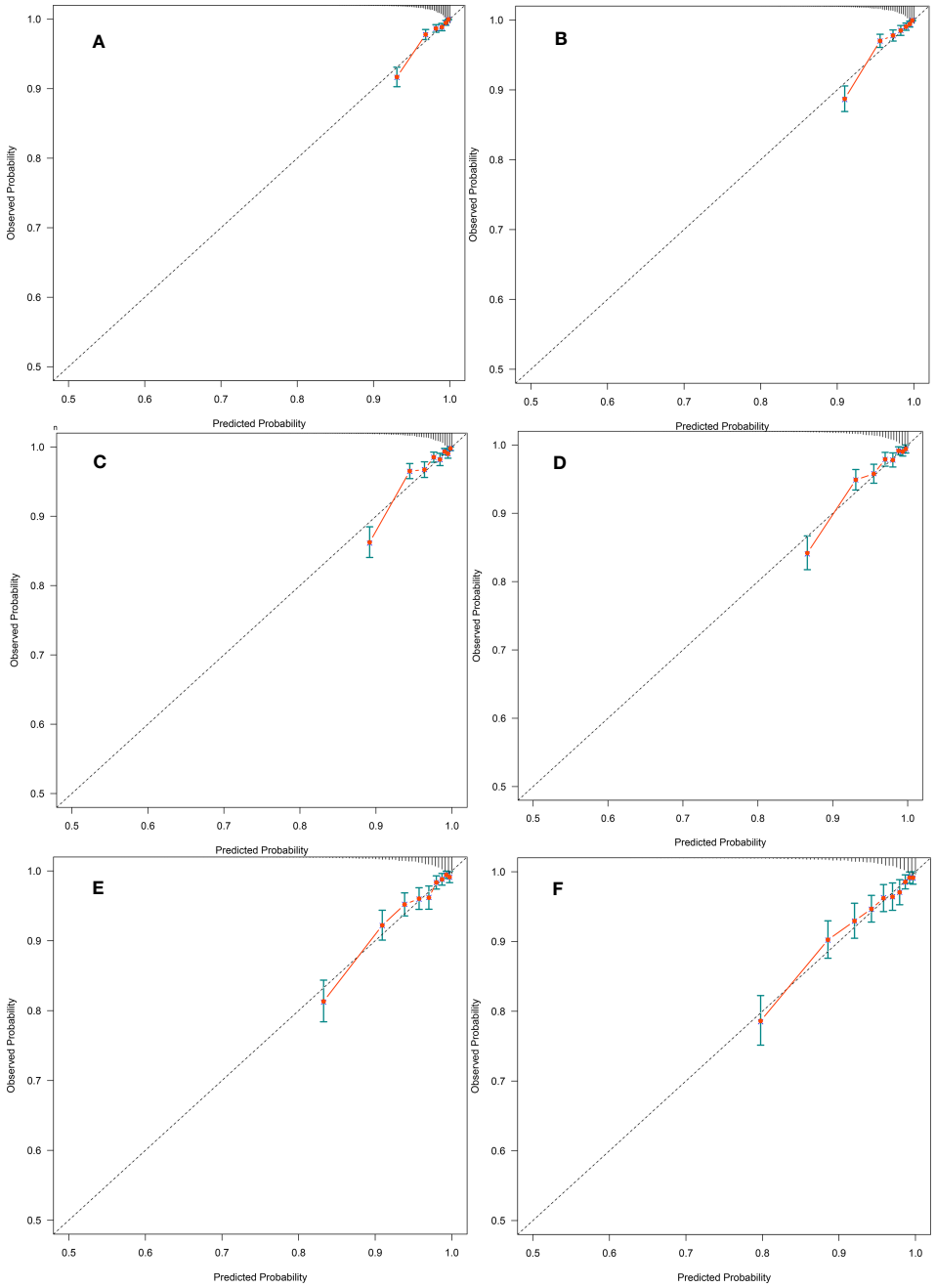



FIGURE 4
**(A–F)** were calibration curves of the prediction for diabetes event at year-7 to year-12, respectively. Dashed lines on the diagonal are reference lines.

risk when METS-VF was located in the Q4 and Q5 intervals than in the Q1 and Q2 intervals. This finding was consistent with the findings of Feng Y et al. who also found a non-linear association between METS-VF and type 2 diabetes in a study of a rural population in Henan, China (17). Therefore, we recommend that both healthy and diabetic people should control their fat intake, body weight, and WC to keep METS-VF below the Q3 interval (METS-VF<6) as much as possible to minimize the risk of diabetes. Furthermore, in line with the conclusions of Bello-Chavolla OY et al, the results of the ROC analysis of the current study showed that METS-VF had a significantly better predictive performance for diabetes compared to WC, WHtR, VAI, and BMI (All $P<0.05$, DeLong test), which may be thanks to its higher predictive accuracy for visceral adiposity (12). It is worth mentioning that in the study by Feng Y et al., they found that although METS-VF had the highest AUC value (0.69) for predicting diabetes compared to other obesity indicators, the power of METS-VF was not significantly different from WC and WHtR for predicting diabetes ($P=0.058$) (17); this result may be related to the smaller study population and relatively short follow-up period (up to 6 years) of Feng Y et al.

Based on the longest 13-year follow-up data of 15,464 subjects, the current study used time-dependent ROC analysis to further

explore the predictive power of METS-VF for the occurrence of diabetes at each time point over the next 2-12 years, showing that the predictive power of METS-VF for diabetes exhibited a slowly increasing trend from year 2 to year 7 of follow-up, while the highest predictive power was reached in years 7 and 8 (AUC=0.79), followed by a gradual decrease in the predictive power of METS-VF from year 9 to year 12. Therefore, it was more accurate to say that METS-VF should be more suitable for predicting future diabetes risk in the medium- and long-term (6-10 years), whereas the maximum 6-year follow-up period in the study by Feng Y et al. may have led them to underestimate the predictive power of METS-VF for future diabetes risk (17). Furthermore, it is worth noting that although the study by Feng Y et al. was also a longitudinal cohort study with follow-up, the time-dependence of diabetes status and METS-VF was not considered in their ROC analysis, which may also lead to some bias in their results (43). In summary, the results of the current study regarding the predictive power of METS-VF for the occurrence of diabetes in different future periods obtained by using time-dependent ROC analysis were more realistic and reliable (44). Given the higher predictive accuracy and more stable predictive thresholds of METS-VF for medium- and long-term diabetes risk, we recommended adding METS-VF to patients' physical examination reports in primary health care and clinical practice as a novel risk marker for predicting future medium- and long-term diabetes risk and, meanwhile, we believed it was relatively safe to keep METS-VF below 6.

### Strengths and limitations

The strengths of the current study are the following: (1) the current study has a larger sample size (n=15,464) and a longer follow-up period (up to 13 years) of the general population cohort compared to the previous studies. (2) The current study explored the predictive power of baseline METS-VF for the occurrence of diabetes in different future periods using time-dependent ROC analysis, and for the first time, it was clear that METS-VF may be most suitable for predicting the risk of diabetes in the medium- and long-term (6-10 years), which provided a more accurate reference for the application of METS-VF for diabetes screening and prevention in primary health care.

This study also has the following limitations: (1) Subjects in the current study did not undergo MRI or DXA examinations to measure visceral fat mass, so we were unable to further compare the correlation between METS-VF and actual visceral fat mass and diabetes risk. (2) Diabetes was defined based on HbA1c ≥6.5% or FPG ≥7.0 mmol/L or subject self-report and did not include patients with abnormal 2-hour postprandial glucose, which may underestimate the correlation between METS-VF and diabetes risk. (3) The current study did not distinguish between types of diabetes, but considering that insulin resistance due to visceral adiposity is the pathogenesis of type 2 diabetes, and that type 2 diabetes accounts for more than 95% of all diabetes, and that type 1 diabetes and type 2 diabetes have different pathogenic characteristics, the results of the current study may be more applicable to type 2 diabetes (45, 46). (4) Although the current study adjusted a large number of confounding factors related to the risk of diabetes, there may still be some risk factors for diabetes that have not been adjusted due to it being an observational study, which may lead to some residual confounding. (5) The current study was a single-center cohort study, so the applicability of the findings to other ethnic populations will need to be further validated in future studies. (6) The current study did not repeat the measurement of all baseline indicators for the subjects during the follow-up period, which limited further exploration of the impact of dynamic changes in METS-VF on the risk of developing diabetes. This aspect needed to be further investigated in future studies.

## Conclusion

In conclusion, the current study demonstrated a significant positive correlation between METS-VF, a novel indicator for assessing visceral adiposity, and the risk of diabetes in the general population. Furthermore, compared to other surrogate indicators for general/visceral adiposity (BMI, WC, WHtR, VAI), baseline METS-VF had a better predictive performance for future diabetes risk and was particularly suitable for predicting future diabetes risk in the medium- and long-term.

## Data availability statement

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

## Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Jiangxi Provincial People's Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## Author contributions

YZ, RY, MK, JQ, CY and GS conceived the research, drafted the manuscript, and did the statistical analysis. YZ revised the manuscript and designed the study. All authors read and approved the final manuscript.

## Funding

This work was supported by Shanghai Hutchison Pharmaceuticals Research Fund (No. 2019009) and the National Key Research and Development Program "Modernization Research of Traditional Chinese Medicine", (No.2017YFC1700500). 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.

## Acknowledgments

Thanks to Professor Okamura and his team for contributing available research data.

## 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.2023.1172323/full#supplementary-material

## References

1. Schmidt AM. Highlighting diabetes mellitus: the epidemic continues. *Arterioscler Thromb Vasc Biol* (2018) 38(1):e1–8. doi: 10.1161/ATVBAHA.117.310221

2. Volpato S, Blaum C, Resnick H, Ferrucci L, Fried LP, Guralnik JM, et al. Comorbidities and impairments explaining the association between diabetes and lower extremity disability: the women's health and aging study. *Diabetes Care* (2002) 25(4):678–83. doi: 10.2337/diacare.25.4.678

3. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Bärnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. *Lancet Diabetes Endocrinol* (2017) 5(6):423–30. doi: 10.1016/S2213-8587(17)30097-9

4. Magliano DJ, Boyko EJIDF Diabetes Atlas 10th edition scientific committee. *IDF DIABETES ATLAS [Internet]. 10th ed*. Brussels: International Diabetes Federation (2021).

5. Schulze MB, Hu FB. Primary prevention of diabetes: what can be done and how much can be prevented? *Annu Rev Public Health* (2005) 26:445–67. doi: 10.1146/annurev.publhealth.26.021304.144532

6. Maggio CA, Pi-Sunyer FX. Obesity and type 2 diabetes. *Endocrinol Metab Clin North Am* (2003) 32(4):805–22,viii. doi: 10.1016/s0889-8529(03)00071-9

7. Miyazaki Y, Glass L, Triplitt C, Wajcberg E, Mandarino LJ, DeFronzo RA. Abdominal fat distribution and peripheral and hepatic insulin resistance in type 2 diabetes mellitus. *Am J Physiol Endocrinol Metab* (2002) 283(6):E1135–43. doi: 10.1152/ajpendo.0327.2001

8. Stefan N. Causes, consequences, and treatment of metabolically unhealthy fat distribution. *Lancet Diabetes Endocrinol* (2020) 8(7):616–27. doi: 10.1016/S2213-8587(20)30110-8

9. Chen P, Hou X, Hu G, Wei L, Jiao L, Wang H, et al. Abdominal subcutaneous adipose tissue: a favorable adipose depot for diabetes? *Cardiovasc Diabetol* (2018) 17(1):93. doi: 10.1186/s12933-018-0734-8

10. Shuster A, Patlas M, Pinthus JH, Mourtzakis M. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. *Br J Radiol* (2012) 85(1009):1–10. doi: 10.1259/bjr/38447238

11. Schreiner PJ, Terry JG, Evans GW, Hinson WH, Crouse JR3rd, Heiss G. Sex-specific associations of magnetic resonance imaging-derived intra-abdominal and subcutaneous fat areas with conventional anthropometric indices. the atherosclerosis risk in communities study. *Am J Epidemiol* (1996) 144(4):335–45. doi: 10.1093/oxfordjournals.aje.a008934

12. Bello-Chavolla OY, Antonio-Villa NE, Vargas-Vázquez A, Viveros-Ruiz TL, Almeda-Valdes P, Gomez-Velasco D, et al. Metabolic score for visceral fat (METS-VF), a novel estimator of intra-abdominal fat content and cardio-metabolic health. *Clin Nutr* (2020) 39(5):1613–21. doi: 10.1016/j.clnu.2019.07.012

13. Feng L, Chen T, Wang X, Xiong C, Chen J, Wu S, et al. Metabolism score for visceral fat (METS-VF): a new predictive surrogate for CKD risk. *Diabetes Metab Syndr Obes* (2022) 15:2249–58. doi: 10.2147/DMSO.S370222

14. Yu P, Meng X, Kan R, Wang Z, Yu X. Association between metabolic scores for visceral fat and chronic kidney disease: a cross-sectional study. *Front Endocrinol (Lausanne)* (2022) 13:1052736. doi: 10.3389/fendo.2022.1052736

15. Feng Y, Yang X, Li Y, Wu Y, Han M, Qie R, et al. Metabolic score for visceral fat: a reliable indicator of visceral obesity for predicting risk for hypertension. *Nutrition* (2022) 93:111443. doi: 10.1016/j.nut.2021.111443

16. Liu XZ, Chen DS, Xu X, Li HH, Liu LY, Zhou L, et al. Longitudinal associations between metabolic score for visceral fat and hyperuricemia in non-obese adults. *Nutr Metab Cardiovasc Dis* (2020) 30(10):1751–7. doi: 10.1016/j.numecd.2020.06.001

17. Feng Y, Yang X, Li Y, Wu Y, Han M, Qie R, et al. Metabolic score for visceral fat: a novel predictor for the risk of type 2 diabetes mellitus. *Br J Nutr* (2022) 128(6):1029–36. doi: 10.1017/S0007114521004116

18. Okamura T, Hashimoto Y, Hamaguchi M, Obora A, Kojima T, Fukui M. Ectopic fat obesity presents the greatest risk for incident type 2 diabetes: a population-based longitudinal study. *Int J Obes (Lond)* (2019) 43(1):139–48. doi: 10.1038/s41366-018-0076-3

19. Okamura T, Hashimoto Y, Hamaguchi M, Obora A, Kojima T, Fukui M. Data from: ectopic fat obesity presents the greatest risk for incident type 2 diabetes: a population-based longitudinal study. *Dryad Dataset* (2019) 17(1):53. doi: 10.5061/dryad.8q0p192

20. Hashimoto Y, Hamaguchi M, Kojima T, Ohshima Y, Ohbora A, Kato T, et al. Modest alcohol consumption reduces the incidence of fatty liver in men: a population-based large-scale cohort study. *J Gastroenterol Hepatol* (2015) 30(3):546–52. doi: 10.1111/jgh.12786

21. Hamaguchi M, Kojima T, Itoh Y, Harano Y, Fujii K, Nakajima T, et al. The severity of ultrasonographic findings in nonalcoholic fatty liver disease reflects the metabolic syndrome and visceral fat accumulation. *Am J Gastroenterol* (2007) 102(12):2708–15. doi: 10.1111/j.1572-0241.2007.01526.x

22. American Diabetes Association. Standards of medical care in diabetes–2011. *Diabetes Care* (2011) 34 Suppl 1(Suppl 1):S11–61. doi: 10.2337/dc11-S011

23. Deng H, Hu P, Li H, Zhou H, Wu X, Yuan M, et al. Novel lipid indicators and the risk of type 2 diabetes mellitus among Chinese hypertensive patients: findings from the guangzhou heart study. *Cardiovasc Diabetol* (2022) 21(1):212. doi: 10.1186/s12933-022-01660-z

24. Hardelid P, Verfuerden M, McMenamin J, Smyth RL, Gilbert R. The contribution of child, family and health service factors to respiratory syncytial virus (RSV) hospital admissions in the first 3 years of life: birth cohort study in Scotland, 2009 to 2015. *Euro Surveill* (2019) 24(1):1800046. doi: 10.2807/1560-7917.ES.2019.24.1.1800046

25. Kim JH. Multicollinearity and misleading statistical results. *Korean J Anesthesiol* (2019) 72(6):558–69. doi: 10.4097/kja.19087

26. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* (2007) 147(8):W163–94. doi: 10.7326/0003-4819-147-8-200710160-00010-w1

27. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* (1988) 44(3):837–45. doi: 10.2307/2531595

28. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA* (2017) 318(14):1377–84. doi: 10.1001/jama.2017.12126

29. Weber MB, Oza-Frank R, Staimez LR, Ali MK, Narayan KM. Type 2 diabetes in asians: prevalence, risk factors, and effectiveness of behavioral intervention at individual and population levels. *Annu Rev Nutr* (2012) 32:417–39. doi: 10.1146/annurev-nutr-071811-150630

30. Yoon KH, Lee JH, Kim JW, Cho JH, Choi YH, Ko SH, et al. Epidemic obesity and type 2 diabetes in Asia. *Lancet* (2006) 368(9548):1681–8. doi: 10.1016/S0140-6736(06)69703-1

31. Chan JC, Malik V, Jia W, Kadowaki T, Yajnik CS, Yoon KH, et al. Diabetes in Asia: epidemiology, risk factors, and pathophysiology. *JAMA* (2009) 301(20):2129–40. doi: 10.1001/jama.2009.726

32. Williams R, Periasamy M. Genetic and environmental factors contributing to visceral adiposity in Asian populations. *Endocrinol Metab (Seoul)* (2020) 35(4):681–95. doi: 10.3803/EnM.2020.772

33. Bragg F, Tang K, Guo Y, Iona A, Du H, Holmes MV, et al. Associations of general and central adiposity with incident diabetes in Chinese men and women. *Diabetes Care* (2018) 41(3):494–502. doi: 10.2337/dc17-1852

34. Wan H, Wang Y, Xiang Q, Fang S, Chen Y, Chen C, et al. Associations between abdominal obesity indices and diabetic complications: Chinese visceral adiposity index and neck circumference. *Cardiovasc Diabetol* (2020) 19(1):118. doi: 10.1186/s12933-020-01095-4

35. Granér M, Siren R, Nyman K, Lundbom J, Hakkarainen A, Pentikäinen MO, et al. Cardiac steatosis associates with visceral obesity in nondiabetic obese men. *J Clin Endocrinol Metab* (2013) 98(3):1189–97. doi: 10.1210/jc.2012-3190

36. Neeland IJ, Turer AT, Ayers CR, Powell-Wiley TM, Vega GL, Farzaneh-Far R, et al. Dysfunctional adiposity and the risk of prediabetes and type 2 diabetes in obese adults. *JAMA* (2012) 308(11):1150–9. doi: 10.1001/2012.jama.11132

37. Smith U. Abdominal obesity: a marker of ectopic fat accumulation. *J Clin Invest* (2015) 125(5):1790–2. doi: 10.1172/JCI81507

38. Szendroedi J, Yoshimura T, Phielix E, Koliaki C, Marcucci M, Zhang D, et al. Role of diacylglycerol activation of PKCθ in lipid-induced muscle insulin resistance in humans. *Proc Natl Acad Sci U.S.A.* (2014) 111(26):9597–602. doi: 10.1073/pnas.1409229111

39. Svedberg J, Strömblad G, Wirth A, Smith U, Björntorp P. Fatty acids in the portal vein of the rat regulate hepatic insulin clearance. *J Clin Invest* (1991) 88(6):2054–8. doi: 10.1172/JCI115534

40. Karlsson T, Rask-Andersen M, Pan G, Höglund J, Wadelius C, Ek WE, et al. Contribution of genetics to visceral adiposity and its relation to cardiovascular and metabolic disease. *Nat Med* (2019) 25(9):1390–5. doi: 10.1038/s41591-019-0563-7

41. Palmer BF, Clegg DJ. The sexual dimorphism of obesity. *Mol Cell Endocrinol* (2015) 402:113–9. doi: 10.1016/j.mce.2014.11.029

42. Ashwell M, Cole TJ, Dixon AK. Ratio of waist circumference to height is strong predictor of intra-abdominal fat. *BMJ* (1996) 313(7056):559–60. doi: 10.1136/bmj.313.7056.559d

43. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol* (2017) 17(1):53. doi: 10.1186/s12874-017-0332-6

44. Tada T, Kumada T, Toyoda H, Kiriyama S, Tanikawa M, Hisanaga Y, et al. Ag Predicts hepatocellular carcinoma development: an analysis using time-dependent receiver operating characteristics. *J Hepatol* (2016) 65(1):48–56. doi: 10.1016/j.jhep.2016.03.013

45. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol* (2018) 14(2):88–98. doi: 10.1038/nrendo.2017.151

46. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* (2006) 444(7121):840–6. doi: 10.1038/nature05482

frontiers | Frontiers in Endocrinology

TYPE Methods
PUBLISHED 25 August 2023
DOI 10.3389/fendo.2023.1213711



OPEN ACCESS

EDITED BY
Youxin Wang,
Capital Medical University, China

REVIEWED BY
Juexian Song,
Capital Medical University, China
Yunfeng Xi,
Inner Mongolia Center for Disease Control and Prevention, China

*CORRESPONDENCE
Houwu Gong
hwgong@hnu.edu.cn

†These authors share first authorship

RECEIVED 28 April 2023
ACCEPTED 04 August 2023
PUBLISHED 25 August 2023

CITATION
Zhang X, Xie J, You X and Gong H (2023) Risk factors and drug discovery for cognitive impairment in type 2 diabetes mellitus using artificial intelligence interpretation and graph neural networks. *Front. Endocrinol.* 14:1213711. doi: 10.3389/fendo.2023.1213711



# Risk factors and drug discovery for cognitive impairment in type 2 diabetes mellitus using artificial intelligence interpretation and graph neural networks

Xin Zhang[1†], Jiajia Xie[2†], Xiong You[3] and Houwu Gong[4,5*]

[1]Department of Pediatric, The First Affiliated Hospital of Hunan College of Traditional Chinese Medicine, Zhuzhou, China, [2]Department of Ultrasound Imaging, The First Hospital of Hunan University of Chinese Medicine, Changsha, China, [3]Center of Rehabilitation diagnosis and Treatment, Hunan Provincial Rehabilitation Hospital, Changsha, China, [4]College of Computer Science and Electronic Engineering, Hunan University, Changsha, China, [5]Military Medical Research Institute, Academy of Military Sciences, Beijing, China

**Background:** Among the 382 million diabetic patients worldwide, approximately 30% experience neuropathy, and one-fifth of these patients eventually develop diabetes cognitive impairment (CI). However, the mechanism underlying diabetes CI remains unknown, and early diagnostic methods or effective treatments are currently not available.

**Objective:** This study aimed to explore the risk factors for CI in patients with type 2 diabetes mellitus (T2DM), screen potential therapeutic drugs for T2DM-CI, and provide evidence for preventing and treating T2DM-CI.

**Methods:** This study focused on the T2DM population admitted to the First Affiliated Hospital of Hunan College of Traditional Chinese Medicine and the First Affiliated Hospital of Hunan University of Chinese Medicine. Sociodemographic data and clinical objective indicators of T2DM patients admitted from January 2018 to December 2022 were collected. Based on the Montreal Cognitive Assessment (MoCA) Scale scores, 719 patients were categorized into two groups, the T2DM-CI group with CI and the T2DM-N group with normal cognition. The survey content included demographic characteristics, laboratory serological indicators, complications, and medication information. Six machine learning algorithms were used to analyze the risk factors of T2DM-CI, and the Shapley method was used to enhance model interpretability. Furthermore, we developed a graph neural network (GNN) model to identify potential drugs associated with T2DM-CI.

**Results:** Our results showed that the T2DM-CI risk prediction model based on Catboost exhibited superior performance with an area under the receiver operating characteristic curve (AUC) of 0.95 (specificity of 93.17% and sensitivity of 78.58%). Diabetes duration, age, education level, aspartate aminotransferase (AST), drinking, and intestinal flora were identified as risk factors for T2DM-CI. The top 10 potential drugs related to T2DM-CI, including Metformin, Liraglutide, and Lixisenatide, were selected by the GNN model. Some

herbs, such as licorice and cuscutae semen, were also included. Finally, we discovered the mechanism of herbal medicine interventions in gut microbiota.

**Conclusion:** The method based on Interpreting AI and GNN can identify the risk factors and potential drugs associated with T2DM-CI.

KEYWORDS

type 2 diabetes mellitus, cognitive impairment, risk factors, drug discovery, graph neural network (GNN)

# Introduction

Cognition is the natural process whereby the brain recognizes and acquires information (1). Cognitive impairment (CI) refers to decreased cognitive processing speed and efficiency, affecting functions such as working memory, task execution, and attention (2). Memory impairment is the most common cognitive change and may progress to dementia in severe cases (3). In recent years, CI has become increasingly recognized as one of the most important cerebrovascular complications of type 2 diabetes (T2DM) (4). There is an increasing consensus suggesting that T2DM is one of the most important causes of CI (5), with reports suggesting that diabetes can lead to a 20%–70% decline in cognitive ability, and the risk of dementia is 60% higher in diabetic patients than in non-diabetic patients (6). Diabetes is the most prevalent metabolic disease worldwide, with 500 million T2DM patients globally, one-third of whom are in China (7). With the changing social structure and the global aging trend, the number of CI cases caused by T2DM is expected to increase exponentially. Studies have shown that the incidence of mild CI in T2DM patients is significantly higher than in non-diabetic patients (8, 9). Mild CI may affect daily activities, such as impaired intelligence, slow thinking speed, reduced flexibility, and lack of concentration (10). CI caused by diabetes can be classified into diabetes-related cognitive decline, mild CI (MCI), and dementia according to severity (11). Therefore, CI can be considered as an intermediate transition between diabetes and dementia, and this process is reversible. Therefore, it is urgent to identify the risk factors for T2DM-CI and prevent its occurrence and development. Research on the risk factors for T2DM-CI has gained significant momentum in recent years. However, no consensus has been reached, and the literature has been predominantly based on foreign populations. The risk factors for T2DM-CI in China have been largely underinvestigated, and the clinical and demographic data included are not comprehensive and cannot reflect the real risk factors for T2DM patients with CI. This study aims to comprehensively analyze the risk factors for T2DM-CI, focusing on demographic characteristics and relevant clinical and physical indicators, to identify T2DM patients with possible CI early, discover potential drugs, improve patient quality of life, and reduce the burden on society.

# Materials and methods

## Study design and patients selection

The study included a population of patients with type 2 diabetes mellitus (T2DM) who were admitted to the Endocrinology Department of the First Affiliated Hospital of Hunan College of Traditional Chinese Medicine and the First Affiliated Hospital of Hunan University of Chinese Medicine between January 2018 and December 2022, and who met the specified inclusion criteria. The selection of research subjects involved a rigorous screening process conducted by at least two medical professionals, who assessed the patients using cognitive scales. Based on the assessment criteria, the patients were divided into two groups: the T2DM group with normal cognition (T2DM-N group) and the T2DM group with cognitive impairment (T2DM-CI group).

## Diagnostic criteria

The diagnosis criteria for T2DM were based on the "Chinese Guidelines for the Prevention and Treatment of Type 2 Diabetes (2013 edition)" (12). According to these criteria, T2DM can be diagnosed if patients presenting with diabetes-related symptoms (such as polyphagia, polydipsia, polyuria, and unexplained weight loss) meet any of the following three conditions: (1) random blood glucose (blood glucose at any time within a day) ≥11.1mmol/L; (2) fasting blood glucose (without calorie intake in 8 h) ≥ 7.0mmol/L; (3) blood glucose value ≥ 11.1mmol/L measured 2 h after 75 g oral glucose tolerance test. For individuals without diabetes symptoms, the blood glucose is re-tested on another day to confirm the diagnosis.

The diagnostic criteria for cognitive impairment are based on the 5th edition of the "Diagnostic and Statistical Manual of Mental Disorders" (DSM-5) and the official manual of the Montreal Cognitive Assessment (MoCA) scale (13). The following three conditions must be met to diagnose cognitive impairment: (1) The Chinese version of the MoCA score is<26 points; (2) the patient, their family, or those who know the patient well provide relevant descriptions of memory decline; (3) the patient has basic

daily living abilities, with a score ≥16 on the instrumental activities of daily living scale (IADL).

## Inclusion and exclusion criteria

The inclusion criteria for the study population were as follows:

1. T2DM-N group: patients diagnosed with type 2 diabetes mellitus according to the diagnostic criteria outlined in the "China Type 2 Diabetes Prevention and Control Guidelines (2013 Edition)."
2. T2DM-CI group: patients diagnosed with both T2DM and cognitive impairment.
3. Age between 30 and 85 years.
4. T2DM disease duration of more than 1 year.
5. Patients with complete data on relevant indicators.

The exclusion criteria were as follows:

1. Patients who have experienced acute metabolic complications of diabetes, such as diabetic ketoacidosis or hyperosmolar hyperglycemic state, within the past month.
2. Patients who recently experienced diseases that may affect glucose and lipid metabolism, such as infection, trauma, stress, or surgery.
3. Patients with severe cardiovascular disease, hematological system disease, malignant tumor, or other serious primary diseases, severe liver or kidney dysfunction, or mental illness.
4. Patients who have experienced serious acute brain diseases within the past 3 months, such as acute cerebral infarction, intracranial hemorrhage, or acute meningitis.
5. Pregnant or lactating women, or those planning to become pregnant.
6. Patients who have participated in other clinical trials within the past 3 months.

## Analysis variables

This study extracted patients' personal information and laboratory examination data from the hospital information system. The analyzed variables included gender, age, body mass index (BMI), heart rate, blood pressure, duration of type 2 diabetes mellitus (T2DM), family history, smoking and drinking history, exercise habits, and more. Laboratory indicators encompassed total cholesterol (TC), triglycerides (TGs), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), very-low-density lipoprotein cholesterol (VLDL-C), homocysteine (HCY), fasting blood glucose (FBG), 2-h postprandial blood glucose (2hPBG), glycosylated hemoglobin (HbA1c), fasting plasma insulin (FINS), fasting C-peptide, creatinine (Crea), aspartate aminotransferase (AST), alanine aminotransferase (ALT), and gut microbiota. In addition, this study utilized data from various databases, including Traditional Chinese Medicine Systems Pharmacology (TCMSP) (14), Online Mendelian Inheritance in Man (OMIM) (15), Therapeutic Target Database (TTD) (16), Pharmacogenomics Knowledgebase (Pharm Gkb) (17), and Drug Bank (18), to conduct drug discovery research for T2DM-CI.

## Machine learning methods

The raw data were processed by organizing and standardizing them. Any feature with missing values exceeding 50% was removed from the dataset. For the remaining features with missing values, continuous features were imputed using the mean and categorical features using the mode. Six machine learning models were selected as candidates for analysis, which included random forest (RF), gradient boosted decision tree model (GBDT), light gradient boosting machine (LGBM), extreme gradient boosting (XGBoost), and categorical features gradient boosting (CatBoost) (19).

- Random Forest is an algorithm that utilizes multiple decision trees to train and predict samples. The output category is determined by the mode of the individual decision tree output categories. Random Forest is insensitive to missing values, capable of handling imbalanced data, and exhibits robustness to outliers.
- Gradient Boosting Decision Tree (GBDT) is a boosting ensemble algorithm based on decision trees incorporating gradient descent. The algorithm consists of multiple decision trees, and the conclusions of all trees are accumulated to provide the final answer. GBDT can handle various types of data, including continuous and discrete values, in a flexible manner. It exhibits high prediction accuracy with relatively less parameter tuning time. Moreover, it demonstrates strong robustness to outliers by utilizing robust loss functions.
- LightGBM is a decision tree algorithm based on histograms, which transforms the storage of feature values into the storage of bin values and does not require the indexing of feature values to samples. LightGBM employs an exclusive feature bundling algorithm to reduce the number of features during the training process, resulting in exceptionally fast training speeds. Therefore, it is highly suitable for classification problems involving high-dimensional datasets.
- XGBoost is a boosting algorithm based on CART trees. XGBoost uses the second-order Taylor expansion of the loss function as a surrogate function, which is then minimized to determine the optimal split point and leaf node output value of the regression tree. XGBoost offers reduced learning time and exhibits high flexibility in its approach.
- CatBoost is an algorithm that utilizes symmetric decision trees (oblivious trees) as its base learner. It incorporates a

specialized method to handle categorical features and employs ordered boosting with combined categorical features to prevent gradient estimation bias. CatBoost demonstrates exceptional performance, reduces the need for hyperparameter tuning, and exhibits strong robustness.

The characteristics of logistic regression are simple calculation and strong interpretability, which are widely used in fields such as finance, healthcare, social networks, and marketing. Random Forest is characterized by no need for feature normalization and feature selection. Random Forest is mainly used for training sets with high square error and low deviation. The characteristics of Adaboost are low generalization error rate, easy coding, and sensitivity to outliers. Adaboost is suitable for baseline classification tasks. CatBoost is particularly adept at handling category features. CatBoost is suitable for processing categorical data. The characteristic of GBDT is high prediction accuracy, suitability for low dimensional data, and ability to handle nonlinear data. GBDT is applicable to regression problems (linear and nonlinear), and it is also applicable to binary classification problems and multiclassification problems. The characteristic of XGBoost is its support for parallel computing, fast training speed, suitability for high bias, low variance training sets, and suitability for numerical vectors.

The entire dataset was randomly split into an 80% training set and a 20% testing set for model training and evaluation. Performance metrics from the validation set were utilized to compare the models and estimate their generalization ability. The Shapley method was employed to enhance the interpretability of the model, providing insights into the factors influencing T2DM-CI at a local level. Furthermore, a graph neural network model was utilized for drug discovery research on T2DM-CI, identifying potential therapeutic drugs with beneficial effects on T2DM-CI.

## Evaluation indicators

This study employed k-fold cross-validation for model validation to evaluate the robustness of the models. The training set was divided into K subsets, with one subset reserved as the validation data, while the remaining K-1 subsets were used for model training. The cross-validation process was repeated K times, with each subset being used as the validation set once, and the results were averaged or combined using other methods to obtain a single estimate. The key advantage of this method is that it repeatedly utilizes randomly generated subsets for training and validation, ensuring a comprehensive evaluation of the models. In this study, the value of k was set to 5.

The experiment adopts the area under the ROC curve (AUC) as the main evaluation indicator and specificity (Spe) and sensitivity (Sen) as secondary indicators. The higher the specificity, the higher the probability of accurate diagnosis; the higher the sensitivity, the lower the probability of missed diagnosis. The calculation formula is as follows:

$$Spe = \frac{TN}{FP + TN} \quad 1$$

$$Sen = \frac{TP}{TP + FN} \quad 2$$

where TP represents the number of true positive samples, TN represents the number of true negative samples, FP represents the number of false-positive samples, and FN represents the number of false-negative samples.

## Statistical analysis

The statistical analysis in this study was conducted using SPSS 22.0 software. Continuous data were reported as mean ± standard deviation ($\bar{x} \pm s$). Prior to analysis, normal distribution and homogeneity of variance tests were performed. If the data satisfied the assumptions of normal distribution and homogeneity of variance, t-tests or ANOVA were employed for analysis. On the other hand, if the data did not meet these assumptions, non-parametric Wilcoxon rank sum tests were utilized. The comparison of count data was assessed using a chi-square test. A p-value< 0.05 was statistically significant.

## GCNN4Micro-Dis model for discovery of potential drugs

We obtained 269 drugs, 598 diseases, and 18,416 disease–drug associations from the Comparative Toxicology Database (CTD). Then, we obtained more information from LTM-TCM, including 1,928 disease symptoms, 9,122 herb medicines, and 1,170,133 associations. In this study, the performance parameters of the ROC and AUPR curves are used as the criteria for selecting drugs based on the graph neural network model. The GCNN4Micro-Dis model evidently performed well and can help identify potential disease–drug associations. The correlation scores were calculated through the model to ensure the relevance between the selected drugs and T2DM-CI.

The model GCNN4Micro-Dis (20), previously developed by a research team, was used to predict potential drugs. The structure of GCNN4Micro-Dis is shown in Figure 1. The model consists of three main steps: (1) performing a graphic Fourier transform on the input data, (2) convolving the transformed result in the spectral domain, and (3) processing the convolution result using inverse Fourier transform.

# Results

## Demographic and clinical characteristics of study participants

This study involved 719 patients, with 255 (33.62%) diagnosed with type 2 diabetes cognitive impairment and 464 (66.38%) without cognitive impairment. A comparison of the data between patients with and without the endpoint event indicated no

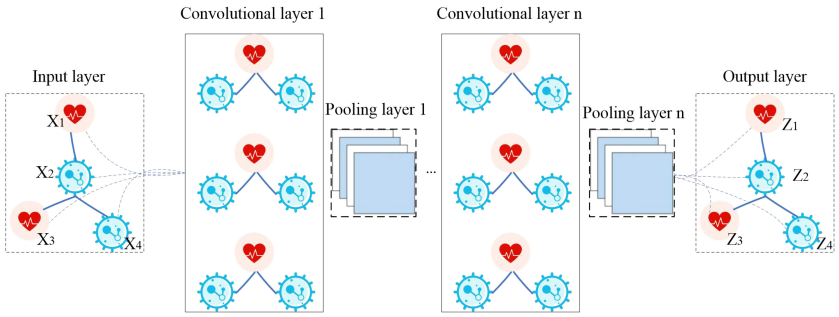



FIGURE 1
The flowchart of GCNN4Micro-Dis.

significant differences in gender, BMI, smoking, total cholesterol (TC), triglycerides (TGs), and other variables ($p>0.05$). However, significant differences were observed in age, education level, duration of diabetes, hypertension, intestinal flora, and LDL-C value ($p<0.05$). More details are provided in Table 1.

## Comparison of performance of T2DM-CI risk prediction models

In this study, the performance of six machine learning algorithms, namely, Logistic Regression, Random Forest, GBDT, Adaboost, XGBoost, and CatBoost, was compared in predicting the risk of T2DM-CI. The results (Table 2, Figure 2) showed that CatBoost exhibited higher AUC and Spe values than the other models in the validation set. The AUC value in the validation set was 95.34%, surpassing the AUC values of the other five models. Additionally, the specificity was 93.17%, outperforming the other four models. The Random Forest model achieved the highest sensitivity (78.58%). Overall, the experimental data from this study demonstrated that the CatBoost model was superior to other models in predicting the risk of T2DM-CI.

## Discovery of risk factors for T2DM-CI

To explore the risk factors influencing T2DM-CI, this study introduced an interpretive T2DM-CI prediction model based on

TABLE 1 Comparison of information among T2DM patients.

| | T2DM-CI (n=255) | T2DM-N (n=464) | $\chi^2$/t/Z | p-value |
|---|---|---|---|---|
| Gender | | | | |
| Male | 135(52.94%) | 244(52.59%) | 1.34 | 0.26 |
| Female | 120(47.06%) | 220(47.41%) | | |
| Age (year) | 64.32±8.32 | 60.12±10.85 | −3.37 | <0.01 |
| Education level | | | | |
| Below middle | 110(43.14%) | 166(35.78%) | 18.51 | <0.01 |
| Middle and above | 145(56.86%) | 298(64.22%) | | |
| BMI(kg/m$^2$) | 24.01±2.25 | 25.93±3.12 | 2.53 | 0.08 |
| Diabetes duration (year) | 13.85±8.11 | 11.36±6.49 | −0.37 | <0.01 |
| Smoke | | | | |
| No | 131(51.37%) | 221(47.63%) | 0.02 | 0.96 |
| Yes | 124(48.63%) | 243(52.37%) | | |
| Drink | | | | |
| No | 117(45.88%) | 207(44.61%) | 1.12 | 0.04 |
| Yes | 138(54.12%) | 257(55.39%) | | |
| Hypertension | | | 7.29 | <0.01 |

*(Continued)*

10.3389/fendo.2023.1213711

TABLE 1 Continued

| | T2DM-CI (n=255) | T2DM-N (n=464) | $\chi^2$/t/Z | p-value |
|---|---|---|---|---|
| Normal | 129(50.59%) | 231(49.78%) | | |
| Abnormal | 126(49.41%) | 233(50.22%) | | |
| Cerebral infarction | | | | |
| Normal | 117(45.88%) | 196(42.24%) | 8.28 | <0.01 |
| Abnormal | 138(54.12%) | 268(57.76%) | | |
| Intestinal flora | | | | |
| Normal | 129(50.59%) | 188(40.52%) | 7.67 | <0.01 |
| Abnormal | 126(49.41%) | 276(59.48%) | | |
| TC (mmol/L) | 8.91±3.01 | 8.85±3.65 | −0.53 | 0.61 |
| TG (mmol/L) | 11.79±3.04 | 11.37±3.86 | −0.96 | 0.38 |
| LDL-C (mmol/L) | 11.15±9.93 | 16.21±10.98 | −1.88 | 0.04 |
| HCY (μmol/L) | 13.72±10.41 | 14.73±11.28 | −0.83 | 0.45 |
| FBG (mmol/L) | 9.11±3.44 | 9.01±3.81 | −0.57 | 0.59 |
| 2hPBG (mmol/L) | 11.85±3.53 | 11.48±3.77 | −0.89 | 0.34 |
| HbA1c (%) | 9.44±2.04 | 9.13±1.99 | −1.89 | 0.05 |
| FINS (μIU/ml) | 12.55±10.21 | 16.94±11.17 | −1.76 | 0.03 |
| HoMA-IR | 4.64±4.11 | 7.94±4.73 | −0.37 | 0.74 |
| Crea (μmol/L) | 73.83±27.36 | 72.11±21.66 | −0.55 | 0.63 |
| AST (U/L) | 15.75±8.23 | 18.93±13.84 | −1.84 | 0.04 |
| ALT (U/L) | 15.42±7.79 | 17.91±7.26 | −0.72 | 0.47 |

CatBoost and TreeSHAP (21). From a global perspective, the importance of features contributing to T2DM-CI was ranked and presented in Figure 3. The analysis revealed that T2DM-CI might be associated with factors such as diabetes duration, age, education level, AST, drinking habits, and intestinal flora.

## Discovery of potential drugs related to T2DM-CI

In the previous section, intestinal flora was identified as a risk factor for T2DM-CI. In this section, we analyzed the relationship

TABLE 2 Comparison of results among different machine learning algorithms.

| DataSet | Algorithms | AUC (%) | Specificity(%) | Sensitivity(%) |
|---|---|---|---|---|
| Training set | Logistic regression | 96.94 | 93.76 | 80.35 |
| | Random Forest | 99.99 | 99.99 | 99.99 |
| | GBDT | 99.46 | 98.83 | 90.32 |
| | Adaboost | 98.35 | 96.18 | 88.86 |
| | XGBoost | 99.99 | 99.99 | 99.99 |
| | CatBoost | 99.81 | 98.12 | 94.51 |
| Test set | Logistic regression | 95.14 | 90.92 | 77.27 |
| | Random Forest | 93.24 | 91.27 | **78.58** |
| | GBDT | 93.17 | 90.23 | 72.73 |
| | Adaboost | 93.15 | 91.15 | 72.73 |
| | XGBoost | 94.28 | 91.19 | 77.27 |
| | **CatBoost** | **95.34** | **93.17** | 77.27 |

The bold values means the highest value.

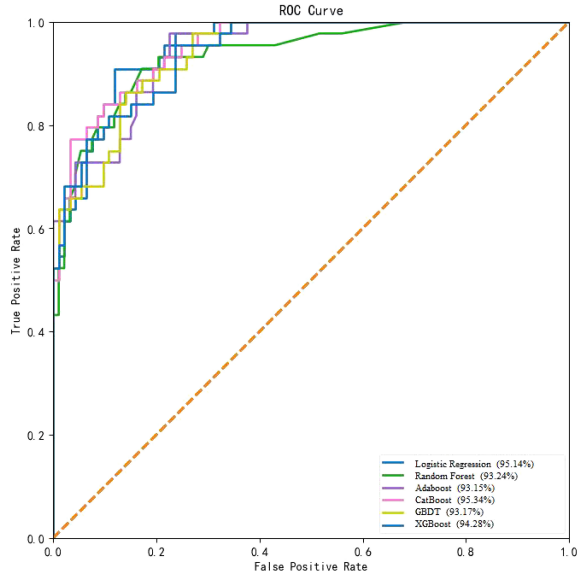



FIGURE 2
ROC curves of different machine learning algorithms on the test set.

between "T2DM-CI_intestinal flora_drug." Subsequently, we utilized the GCNN4Micro-Dis model (20) to identify potential drugs associated with T2DM-CI. Table 3 presents the top 10 drugs ranked by their association scores with T2DM-CI. Some herbs were included, such as licorice and cuscutae semen. It is worth mentioning that the results obtained have been validated in the published literature (22).

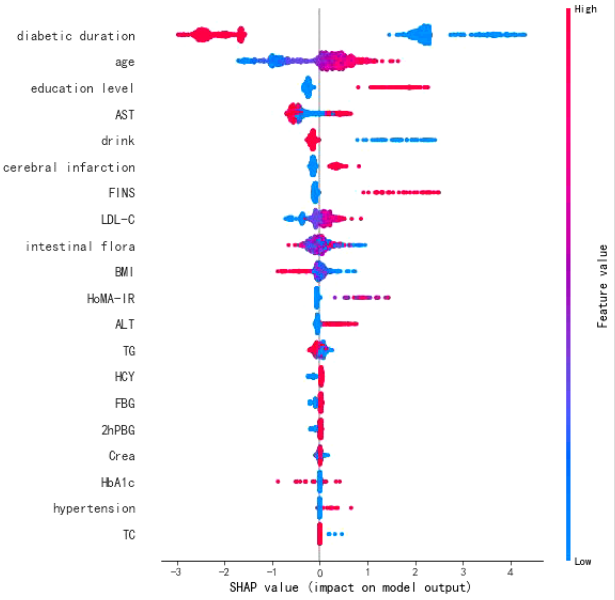



FIGURE 3
T2DM-CI features importance.

## Discussion

In this study, our approach based on artificial intelligence interpretation and graph neural networks enabled the identification of risk factors and potential drugs that impact the progression of T2DM to cognitive impairment. These findings offer valuable insights for the comprehensive treatment of T2DM and the prevention of dementia. The analysis highlighted the significance of diabetes duration, age, education level, AST, alcohol consumption, and intestinal flora as important risk factors for T2DM-CI. Importantly, the present study focused on the T2DM population and assessed relevant risk factors, enabling more accurate and convenient screening and early prevention in clinical practice. Furthermore, this study encompassed a comprehensive range of potential risk indicators. While previous research primarily concentrated on common clinical indicators, this study incorporated emerging potential risk indicators such as HoMA-IR, FINS, and intestinal flora. This expansion of the risk screening scope provides a valuable reference value for future research and enhances our understanding of the multifaceted nature of T2DM-CI.

However, it should be borne in mind that this study has some limitations. The available case data were limited, which restricted the ability to conduct a stratified analysis of certain influencing factors, and the findings may be biased to some extent. Therefore, our results can only reflect the influencing factors of cognitive impairment in the T2DM population to some extent and should be interpreted with caution. Nonetheless, the findings still provide valuable guidance for preventing and treating cognitive impairment in T2DM patients. Clinical data comprise patient visit information, yet accurately reflecting all patients' symptoms through electronic medical records can be challenging for doctors, resulting in incomplete data. Indeed, some symptoms that go unnoticed by doctors may go unrecorded, leading to missing records in hospital documentation of patient visits. Furthermore, different hospitals may have varying records for the same disease, and symptoms can vary among patients. Consequently, there is a limited availability of clinical samples for real-world data. The sample size in this study was determined based on the existing data, without prior power calculation for sample size. Consequently, the study is limited by a small sample size of clinical samples, which impacts the research quality. To enhance the robustness of the results, this study necessitates a larger sample size and a more standardized research paradigm. On the one hand, we plan to explore alternative methods to increase the sample size or utilize additional data sources from public databases to complete multicenter validation studies, such as the Pima Indians Diabetes Database. On the other hand, we plan to create a questionnaire and distribute it to third-party survey teams, such as the PowerCX Wind Chime System, which can target a sample of people to answer the questionnaire. Over the past decade, third-party survey teams have become increasingly popular and even trusted by professional research companies. With the advent of big data and the continuous improvement of multisystem network connections, favorable conditions should be established to facilitate further research into the influencing factors. This will contribute to

TABLE 3 Top 10 potential drugs related to T2DM-CI.

| Rank | Related drugs | Scores | Evidence |
|---|---|---|---|
| 1 | Metformin | 0.00061 | PMID: 31975558 |
| 2 | Liraglutide | 0.00060 | PMID: 31790314 |
| 3 | Lixisenatide | 0.00059 | PMID: 21391833 |
| 4 | Liquorice | 0.00058 | PMID: 36232291 |
| 5 | Dulaglutide | 0.00057 | PMID: 30394576 |
| 6 | 3-n-Butylphthalide | 0.00056 | Unconfirmed |
| 7 | Cuscutae Semen | 0.00056 | Unconfirmed |
| 8 | Lycii Fructus | 0.00055 | PMID: 16689001 |
| 9 | DPP-4i | 0.00054 | PMID: 30394576 |
| 10 | Rhizoma Dioscoreae | 0.00054 | PMID: 31717456 |

the generation of more optimized clinical evidence, enabling a deeper understanding of the complex interactions and variables involved in various medical conditions.

The results of this study highlight several important findings regarding the relationship between type 2 diabetes and mild cognitive impairment. First, the duration of diabetes was identified as a potential risk factor for cognitive impairment. A longer duration of diabetes (more than 20 years) was associated with a higher likelihood of cerebral vascular injury, brain atrophy, and impaired cognitive function. This can be attributed to the chronic metabolic dysfunction associated with diabetes, which leads to ischemic and hypoxic changes in brain tissue and increased inhibitory neurotransmitters (23). Additionally, age was a significant factor in the development of mild cognitive impairment in patients with type 2 diabetes. Older patients, particularly those between 60 and 75, were more susceptible to cognitive impairment. This observation is consistent with previous research, suggesting that age-related decline in dopamine neurotransmission efficiency and frontal gyrus system function contribute to the deterioration of cognitive function over time (24).

Furthermore, education level was identified as a strong determinant of cognitive impairment in individuals with type 2 diabetes. Higher education levels were associated with better cognitive function, attributed to engaging in intellectual labor, maintaining good learning habits, and keeping brain cells active. Conversely, lower education levels, often associated with more physical labor and limited brain usage, led to a decline in brain neuron reserve and decreased awareness of health management (25).

Furthermore, this study revealed that intestinal flora may be a potential risk factor for mild cognitive impairment in patients with type 2 diabetes. Intestinal flora primarily influences the host through its bacterial bodies and metabolic byproducts (26). Intestinal dysbiosis in individuals with diabetes can directly affect central function and promote other pathways that impact cognitive function. These pathways are interconnected. Intestinal flora can influence metabolic and neurological diseases, offering a novel perspective for treating T2DM-CI. The altered flora in diabetic patients plays a crucial role in their cognitive impairment, highlighting the potential of regulating intestinal flora as an effective treatment target for T2DM-CI (Figure 4).

Most traditional Chinese medicine formulas can modulate the composition of the symbiotic flora. A multicenter, randomized, open-label clinical trial demonstrated that a combination of metformin and a traditional Chinese medicine formula containing *Salvia miltiorrhiza*, *Anemarrhena asphodeloides*, *Schisandra chinensis*, *Coptis chinensis*, red yeast rice, aloe vera, bitter melon, and dried ginger could improve type 2 diabetes with hyperlipidemia by promoting the growth of beneficial flora, such as Blautia and Faecalibacterium (27). Furthermore, another Chinese medicine formula Ge-Gen-Qin-Lian decoction, has been found to enrich beneficial flora, including Faecalibacterium, in the gut, associated with its anti-diabetic effects (28). Chinese medicine exerts its regulatory effects through intricate chemical interactions in the gut, thereby maintaining a healthy gut ecosystem, controlling insulin resistance, and reducing host inflammation.

Considering further experimental validation of our results, the planned experiments and validation methods are as follows. First are the molecular and cellular experiments. *In vitro* experiments

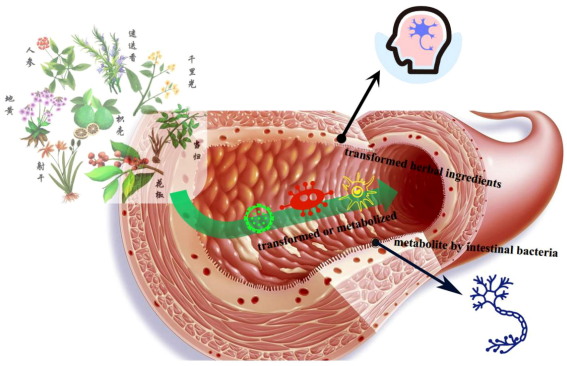



FIGURE 4
Mechanism of herbal medicine interventions in gut microbiota.

involve applying this candidate drug to the cell model of the relevant disease, observing whether it can affect the related pathological changes of this disease model. The techniques we may use include immunofluorescence staining, Western blot, qPCR, etc., to detect changes in key biomarkers. Second are animal experiments. If *in vitro* experiments prove that the drug has an effect on specific targets or pathways, then *in vivo* research is conducted, usually in animal models. At this stage, we need to observe whether the administration of the candidate drug in a specific disease model can improve symptoms or pathological changes. Third are clinical trials. If in both *in vitro* and *in vivo* experiments, the drug demonstrates the potential to alter biological processes and exhibits good safety, a clinical trial is then conducted to verify the drug's effects and safety in humans. This is a key step in our final confirmation of the drug's applicability and safety.

## 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 author.

## Author contributions

Study concept and design: HG. Acquisition of data: XZ, JX. Analysis and interpretation of data: XY. Drafting of the manuscript: XZ, JX. Critical revision of the manuscript for important intellectual content: HG. Statistical analysis: XY. Obtained funding: HG. Technical or material support: XZ, JX. Study supervision: HG. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Natural Sciences Foundation of Hunan Province (Grant No. 2021JJ30139), the National Natural Science Foundation of China (Grant No. 61773157), the China Postdoctoral Science Foundation (Grant No. 2022M711113), the 2021 Scientific Research Plan Project of Hunan Provincial Health Commission (202109021496), and the Rehabilitation Project of Hunan Disabled Persons' Federation in 2022 (Grant No. 2022XK0305).

## 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



## References

1. Desikan RS, Cabral HJ, Hess CP, Dillon WP, Glastonbury CM, Weiner MW, et al. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. *Brain* (2009) 132(Pt 8):2048–57. doi: 10.1093/brain/awp123

2. Petersen RC, Negash S, Selamawit . Mild cognitive impairment: an overview. *CNS spectrums* (2008) 13(1):45–53. doi: 10.1017/S1092852900016151

3. Mezrar S, Bendella F. A systematic review of serious games relating to cognitive impairment and dementia. *J digital Inf Manage* (2022) 2022(1):20. doi: 10.6025/jdim/2022/20/1/01-09

4. Ebady AS, Arami MA, Shafigh MH. Investigation on the relationship between diabetes mellitus type 2 and cognitive impairment. *Diabetes Res Clin Pract* (2008) 82 (3):305–9. doi: 10.1016/j.diabres.2008.08.020

5. Zhang Y, Zhang X, Zhang J, Liu C, Yuan Q, Yin X, et al. Gray matter volume abnorMalities in type 2 diabetes mellitus with and without mild cognitive impairment. *Neurosci Lett* (2014) 562:1–6. doi: 10.1016/j.neulet.2014.01.006

6. Bahorun T, Narrain D, Indelicato J, Aruoma OI, Murad F. Cognitive impairment in patients with type 2 diabetes mellitus: Perspectives and challenges. *Arch Med Biomed Res* (2014).

7. Lee Y, Smofsky A, Nykoliation P, Allain SJ, Lewis-Daly L, Schwartz J, et al. Cognitive impairment mediates workplace impairment in adults with type 2 diabetes mellitus: results from the motivaction study. *Can J Diabetes* 42(3):289–95. doi: 10.1016/j.jcjd.2017.06.013

8. Rizzo MR, Meo ID, Polito R, Auriemma MC, Gambardella A, Mauro GD, et al. Cognitive impairment and type 2 diabetes mellitus: Focus of SGLT2 inhibitors treatment. *Pharmacol Res* (2022) 176:106062. doi: 10.1016/j.phrs.2022.106062

9. Chen Y, Zhou Z, Liang Y, Tan X, Li Y, Qin C, et al. Classification of type 2 diabetes mellitus with or without cognitive impairment from healthy controls using high-order functional connectivity. *Hum Brain Mapp* (2021) 42(14):4671–84. doi: 10.1002/hbm.25575

10. Qin D, Qian H, Shouliang QI, Teng Y, Jianlin WU. ANALYSIS OF RS-FMRI IMAGES CLARIFIES BRAIN ALTERATIONS IN TYPE 2 DIABETES MELLITUS PATIENTS WITH COGNITIVE IMPAIRMENT. *J Mechanics Med Biol* (2021) 21:2140015. doi: 10.1142/S0219519421400157

11. Balashova AV, Glinkina IV, Fadeev V. The Epidemiology of Cognitive impairment in elderly patients with type 2 diabetes mellitus: prevalence and related clinical factors. *Russian J Geriatric Med* (2021) 2021(2):259. doi: 10.37586/2686-8636-2-2021-249

12. Ji L, Lu J, Weng J, Jia W, Tian H, Zhu D, et al. China type 2 diabetes treatment status survey of treatment pattern of oral drugs users. *J Diabetes* (2015) 7(2):166–73. doi: 10.1111/1753-0407.12165

13. Adrian, Wong, Lorraine, et al. Montreal cognitive assessment. *Stroke* (2015).

14. Fan Y, Sun L, Xu LJ, Peng Y, Liu HB, He CN, et al. In silico approach for anti-thrombosis drug discovery: P2Y1R structure-based TCMs screening. *Front Pharmacol* (2016) 7:531. doi: 10.3389/fphar.2016.00531

15. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* (2002) 33:D514–7. doi: 10.1093/nar/30.1.52

16. Chen X, Ji ZL, Chen YZ. TTD: therapeutic target database. *Nucleic Acids Res* (2002) 30(1):412–5. doi: 10.1093/nar/30.1.412

17. Hewett M, Oliver DE, Rubin DL, Easton KL, Stuart JM, Altman RB, et al. PharmGKB: the pharmacogenetics knowledge base. *Nucleic Acids Res* (2002) 311:179–91. doi: 10.1093/nar/30.1.163

18. Wishart DS. DrugBank: A general resource for pharmaceutical and pharmacological research. *Mol Cell Pharmacol* (2010) 2(1):D901–6. doi: 10.1093/nar/gkm958

19. Gong H, Wang M, Zhang H, Elahe MF, Jin M. An explainable AI approach for the rapid diagnosis of COVID-19 using ensemble learning algorithms. *Front Public Health* (2022) 10:874455. doi: 10.3389/fpubh.2022.874455

20. Gong H, You X, Jin M, Meng Y, Zhang H, Yang S, et al. Graph neural network and multi-data heterogeneous networks for microbe-disease prediction. *Front Microbiol* (2022) 13:1077111. doi: 10.3389/fmicb.2022.1077111

21. Zeng W, Davoodi A, Topaloglu RO. (2020). Explainable DRC hotspot prediction with random forest and SHAP tree explainer[C]// 2020 Design, in: *2020 Design, Automation & Test in Europe Conference & Exhibition*, Grenoble, France: Institute of Electrical and Electronics Engineers (IEEE).

22. Wang J, Yue RS, Wang XM, Liu RR, Zhao WW, Yang MY, et al. Effects of the method of promoting spleen and dispersing essence on intestinal flora and immune function of T2DM patients based on the theory of spleen qi dispersing essence. *China J Traditional Chin Med Pharm* (2019).

23. Ghatan S, Ahmadizar F, Li R, Ling O. Type 2 diabetes clusters indicate diabetes duration key in fracture risk. *J Endocrine Soc* (2021). doi: 10.1530/endoabs.73.PEP2.2

24. Vega T, Gil M, Lozano J. Age and sex differences in the incidence of diabetes mellitus in a population-based Spanish cohort. *J Diabetes* (2015) 7(3):411. doi: 10.1111/1753-0407.12183

25. Dunkley AJ, Tyrer F, Spong R, Gray LJ, Gillett M, Doherty Y, et al. Screening for glucose intolerance and development of a lifestyle education programme for prevention of Type 2 diabetes in a population with intellectual disabilities. *Programme Grants Appl Res* (2016) 5(11). doi: 10.3310/pgfar05110

26. Cai X, Lei W, Wang X, Liu S. Effect of high dietary fiber low glycemic index diet on intestinal flora, blood glucose and inflammatory response in T2DM patients. *Allied Academies* (2018) 2017(21):9371–5.

27. Ma Q, Li Y, Li P, Wang M, Wang J, Tang Z, et al. Research progress in the relationship between type 2 diabetes mellitus and intestinal flora. *Biomed Pharmacother* (2019) 117:109138–. doi: 10.1016/j.biopha.2019.109138

28. Li Y, Chen D, Zhang F, Lin Y, Liu J. Preventive effect of pressed degreased walnut meal extracts on T2DM rats by regulating glucolipid metabolism and modulating gut bacteria flora. *J Funct Foods* (2020) 64:103694. doi: 10.1016/j.jff.2019.103694

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 08 September 2023
DOI 10.3389/fendo.2023.1229900



OPEN ACCESS

EDITED BY
Di Liu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Elsa Vitale,
Bari Local Health Authority, Italy
Khadijeh Irandoust,
Imam Khomeini International University, Iran

*CORRESPONDENCE
Qingling Ren
yfy0047@njucm.edu.cn

RECEIVED 27 May 2023
ACCEPTED 24 August 2023
PUBLISHED 08 September 2023

CITATION
Si Y, Fei Y, Ma H, Xu Y, Ning L, Li X and Ren Q (2023) The causal association between polycystic ovary syndrome and susceptibility and severity of COVID-19: a bidirectional Mendelian randomization study using genetic data.
*Front. Endocrinol.* 14:1229900.
doi: 10.3389/fendo.2023.1229900



# The causal association between polycystic ovary syndrome and susceptibility and severity of COVID-19: a bidirectional Mendelian randomization study using genetic data

Yu Si, Yuye Fei, Hua Ma, Yating Xu, Li Ning, Xiu Li and Qingling Ren*

The Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China

**Introduction:** Observational studies have reported an association between polycystic ovary syndrome (PCOS) and COVID-19, but a definitive causal relationship has not been established. This study aimed to assess this association using two-way two-sample Mendelian randomization (MR).

**Methods:** A summary of PCOS characteristics was compiled using the PCOS summary statistics from the Apollo University of Cambridge Repository. COVID-19 susceptibility and severity statistics, including hospitalization and extremely severe disease, were obtained from genome-wide association studies from the COVID-19 Host Genetics Initiative. The primary analysis used the inverse variance-weighted method, supplemented by the weighted median, MR-Egger, and MR-PRESSO methods.

**Results:** The forward MR analysis showed no significant impact of PCOS on COVID-19 susceptibility, hospitalization, or severity (OR = 0.983, 1.011, 1.014; 95% CI = 0.958–1.008, 0.958–1.068, 0.934–1.101; and $p$ = 0.173, 0.68, 0.733; respectively). Similarly, reverse MR analysis found no evidence supporting COVID-19 phenotypes as risk or protective factors for PCOS (OR = 1.041, 0.995, 0.944; 95% CI = 0.657–1.649, 0.85–1.164, 0.843–1.058; and $p$ = 0.864, 0.945, 0.323; respectively). Consequently, no significant association between any COVID-19 phenotype and PCOS was established.

**Conclusion:** This MR study suggested that PCOS is not a causal risk factor for the susceptibility and severity of COVID-19. The associations identified in previous observational studies might be attributable to the presence of comorbidities in the patients.

KEYWORDS

**COVID-19, polycystic ovarian syndrome, Mendelian randomization, SARS-CoV-2, diabetes**

# 1 Introduction

COVID-19 is a systemic disease caused by the SARS-CoV-2 virus, primarily affecting the lungs. The pathophysiological mechanisms underlying COVID-19 involve the binding of SARS-CoV-2 to angiotensin-converting enzyme 2 (ACE2) on cell membranes, triggering local and systemic inflammatory reactions, oxidative stress, and tissue hypoxia (1). These processes involve multiple organs, including the lungs, spleen, liver, heart, and kidneys. While mild cases may be asymptomatic, severe cases can lead to dyspnea/hypoxemia, acute respiratory distress syndrome, septic shock, metabolic acidosis, and multiple organ dysfunction syndrome, often culminating in death (2). The COVID-19 pandemic, a serious global epidemic and a significant public health concern, had reached unprecedented levels of incidence and mortality at the time of writing (3, 4).

PCOS is one of the most common gynecological endocrine disorders affecting women of reproductive age, with a global incidence ranging from 8% to 13% (5). The main clinical features of PCOS include hyperandrogenism, anovulation, insulin resistance, hyperinsulinemia, abnormal menstruation, and reproductive disorders. Moreover, PCOS is associated with an increased risk of developing metabolic syndrome, cardiovascular and cerebrovascular diseases, tumors, and type 2 diabetes mellitus (6).

Epidemiological research has indicated that individuals with metabolic syndrome, which encompasses conditions such as type 2 diabetes mellitus, obesity, dyslipidemia, and hypertension, are more susceptible to severe clinical outcomes of COVID-19 (7–10). Although PCOS is not explicitly implicated in these findings, it shares common complications with these metabolic conditions (11). Consequently, it is hypothesized that women with PCOS may be more vulnerable to contracting COVID-19 and experiencing severe clinical symptoms. Several small observational studies have suggested a potential predisposing role of PCOS in COVID-19. These studies have shown that compared to healthy women, those with PCOS have a 28%–50% higher likelihood of SARS-CoV-2 infection, coupled with increased incidence rates of hospitalization and mortality (12). Hyperandrogenism and chronic low-grade inflammation, which are pivotal factors in the pathogenesis of PCOS, may contribute to the progression of COVID-19 infection (13). Furthermore, COVID-19 may induce pancreatic beta-cell failure and adipocyte dysfunction, resulting in insulin resistance and potentially augmenting the risk of future PCOS development (14). However, the observed association between COVID-19 infection and PCOS in observational studies remains subject to confounding factors and the reversal of causal relationships, necessitating further investigation to establish a robust causal link between them.

MR represents a novel epidemiological method that employs genetic data to explore causal relationships between exposures and outcomes (15). By leveraging the random distribution of genetic variants during meiosis, MR helps to overcome the limitations associated with confounding and reverse causality that are commonly encountered in observational studies (16). In this study, we conducted a two-way MR analysis to elucidate the potential causal relationship between PCOS and the risk of COVID-19 infection.

# 2 Methods

## 2.1 Study design

A bidirectional two-sample MR research was conducted to evaluate the causal association between PCOS and COVID-19 susceptibility and severity. The instrumental variables employed in the analysis were chosen based on three key principles: (1) a robust correlation between genetic variation and the exposure of interest; (2) a minimal correlation between genetic variation and potential confounding factors; and (3) genetic variation that does not directly influence the outcomes under investigation (17). In this study, bidirectional MR was used to assess the effects of PCOS on COVID-19 (forward MR) and the effects of COVID-19 on PCOS (reverse MR), using the genome-wide association study (GWAS) data for PCOS and COVID-19, respectively.

## 2.2 Data sources

The data utilized in this study were derived from the COVID-19 Host Genetics Initiative GWAS round 7 meta-analyses (18). The dataset comprised 112,612 European patients with COVID-19 who were enrolled for the susceptibility phenotypic study. These patients were categorized based on laboratory-confirmed SARS-CoV-2 infection, which was identified through electronic health records (identified through the International Classification of Diseases codes or physician annotations) or self-reporting. The control group consisted of 2,474,079 individuals without confirmed COVID-19 infection. To evaluate the severity of COVID-19, two separate cohorts were utilized. The first cohort compared 24,274 hospitalized patients with 2,061,529 control patients, whereas the second cohort compared 8,779 individuals with very severe COVID-19 outcomes with 1,001,875 control individuals who were not part of the case group (source: https://www.covid19hg.org/results/r7/). Further information on the phenotypic characteristics of the study participants is presented in Table 1.

A comprehensive analysis of statistical data on PCOS was conducted, utilizing data obtained from the authoritative repository available at https://www.repository.cam.ac.uk/items/3ccbf35f-5f69-4b46-b4fc-12cb67af71ae. The summary encompasses the findings of a GWAS comprising seven independent cohorts (19). The PCOS cohort consisted of 10,074 individuals diagnosed with PCOS, whereas the control group comprised 103,164 healthy female participants. To

TABLE 1 Data sources for the analysis.

| Phenotype | Source of Genetic Variants | |
|---|---|---|
| | Consortium | Participants |
| Polycystic ovarian syndrome | – | Cases: 10,074 participants diagnosed with PCOS. |
| | | Controls: 103,164 individuals not diagnosed with PCOS. |
| COVID-19 susceptibility | Susceptibility | Cases: 112,612 participants who were verified as COVID-19 through laboratory testing for SARS-CoV-2 infection, electrical health records, or self-reporting. |
| | | Controls: 2,474,079 participants who were a part of the cohorts but were not counted as cases. |
| COVID-19 severity | Hospitalized | Cases: 9,986 patients with COVID-19 were hospitalized. |
| | | Controls: 1,877,672 participants who were excluded from the analysis. |
| | Very severe disease | Cases: There were 8,779 individuals classified as very severe patients who either passed away or needed breathing assistance<br>Controls: 1,001,875 individuals who were excluded from the analysis and served. |

ensure accuracy, the statistical analysis accounted for variables such as age, age squared, and gender. For further information regarding the characteristics of the seven independent cohorts, interested readers can refer to the original article (19). Information pertaining to ethical approval and consent for data utilization was obtained by referencing the original article.

## 2.3 Selection of genetic instruments

To ensure the reliability of the analysis and minimize potential statistical biases that could originate from the original GWAS, rigorous criteria were applied during the selection of genetic instruments. Specifically, only single nucleotide polymorphisms (SNPs) with a significance threshold of $p < 5 \times 10^{-8}$ were considered for inclusion. Subsequently, only SNPs exhibiting linkage imbalance ($R^2 < 0.01$) and clustering within genomic regions separated by at least 10 Mb were retained for further analysis.

Utilizing the PhenoScanner database (http://www.phenoscanner.medschl.cam.ac.uk/), several instrumental variables related to other phenotypes that could potentially influence the outcomes were identified. These variables included rs9264740 (associated with diabetes mellitus diagnosed by a doctor, self-reported type 1 diabetes, and treatment with insulin product), rs1128175 (associated with diabetes diagnosed by a doctor; medication for cholesterol, blood pressure, or diabetes mellitus: insulin, started insulin within one year of diagnosis of diabetes mellitus, treatment with insulin product), rs550057 (associated with type 2 diabetes mellitus; medication for cholesterol, blood pressure, or diabetes mellitus: cholesterol-lowering medication), rs1498399 (associated with body mass index and weight), and rs1634761 (associated with weight). These selected SNPs for exposure to COVID-19 were found to be significantly correlated with diabetes phenotypes (e.g., diagnosed with diabetes mellitus, history of insulin therapy) and body mass phenotypes (e.g., weight, body mass index). Noteworthy studies have indicated that diabetes mellitus represents a prominent risk factor for susceptibility to and severity of SARS-CoV-2 infection. Compared to non-diabetic patients, patients with diabetes mellitus who contract SARS-CoV-2 exhibit elevated IL-6 and CRP levels. This phenomenon might be attributed to the inherent proinflammatory effect of diabetes mellitus, which may contribute to the systemic inflammatory response observed in COVID-19 (20). Moreover, patients with diabetes mellitus experience prolonged hospital stays, more severe pneumonia symptoms, and higher clinical mortality rates (21). Recent evidence further suggests that SARS-CoV-2 can directly induce acute or chronic damage to the pancreas, thereby influencing the regulation of glucose metabolism and insulin sensitivity, and even potentially inducing diabetes mellitus in individuals without prior diabetic conditions (22). Meanwhile, inflammation and the immune system in obese individuals could play a role in relation to viral diseases. Adipose tissue produces pro-inflammatory cytokines in high amounts, causing chronic low-grade inflammation and immune dysregulation (23). Consequently, these instrumental variables were excluded both before and after conducting the MR analysis.

The coefficient of determination ($R^2$) reflects the potential of genetic factors to account for variations in exposure scenarios. Additionally, F statistics ($F > 10$) were utilized to ensure the inclusion of robust instrumental variables while excluding weaker ones. Comprehensive details regarding the chosen SNP are available in the Supplementary Document.

## 2.4 MR analysis

Data on PCOS were analyzed using a two-sample MR approach. The primary method employed in this study was the inverse-variance weighted (IVW) method (24), followed by MR Egger and weighted median as secondary methods (25). It is important to note that the MR Egger method typically yields larger standard

errors of causal estimation and lower causal effect estimates compared to IVW (26). Thus, IVW was utilized to investigate the causal link between exposure and outcomes, and the findings were presented as ORs with corresponding 95% confidence intervals (CIs). Additionally, three sensitive assessments were performed, including weighted median and MR-Egger methods, to assess the impact of different assumptions and investigate potential pleiotropy-induced biases (25, 27, 28). If more than 50% of the instrumental variables are reliable, the weighted median approach determines the median of the empirical distribution of MR estimates, providing trustworthy estimates (25).

Several other sensitivity analyses were performed to evaluate and address the potential sources of bias in the study. Cochran's Q-test was used to assess heterogeneity among the instrumental variables, whereas MR-PRESSO was employed to detect and correct violations of the instrumental variable assumptions (28). MR-PRESSO is particularly useful when the horizontal pleiotropic effects account for less than 10% of the total variation (29). Leave-one-out analysis was utilized to examine the influence of individual SNPs on the MR findings. Additionally, the statistical power of the studies was assessed using the website (http://glimmer.rstudio.com/kn3in/mrnd/) (30). A flowchart illustrating the step-by-step process of the MR analysis is presented in Figure 1.

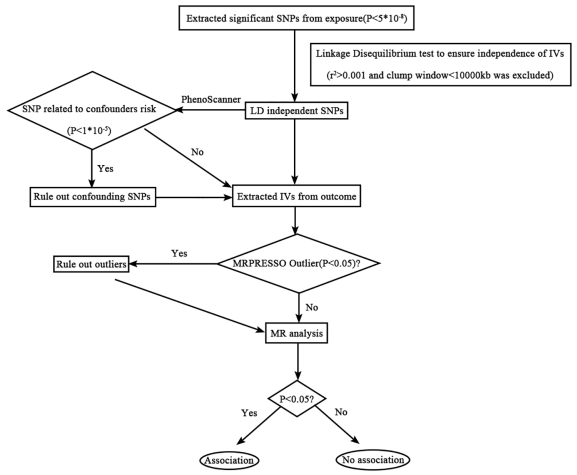



FIGURE 1
Flowchart detailing the analytical techniques and the step-by-step process of the MR analysis.

This study utilized the two-sample MR and MR-PRESSO software packages for conducting the MR analysis. All research involving statistics was conducted using R software (version 4.2.1), and STATA 12.0 and R software were combined to visualize the data.

# 3 Results

## 3.1 PCOS and COVID-19: a causal link

PCOS was neither a risk factor nor a protective factor in the susceptibility to COVID-19, hospitalization, or a severe disease phenotype. The IVW method yielded an OR of 0.983 (95% CI = 0.958–1.008, $p$ = 0.173) for susceptibility to COVID-19, an OR of 1.011 (95% CI = 0.958–1.068, $p$ = 0.68) for hospitalization, and an OR of 1.014 (95% CI = 0.934–1.101, $p$ = 0.733) for a severe disease phenotype. Furthermore, evaluations using the Q test, MR Egger intercept, and MR-PRESSO showed no notable heterogeneity, pleiotropy level, or outliers linking PCOS to the risk of COVID-19. Table 2 and Figure 2 demonstrate the comprehensive outcomes of the various MR analyses. Scatter plots, funnel plots, and leave-one-out analyses are available in the supplementary documents.

## 3.2 COVID-19 and PCOS: a causal link

Likewise, our investigation found no supportive evidence indicating that susceptibility to COVID-19, hospitalization, or the manifestation of a severe disease phenotype has a significant impact on the risk or protection against PCOS. Employing the IVW method, the ORs were estimated as 1.041 (95% CI = 0.657–1.649, $p$ = 0.864) for susceptibility to PCOS, 0.995 (95% CI = 0.85–1.164, $p$ = 0.945) for hospitalization, and 0.944 (95% CI = 0.843–1.058, $p$ = 0.323) for very severe disease phenotype due to COVID-19. None of these associations reached statistical significance. During the analysis, we identified instrumental variables such as rs9264740, rs1128175, rs550057, rs1498399, and rs1634761, which were significantly associated with diabetes mellitus and body mass which may act as confounding factors in the COVID-19 phenotype. To address this concern, these instrumental variables were excluded from the analysis, and the MR analysis was re-evaluated, leading to consistent conclusions with the initial analysis.

TABLE 2 MR-derived evaluation of the causal impact of PCOS on COVID-19.

| | Outcome | nSNP | IVW | | Weighted Median | | MR-Egger | |
|---|---|---|---|---|---|---|---|---|
| | | | OR (95%CI) | *P* | OR (95%CI) | *P* | OR (95%CI) | *P* |
| PCOS | COVID-19 susceptibility | 12 | 0.983 (0.958,1.008) | 0.173 | 0.980 (0.949,1.013) | 0.231 | 1.039 (0.918,1.175) | 0.562 |
| | COVID-19 hospitalization | 12 | 1.011 (0.958,1.068) | 0.684 | 1.012 (0.940,1.090) | 0.743 | 1.223 (0.931,1.606) | 0.178 |
| | COVID-19 severity | 12 | 1.014 (0.934,1.101) | 0.733 | 0.995 (0.892,1.110) | 0.925 | 1.174 (0.768,1.795) | 0.476 |

10.3389/fendo.2023.1229900

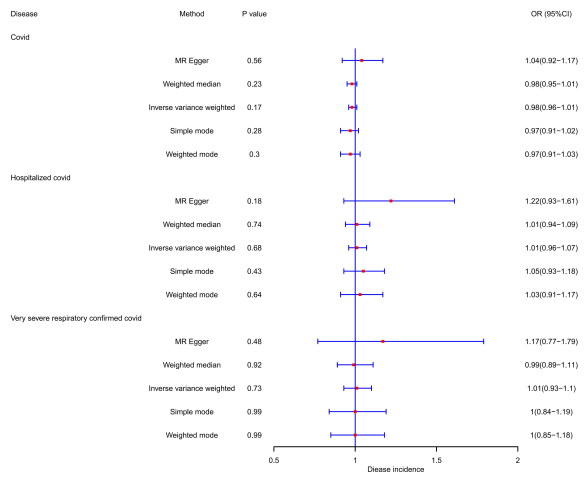



FIGURE 2
OR and 95% CI of the causal relationship between PCOS and the risk of COVID-19.

Furthermore, our study revealed that none of the COVID-19 phenotypes emerged as protective factors for PCOS ($p > 0.05$). To mitigate the potential heterogeneity in research findings, we employed the IVW random effects approach. The comprehensive results of this analysis are shown in Table 3 and Figure 3. Scatter plots, funnel plots, and leave-one-out analyses are available in the Supplementary Documents.

# 4 Discussion

COVID-19, which has been declared by the WHO as a global public health epidemic, has become one of the most concerning diseases in recent years. As of September 26, 2022, there were nearly 620 million laboratory-reported cases of COVID-19 worldwide, with more than 6.5 million deaths (https://www.worldometers.info/coronavirus/). Although COVID-19 can affect people of all ages and backgrounds, patients with pre-existing medical conditions are at an increased risk of experiencing severe outcomes and increased mortality rates (31, 32).

Our bidirectional two-sample MR study incorporated data from two distinct consortiums, particularly focusing on individuals of European ancestry. To ensure the reliability and accuracy of our genetic instruments, a meticulous screening process was conducted utilizing the PhenoScanner database.

Large sample size and a homogeneous study population are the prerequisites for the validity of the MR analysis results, and the assumptions of association, independence, and exclusion should be met to obtain valid causal inferences between the exposure and outcome variables.

In conclusion, our extensive MR analysis indicated no evidence of a causal relationship between PCOS and susceptibility to or severity of COVID-19. The results suggest that PCOS has minimal to no effect on the likelihood or severity of COVID-19 infection.

Previous observational studies have suggested a potential association between PCOS and COVID-19, which may be influenced by confounding factors and reverse causality. A closed cohort study conducted in the British population (12) aimed to investigate this relationship. The study included 21,292 women diagnosed with PCOS and randomly selected 78,310 healthy women. The incidence rate of COVID-19 was higher in women diagnosed with PCOS, at 18.1 per 1000 person-years, compared to women without PCOS, at 11.9 per 1000 person-years.

Findings from basic research indicate that androgens may contribute to the progression of COVID-19 by modifying androgen-mediated immune control and upregulating the expression of TMPRSS2, a cellular co-receptor required for SARS-CoV-2 infection (33). Immune dysfunction and a persistent inflammatory state are brought on by the endocrine–immune axis of patients with PCOS. The compensatory hyperglycemia, hyperandrogenism, and insulin resistance associated with PCOS may render individuals more susceptible to COVID-19 (13, 34, 35). Additionally, PCOS may increase susceptibility to COVID-19 through comorbidities such as obesity (36, 37). Notably, a study revealed significantly increased mRNA expression of ACE2 and TMPRSS2, key molecules involved in SARS-CoV-2 cell entry, in the livers of patients with advanced nonalcoholic fatty liver disease (NAFLD) (38).

Owing to the existence of conflicting findings among studies, not all studies have reached a consistent conclusion regarding the association between PCOS and the susceptibility to and severity of COVID-19. One such study conducted by HIPAA Limited, a subsidiary of the University of California Coronavirus Disease Research Dataset (UC CORDS), utilized health records and statistical analyses of patients undergoing COVID-19 testing at the University of California medical institutions. The results of this study

TABLE 3 MR-derived estimates of the causal influence of COVID-19 on PCOS.

| | Outcome | nSNP | IVW | | Weighted Median | | MR-Egger | |
|---|---|---|---|---|---|---|---|---|
| | | | OR (95%CI) | *P* | OR (95%CI) | *P* | OR (95%CI) | *P* |
| COVID-19 susceptibility | PCOS | 13 | 1.041 (0.657,1.649) | 0.864 | 1.271 (0.752,2.145) | 0.370 | 1.282 (0.538,3.055) | 0.586 |
| COVID-19 hospitalization | | 30 | 0.995 (0.850,1.164) | 0.945 | 1.078 (0.875,1.328) | 0.480 | 1.072 (0.807,1.424) | 0.636 |
| COVID-19 severity | | 25 | 0.944 (0.843,1.058) | 0.323 | 1.052 (0.917,1.206) | 0.471 | 0.975 (0.796,1.945) | 0.810 |

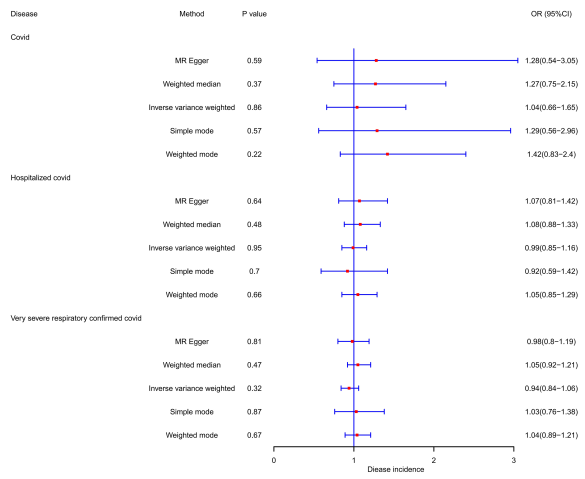



FIGURE 3
OR and 95% CI of the causal relationship between the risk of COVID-19 and PCOS.

did not provide evidence to support an increased risk of COVID-19 infection, hospitalization, or mortality among women with acne vulgaris, PCOS, or hirsutism. Consequently, establishing a causal relationship between PCOS and susceptibility to and severity of COVID-19 based solely on observational studies is challenging.

Although our study did not yield positive results, we formulated certain assumptions to understand the relationship between PCOS and COVID-19. Notably, PCOS is influenced by both genetic and environmental factors and is associated with comorbidities such as obesity, insulin resistance, diabetes mellitus, NAFLD, cardiovascular diseases, and cerebrovascular diseases. These comorbidities are established risk factors for COVID-19 susceptibility and severity, suggesting their potential role in the pathogenesis of COVID-19. The interaction between PCOS and COVID-19 involves a complex interplay of comorbidities rather than a straightforward causal relationship. The use of instrumental variables in MR analysis to control for confounding factors, such as diabetes mellitus and body mass phenotype, is a robust approach to strengthen the validity of the findings. By excluding these variables from the analysis, we aimed to minimize potential biases and ensure a more accurate assessment of the potential causal relationship between PCOS and COVID-19 infection. While the study did not establish a direct causal relationship between PCOS and COVID-19 infection and severity, it highlights the importance of considering the broader context of comorbidities and underlying risk factors that may influence disease outcomes.

MR analysis aims to minimize the impact of confounding variables to the greatest extent possible. Furthermore, we employed an extensive collection of independent datasets to secure genetic instruments, thereby diminishing the probability of bias introduced by a limited number of cases. However, given that our study exclusively focuses on individuals of European ancestry, there are inherent limitations to our conclusions, highlighting the importance of expanding the sample study to other populations.

Although MR studies are valuable tools for investigating causal relationships between exposures and outcomes using genetic variants as instrumental variables, they cannot account for other non-genetic factors, such as lifestyle and environmental factors, which may also play a role in the observed association. Therefore, this conclusion should be interpreted with caution, and further research is needed to fully understand the causal relationship between PCOS and COVID-19.

## 5 Conclusion

In conclusion, MR analysis does not provide evidence supporting PCOS as a causal risk factor influencing the susceptibility or severity of COVID-19. The previously observed correlation between PCOS and COVID-19 may be attributed to the influence of comorbidity factors. These comorbidities, such as obesity, insulin resistance, diabetes, and other cardiovascular and metabolic conditions, rather than PCOS itself, could be contributing to the association observed in these 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 author.

## Author contributions

The study was designed by YS and YF. Data analysis was conducted by YS, XL, and LN. YS, HM, and YX contributed significantly to the inception of the project and the manuscript. QR provided revisions and approved the paper. YS and YF made equal contributions to this article. All authors contributed to the article and approved the submitted version.

## Funding

This study was facilitated by funding from the National Natural Science Foundation of China (General Program, 8207150903), the Chinese Clinical Medicine Innovation Center of Obstetrics, Gynecology, and Reproduction in Jiangsu Province (ZX202102), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province, China (SJCX22_0770). None of these funding bodies partook in the design of the research, data collection and analysis, decision to publish, or manuscript preparation.

## Acknowledgments

We are grateful to the COVID19-hg Consortium and the Apollo University of Cambridge Repository for providing valuable summary statistics related to PCOS. We thank Bullet Edits Limited for the linguistic editing and proofreading of the manuscript.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1229900/full#supplementary-material

## References

1. Gusev E, Sarapultsev A, Solomatina L, Chereshnev V. SARS-coV-2-specific immune response and the pathogenesis of COVID-19. *IJMS* (2022) 23(3):1716. doi: 10.3390/ijms23031716

2. Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir Med* (2020) 8(4):420–2. doi: 10.1016/S2213-2600(20)30076-X

3. Chow EJ, Uyeki TM, Chu HY. The effects of the COVID-19 pandemic on community respiratory virus activity. *Nat Rev Microbiol* (2023) 21(3):195–210. doi: 10.1038/s41579-022-00807-9

4. Koelle K, Martin MA, Antia R, Lopman B, Dean NE. The changing epidemiology of SARS-CoV-2. *Science* (2022) 375(6585):1116–21. doi: 10.1126/science.abm4915

5. Teede H, Misso M, Tassone EC, Dewailly D, Ng EH, Azziz R, et al. Anti-müllerian hormone in PCOS: A review informing international guidelines. *Trends Endocrinol Metab* (2019) 30(7):467–78. doi: 10.1016/j.tem.2019.04.006

6. Chen W, Pang Y. Metabolic syndrome and PCOS: pathogenesis and the role of metabolites. *Metabolites* (2021) 11(12):869. doi: 10.3390/metabo11120869

7. Kalligeros M, Shehadeh F, Mylona EK, Benitez G, Beckwith CG, Chan PA, et al. Association of obesity with disease severity among patients with coronavirus disease 2019. *Obes (Silver Spring)* (2020) 28(7):1200–4. doi: 10.1002/oby.22859

8. Lighter J, Phillips M, Hochman S, Sterling S, Johnson D, Francois F, et al. Obesity in patients younger than 60 years is a risk factor for Covid-19 hospital admission. *Clin Infect Dis* (2020) 71(15):896–7. doi: 10.1093/cid/ciaa415

9. Palaiodimos L, Kokkinidis DG, Li W, Karamanis D, Ognibene J, Arora S, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York. *Metabolism* (2020) 108:154262. doi: 10.1016/j.metabol.2020.154262

10. Pranata R, Henrina J, Raffaello WM, Lawrensia S, Huang I. Diabetes and COVID-19: The past, the present, and the future. *Metabolism* (2021) 121:154814. doi: 10.1016/j.metabol.2021.154814

11. Joham AE, Norman RJ, Stener-Victorin E, Legro RS, Franks S, Moran LJ, et al. Polycystic ovary syndrome. *Lancet Diabetes Endocrinol* (2022) 10(9):668–80. doi: 10.1016/S2213-8587(22)00163-2

12. Subramanian A, Anand A, Adderley NJ, Okoth K, Toulis KA, Gokhale K, et al. Increased COVID-19 infections in women with polycystic ovary syndrome: a population-based study. *Eur J Endocrinol* (2021) 184(5):637–45. doi: 10.1530/EJE-20-1163

13. Morgante G, Troìa L, De Leo V. Coronavirus Disease 2019 (SARS-CoV-2) and polycystic ovarian disease: Is there a higher risk for these women? *J Steroid Biochem Mol Biol* (2021) 205:105770. doi: 10.1016/j.jsbmb.2020.105770

14. Reiterer M, Rajan M, Gómez-Banoy N, Lau JD, Gomez-Escobar LG, Ma Lunkun, et al. Hyperglycemia in acute COVID-19 is characterized by insulin resistance and adipose tissue infectivity by SARS-CoV-2. *Cell Metab* (2021) 33(11):2174–2188.e5. doi: 10.1016/j.cmet.2021.09.009

15. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. *Res Synth Methods* (2019) 10(4):486–96. doi: 10.1002/jrsm.1346

16. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* (2018) 362:k601. doi: 10.1136/bmj.k601

17. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* (2017) 26(5):2333–55. doi: 10.1177/0962280215597579

18. The COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum Genet* (2020) 28(6):715–8. doi: 10.1038/s41431-020-0636-6

19. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, et al. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PloS Genet* (2018) 14(12):e1007813. doi: 10.1371/journal.pgen.1007813

20. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res Rev* (2020) 36(7):e3319. doi: 10.1002/dmrr.3319

21. Singh AK, Khunti K. COVID-19 and diabetes. *Annu Rev Med* (2022) 73:129–47. doi: 10.1146/annurev-med-042220-011857

22. Iacobellis G. COVID-19 and diabetes: Can DPP4 inhibition play a role? *Diabetes Res Clin Pract* (2020) 162:108125. doi: 10.1016/j.diabres.2020.108125

23. Freuer D, Linseisen J, Meisinger C. Impact of body composition on COVID-19 susceptibility and severity: A two-sample multivariable Mendelian randomization study. *Metabolism* (2021) 118:154732. doi: 10.1016/j.metabol.2021.154732

24. The International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* (2011) 478(7367):103–9. doi: 10.1038/nature10405

25. 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

26. 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

27. 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

28. Ong JS, MacGregor S. Implementing MR-PRESSO and GCTA-GSMR for pleiotropy assessment in Mendelian randomization studies from a practitioner's perspective. *Genet Epidemiol* (2019) 43(6):609–16. doi: 10.1002/gepi.22207

29. 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

30. Brion MJA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* (2013) 42(5):1497–501. doi: 10.1093/ije/dyt179

31. Romdhani M, Vedasalam S, Souissi A, Fessi MS, Varma A, Taheri M, et al. Is there a diurnal variation of COVID-19 patients warranting presentation to the health centre? A chronobiological observational cross-sectional study. *Ann Med* (2022) 54(1):3060–8. doi: 10.1080/07853890.2022.2136399

32. Taheri M, Esmaeili A, Irandoust K, Mirmoezzi M, Souissi A, Laher I, et al. Mental health, eating habits and physical activity levels of elite Iranian athletes during the COVID-19 pandemic. *Science&Sports* (2023) 29. doi: 10.1016/j.scispo.2023.01.002

33. Foresta C, Rocca MS, Di Nisio A. Gender susceptibility to COVID-19: a review of the putative role of sex hormones and X chromosome. *J Endocrinol Invest* (2021) 44(5):951–6. doi: 10.1007/s40618-020-01383-6

34. Moradi F, Enjezab B, Ghadiri-Anari A. The role of androgens in COVID-19. *Diabetes Metab Syndr* (2020) 14(6):2003–6. doi: 10.1016/j.dsx.2020.10.014

35. Kyrou I, Karteris E, Robbins T, Chatha K, Drenos F, Randeva HS. Polycystic ovary syndrome (PCOS) and COVID-19: an overlooked female patient population at potentially higher risk during the COVID-19 pandemic. *BMC Med* (2020) 18:220. doi: 10.1186/s12916-020-01697-5

36. Michalakis K, Panagiotou G, Ilias I, Pazaitou-Panayiotou K. Obesity and COVID-19: A jigsaw puzzle with still missing pieces. *Clin Obes* (2021) 11(1):e12420. doi: 10.1111/cob.12420

37. Michalakis K, Ilias I. SARS-CoV-2 infection and obesity: Common inflammatory and metabolic aspects. *Diabetes Metab Syndr* (2020) 14(4):469–71. doi: 10.1016/j.dsx.2020.04.033

38. Fondevila MF, Mercado-Gómez M, Rodríguez A, Gonzalez-Rellan MJ, Iruzubieta P, Valentí V, et al. Obese patients with NASH have increased hepatic expression of SARS-CoV-2 critical entry points. *J Hepatol* (2021) 74(2):469–71. doi: 10.1016/j.jhep.2020.09.027

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 08 September 2023
DOI 10.3389/fendo.2023.1215666



OPEN ACCESS

EDITED BY
Youxin Wang,
Capital Medical University, China

REVIEWED BY
Luca Soraci,
IRCCS INRCA, Italy
Andrea Giannini,
Sapienza University of Rome, Italy

*CORRESPONDENCE
Xiaoqiang Liu
xiaoqiangliu1@163.com

†These authors have contributed equally to this work

RECEIVED 02 May 2023
ACCEPTED 24 August 2023
PUBLISHED 08 September 2023

CITATION
Liu L, Xu M, Zhou H, Hao X, Chen X and Liu X (2023) Association of serum 25-hydroxyvitamin D with urinary incontinence in elderly men: evidence based on NHANES 2007-2014. *Front. Endocrinol.* 14:1215666. doi: 10.3389/fendo.2023.1215666



# Association of serum 25-hydroxyvitamin D with urinary incontinence in elderly men: evidence based on NHANES 2007-2014

Li Liu†, Mingming Xu†, Hang Zhou†, Xuexue Hao, Xiangyu Chen and Xiaoqiang Liu*

Department of Urology, Tianjin Medical University General Hospital, Tianjin, China

**Background:** The correlation between serum 25-hydroxyvitamin D (25(OH)D) and different sub-types of urinary incontinence in elderly men continues to be uncertain. Hence, we performed this research to evaluate whether serum 25(OH)D levels are correlated with urinary incontinence among elderly men.

**Methods:** The present study incorporated the male population aged 50 years and above from four cycles of the NHANES database spanning from 2007 to 2014, for the purpose of analysis. The assessment of urinary incontinence was carried out through a correlation questionnaire, while standardized liquid chromatography-tandem mass spectrometry (LC-MS/MS) was adopted to quantify serum 25(OH)D. A weighted multi-factorial logistic regression analysis was carried out to ascertain and investigate any potential correlation that may exist between serum 25(OH)D and urinary incontinence in senior males.

**Results:** Ultimately, a sum of 4663 elderly men were involved in our analysis. The outcomes of the univariable analysis illustrated that the group with vitamin D deficiency exhibited augmented odds of all three urinary incontinence types in comparison to the vitamin D-sufficient group. After accounting for age, race, and BMI, no appreciable variations in the outcomes were noticed. However, after accounting for all covariates, only SUI (OR = 1.677; 95% confidence interval (CI) = 1.074–2.618) and MUI (OR = 1.815; 95% confidence interval (CI) = 1.010–3.260) demonstrated statistical significance.

**Conclusion:** Decreased serum 25(OH)D levels were connected with stress urinary incontinence and mixed urinary incontinence in elderly men.

KEYWORDS

**urinary incontinence, 25-hydroxyvitamin D, NHANES, elderly men, vitamin D deficiency**

10.3389/fendo.2023.1215666

# Introduction

Urinary incontinence (UI) is a prevalent disorder characterized by the uncontrollable leakage of urine (1). This disorder has a profound effect on the affected individual's standard of living and imposes a huge cost on society (2). Stress urinary incontinence (SUI), urge urinary incontinence (UUI), and mixed urinary incontinence (MUI) are the three most common forms of urinary incontinence (3). While UI is more prevalent in women, it is not limited to this gender and is also widespread in a significant percentage of men, particularly in older age groups. According to the European Association of Urology guidelines, UI affects approximately 11% of men between the ages of 60–64 years and up to 31% of men aged 85 years (1). Hence, the impact of UI on men, especially the elderly, should not be overlooked.

Vitamin D is a fat-soluble vitamin synthesized primarily by human skin after exposure to ultraviolet radiation and secondarily by food or other forms of supplementation (4). Its primary function is to maintain calcium and phosphorus metabolism and bone health in the body, with its main form in the body being 25-hydroxyvitamin D (25(OH)D) (5). A serum level of 25(OH)D below 50 nmol/L is typically considered a sign of vitamin D deficiency (6, 7). Vitamin D deficiency is a pervasive condition affecting a significant population worldwide. Studies have shown that approximately 1 billion people are impacted by vitamin D insufficiency or deficiency, particularly among the elderly population (8). Notably, there is a clear correlation between vitamin D deficiency and bone diseases (9). Furthermore, research has revealed that vitamin D insufficiency or deficiency is also linked to various conditions such as muscle weakness, tumors, depression, metabolic diseases, and cardiovascular diseases (8, 10–12).

Numerous studies have demonstrated a significant correlation between vitamin D deficiency and urinary incontinence in women (13, 14). Interestingly, a study has even suggested that vitamin D supplementation may improve urinary incontinence in premenopausal women (15). However, relatively little is known regarding the association between vitamin D deficiency and urinary incontinence in older men. Accordingly, the primary aim of this research endeavor is to explore the correlation between vitamin D deficiency and urinary incontinence in older men, using data obtained from the National Health and Nutrition Examination Survey (NHANES).

# Methods

## Study participants

The NHANES database is a publicly available repository of comprehensive health and nutrition data for the United States population. In this study, we extracted data from four NHANES cycles, spanning from 2007 to 2014, which contained consolidated serum 25(OH)D data. The study enrolled exclusively males aged 50 years and above, with exclusion criteria eliminating older men who did not provide urinary incontinence information or vitamin D data (Figure 1).

## Serum 25(OH)D levels

Standardized liquid chromatography-tandem mass spectrometry (LC-MS/MS) was utilized to measure serum 25(OH)D concentrations. It must be noted that total serum 25(OH)D levels below 50 nmol/L were classified as vitamin D deficient, while levels ranging from 50–75 nmol/L were considered vitamin D insufficient. On the other hand, levels of 75 nmol/L or higher were categorized as vitamin D sufficient, as per established guidelines (6).

## Definition of urinary incontinence

The assessment of urinary incontinence was conducted through the administration of a standardized questionnaire. In accordance with the criteria, any instance of urine leakage or loss of control during activities such as coughing, weight lifting, or exercise within the preceding 12-month period was classified as stress urinary incontinence (SUI). Urge urinary incontinence (UUI) was defined as involuntary urine leakage or loss of control as a result of the urge or pressure to urinate without the ability to reach a toilet in time. Concurrent manifestations of both SUI and UUI are categorized as mixed urinary incontinence (MUI).

## Covariates

In this study, hypertension is defined as having received a diagnosis of hypertension or having an average blood pressure of 140/90 mmHg or higher as indicated by a medical professional. Smoking status is categorized as either never, former, or current.

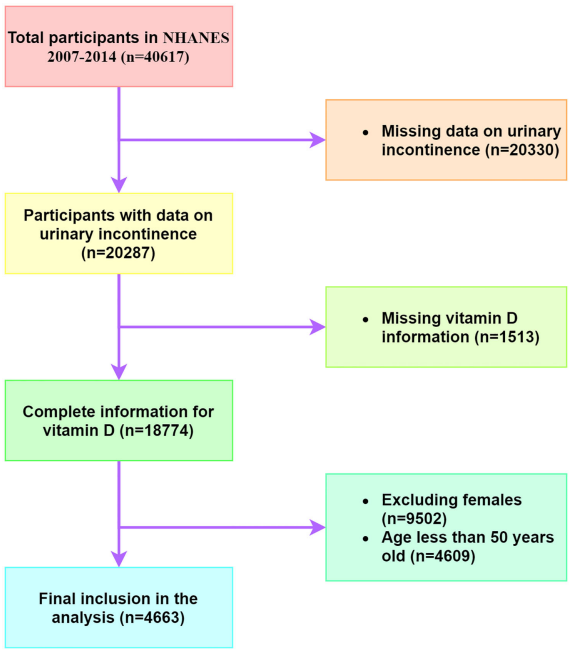



FIGURE 1
Flowchart for research.

Alcohol use is defined as having consumed at least 12 alcoholic beverages within one year. Depression status is evaluated using the PHQ-9, with a score of 10 or higher indicating depression. Additional covariates in the study include age categories (50–59, 60–69, and 70 or older), race, BMI (<25, 25–29.99, and 30 or higher), diabetes status (yes, no, or borderline), and family poverty ratio (<1.3, 1.3–3.49, and 3.5 or higher).

## Statistical analysis

Categorical variables were described using quantitative or proportional (%), while the mean ± standard error (SE) was used to express continuous variables. The analysis of the correlation between serum 25(OH)D and urinary incontinence was conducted utilizing weighted multivariate logistic regression models. Furthermore, in our research, we developed three models and determined their respective odds ratios (ORs). The first model was univariable and did not account for covariates, while model 1 adjusted for age, race, and BMI, and model 2 adjusted for age, race, BMI, alcohol use, diabetes, hypertension, family poverty rate, depression, and smoking. The level of statistical significance utilized in this study was determined to be a two-tailed p-value less than 0.05. The statistical analyses were conducted using software tools such as Stata and SPSS.

# Results

Our study included a total of 4663 older males who were selected for analysis. Out of these participants, 1263 were identified as vitamin D deficient, 1833 were classified as insufficient in vitamin D, and 1567 were found to have sufficient levels of vitamin D. A comprehensive overview of the study demographics can be found in Table 1. The prevalence of SUI, UUI, and MUI in the vitamin D deficient group was 8.3%, 26.8%, and 5.5%, respectively. The vitamin D insufficient group had a prevalence of 7.2%, 22.7%, and 4.3% for SUI, UUI, and MUI, respectively. Finally, the prevalence of SUI, UUI, and MUI in the vitamin D sufficient group had been recorded as 7.8%, 23.4%, and 4.4%, respectively.

The findings of the univariable analysis revealed that the group with vitamin D deficiency exhibited augmented odds of all three urinary incontinence types in comparison to the vitamin D

TABLE 1 A comprehensive overview of the study demographics.

| 25(OH)D | < 50 nmol/l | 50–75 nmol/l | ≥ 75 nmol/l | P Value |
|---|---|---|---|---|
| **Participants (n)** | 1263 | 1833 | 1567 | |
| **Age, Mean ± SE** | 62.68 ± 0.25 | 64.61 ± 0.22 | 66.87 ± 0.24 | <0.001 |
| **BMI, n (%)** | | | | <0.001 |
| <25 | 298(23.6) | 356(19.4) | 427(27.2) | |
| 25-29.99 | 426(33.7) | 761(41.5) | 690(44.0) | |
| ≥30 | 509(40.3) | 699(38.1) | 426(27.2) | |
| missing | 30(2.4) | 17(0.9) | 24(1.5) | |
| **Race, n (%)** | | | | <0.001 |
| Non-Hispanic White | 340(26.9) | 919(50.1) | 1064(67.9) | |
| Non-Hispanic Black | 502(39.7) | 301(16.4) | 180(11.5) | |
| Other Hispanic | 117(9.3) | 199(10.9) | 105(6.7) | |
| Other races | 304(24.1) | 414(22.6) | 218(13.9) | |
| **Alcohol use, n (%)** | | | | 0.967 |
| yes | 1035(81.9) | 1510(82.4) | 1299(82.9) | |
| no | 226(17.9) | 321(17.5) | 266(17.0) | |
| missing | 2(0.2) | 2(0.1) | 2(0.1) | |
| **Hypertension, n (%)** | | | | 0.009 |
| yes | 812(64.3) | 1108(60.4) | 1004(64.1) | |
| no | 434(34.4) | 709(38.7) | 556(35.5) | |
| missing | 17(1.3) | 16(0.9) | 7(0.4) | |
| **Smoking, n (%)** | | | | <0.001 |

*(Continued)*

10.3389/fendo.2023.1215666

TABLE 1 Continued

| 25(OH)D | < 50 nmol/l | 50–75 nmol/l | ≥ 75 nmol/l | P Value |
|---|---|---|---|---|
| never | 452(35.8) | 725(39.6) | 579(36.9) | |
| former | 449(35.6) | 808(44.1) | 758(48.4) | |
| current | 360(28.5) | 300(16.4) | 230(14.7) | |
| missing | 2(0.2) | 0(0.0) | 0(0.0) | |
| **Depression, n (%)** | | | | 0.004 |
| yes | 112(8.9) | 101(5.5) | 99(6.3) | |
| no | 1127(89.2) | 1705(93.0) | 1442(92.0) | |
| missing | 24(1.9) | 27(1.5) | 26(1.7) | |
| **Diabetes, n (%)** | | | | 0.027 |
| yes | 313(24.8) | 368(20.1) | 319(20.4) | |
| no | 909(72.0) | 1385(75.6) | 1189(75.9) | |
| borderline | 41(3.2) | 78(4.3) | 58(3.7) | |
| missing | 0(0.0) | 2(0.1) | 1(0.1) | |
| **Family poverty ratio** | | | | <0.001 |
| <1.3 | 397(31.4) | 449(24.5) | 330(21.1) | |
| 1.3-3.49 | 452(35.8) | 654(35.7) | 511(32.6) | |
| ≥3.5 | 314(24.9) | 584(31.9) | 600(38.3) | |
| missing | 100(7.9) | 146(8.0) | 126(8.0) | |

sufficient group. Upon making adjustments for age, race, and BMI, no significant changes in the results were observed. However, after accounting for all covariates, only SUI (OR = 1.677; 95% confidence interval (CI) = 1.074–2.618) and MUI (OR = 1.815; 95% confidence interval (CI) = 1.010–3.260) demonstrated statistical significance. A detailed account of the results can be referenced in Table 2 and Figure 2.

## Discussion

The present study seeks to examine the correlation between serum 25(OH)D and urinary incontinence in elderly males. This study is novel in that it is the first national study to examine the association between serum 25(OH)D and various sub-types of urinary incontinence in older men. Our findings indicate that, in both the univariable model and model 1, vitamin D deficiency was linked with increased odds of urinary incontinence, including urge, stress, and mixed. However, model 3, which adjusted for all covariates, revealed that serum 25(OH)D was linked independently with stress and mixed urinary incontinence in older men. This discovery constitutes the most significant finding of our study.

Urinary incontinence and vitamin D deficiency were prevalent disorders in the elderly population. For females, there was substantial evidences supporting a strong association between vitamin D and urinary incontinence (13). A randomized controlled trial demonstrated that vitamin D supplementation improved urinary incontinence in premenopausal women (15). However, conflicting results have been reported that moderate doses of vitamin D supplements did not diminish urinary incontinence in older women, and the underlying causes of this discrepancy require further exploration (16). In contrast, research on the relationship between vitamin D and urinary incontinence in males was relatively limited. One previous study found a correlation between vitamin D deficiency and moderate to severe urinary incontinence in adult men, though this data may be subject to bias due to testing technology at the time (17). Moreover, they failed to investigate the relationship further with older men. Instead, the present study focused on older men and investigated the association between vitamin D deficiency and urinary incontinence in older men through a large national sample.

The precise mechanisms behind the correlation between vitamin D and urinary incontinence in elderly men have yet to be fully elucidated and require further investigation. Nonetheless, multiple mechanisms may potentially be at play. Initially, it was recognized that the processes that govern urination were intricate and predominantly governed by neural, muscular, and other influences. Studies on neuromodulation evidenced that vitamin D plays a protective role by elevating antioxidant levels in neurons (18). Furthermore, vitamin D has been found to regulate the expression of several neurotransmitters, including acetylcholine and dopamine (18, 19). Also, as demonstrated in *in vitro* experiments, 1,25-dihydroxyvitamin D3 (1,25(OH)2D3) was

10.3389/fendo.2023.1215666

TABLE 2 Results of weighted logistic regression analysis.

| | Prevalences (%) | SUI | | | Prevalences (%) | UUI | | | Prevalences (%) | MUI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Univariable Model | Model 1 | Model 2 | | Univariable Model | Model 1 | Model 2 | | Univariable Model | Model 1 | Model 2 |
| 25(OH)D | | | | | | | | | | | | |
| <50 nmol/l | 8.3 | 1.517 (1.025-2.246) | 1.741 (1.132-2.679) | 1.677 (1.074-2.618) | 26.8 | 1.294 (1.020-1.642) | 1.316 (1.004-1.725) | 1.177 (0.893-1.550) | 5.5 | 1.844 (1.119-3.039) | 2.197 (1.246-3.873) | 1.815 (1.010-3.260) |
| 50-75 nmol/l | 7.2 | 1.053 (0.739-1.500) | 1.126 (0.782-1.621) | 1.131 (0.783-1.632) | 22.7 | 1.120 (0.900-1.393) | 1.169 (0.927-1.473) | 1.157 (0.916-1.462) | 4.3 | 1.263 (0.806-1.981) | 1.373 (0.857-2.201) | 1.336 (0.838-2.130) |
| ≥75 nmol/l | 7.8 | 1.0 (Reference) | 1.0 (Reference) | 1.0 (Reference) | 23.4 | 1.0 (Reference) | 1.0 (Reference) | 1.0 (Reference) | 4.4 | 1.0 (Reference) | 1.0 (Reference) | 1.0 (Reference) |

SUI, Stress urinary incontinence; UUI, Urge urinary incontinence; MUI, Mixed urinary incontinence.
Univariable Model: not adjusted; Model 1: adjusted for age, race, and BMI. Model 2: adjusted for age, race, BMI, alcohol use, diabetes, hypertension, family poverty rate, depression, and smoking.

effective in mitigating neuroinflammation through the inhibition of MAPK pathways, as well as promoting neural stem cell proliferation and enhancing their differentiation into neurons and oligodendrocytes (20, 21).

Secondly, the presence of vitamin D receptors throughout various tissues and organs, such as muscle, bladder, prostate, and urethra, has been well established (22). Recent studies have demonstrated that vitamin D has an impact on muscle strength and function by influencing various cellular processes such as cell proliferation, differentiation, protein synthesis, and myotube size (23). Experiments *in vitro* have displayed that the signaling of vitamin D and the vitamin D receptor effectively abrogated skeletal muscle atrophy through its inhibitory effect on the renin-angiotensin system (24). Furthermore, deficient levels of vitamin D may give rise to an array of consequences. Such consequences may encompass the deregulation of calcium metabolism, which can lead to anomalous contraction of the detrusor muscle. Abnormalities of the detrusor muscle are usually one of the most significant factors causing urinary incontinence, and surgery can improve the associated symptoms to a certain extent (25). And some studies have reported that a postoperative combination of Ospemifene can significantly improve patients' symptoms without increasing postoperative adverse events (26). Additionally, recent findings suggest that low levels of vitamin D may facilitate increased inflammatory cytokine activity, provoking inflammation in the bladder wall and thereby impairing urinary function (27).

In addition to its direct impact, vitamin D may also have an indirect influence on urinary incontinence through various means. Benign prostatic hyperplasia (BPH) was a frequent contributor to bladder outlet obstruction and a significant factor in urinary incontinence experienced by men in their senior years (28). Several research studies have pointed out a possible correlation between vitamin D deficiency and BPH. Vitamin D supplementation, on the other hand, has been suggested as a potential option to alleviate BPH symptoms. To elaborate, VDR agonists like elocalcitol could potentially enhance bladder contractility and obstruct prostate enlargement, thereby alleviating BPH symptoms (29). Furthermore, low serum vitamin D levels have been attributed to an increased risk of falls in old age, which could possibly increase the risk of urinary incontinence to some extent. Other than that, evidence-based medicine pointed out that individuals suffering from depression encountered more severe urine incontinence symptoms. According to one study, vitamin D supplementation may alleviate depression (30, 31). Beyond that, there may be other mechanisms, so more research was required to explore them.

Finally, this research presents various advantages. At first, it represents the first nationwide research assessing the relationship of serum 25(OH)D with distinct sub-types of urine incontinence in older males. In the second place, we incorporated data from four cycles, yielding a huge sample size. And third, serum vitamin D in the present investigation was detected *via* the LC-MS/MS method, which provided higher accuracy compared to previous assay techniques. Lastly, in the multi-factorial logistic regression analysis, we controlled for plenty of variables. However, several flaws remained. This was an observational study, and even though we controlled for numerous confounders, we couldn't rule out the

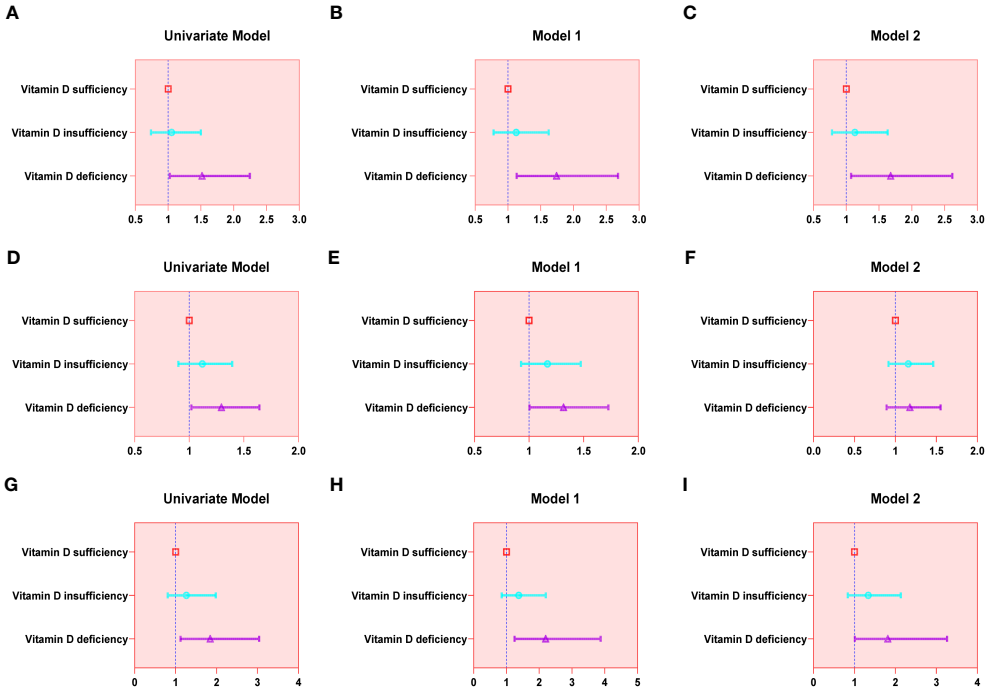



FIGURE 2
Forest plot for logistic regression analysis. **(A-C)**: SUI; **(D-F)**: UUI; **(G-I)**: MUI.

influence of any unmeasured confounding factors. Additionally, attrition bias may exist. Hence, further studies are still required to confirm the causal relationship between them and the underlying mechanisms.

## Conclusion

Decreased serum 25(OH)D levels were connected with stress urinary incontinence and mixed urinary incontinence in elderly men.

## 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 author.

## Author contributions

LL and MX conceptualized the design of this work. XH and XC were engaged in collecting and organizing the data. The data processing and image creation were executed by MX and HZ. The initial version of the paper was accomplished by MX and LL. XL was in control of reviewing and proofreading the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. The National Natural Science Funds of China (82171594) provided funding for the research.

## Acknowledgments

We are appreciative of the NHANES database for providing public availability data.

## 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.

# References


1. Gacci M, Sakalis VI, Karavitakis M, Cornu JN, Gratzke C, Herrmann TRW, et al. European association of urology guidelines on male urinary incontinence. *Eur Urol* (2022) 82:387–98. doi: 10.1016/j.eururo.2022.05.012

2. Lukacz ES, Santiago-Lastra Y, Albo ME, Brubaker L. Urinary incontinence in women: A review. *Jama* (2017) 318:1592–604. doi: 10.1001/jama.2017.12137

3. Vaughan CP, Markland AD. Urinary incontinence in women. *Ann Internal Med* (2020) 172:Itc17–itc32. doi: 10.7326/AITC202002040

4. Bikle D, Christakos S. New aspects of vitamin D metabolism and action - addressing the skin as source and target. *Nat Rev Endocrinol* (2020) 16:234–52. doi: 10.1038/s41574-019-0312-5

5. de la Guía-Galipienso F, Martínez-Ferran M, Vallecillo N, Lavie CJ, Sanchis-Gomar F, Pareja-Galeano H, et al. and cardiovascular health. *Clin Nutr (Edinburgh Scotland)* (2021) 40:2946–57. doi: 10.1016/j.clnu.2020.12.025

6. Xiao Q, Cai B, Yin A, Huo H, Lan K, Zhou G, et al. L-shaped association of serum 25-hydroxyvitamin D concentrations with cardiovascular and all-cause mortality in individuals with osteoarthritis: results from the NHANES database prospective cohort study. *BMC Med* (2022) 20:308. doi: 10.1186/s12916-022-02510-1

7. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* (2011) 96:1911–30. doi: 10.1210/jc.2011-0385

8. Sahota O. Understanding vitamin D deficiency. *Age Ageing* (2014) 43:589–91. doi: 10.1093/ageing/afu104

9. Bouillon R, Marcocci C, Carmeliet G, Bikle D, White JH, Dawson-Hughes B, et al. Skeletal and extraskeletal actions of vitamin D: current evidence and outstanding questions. *Endocrine Rev* (2019) 40:1109–51. doi: 10.1210/er.2018-00126

10. Bartley J, Vitamin D. emerging roles in infection and immunity. *Expert Rev Anti-infective Ther* (2010) 8:1359–69. doi: 10.1586/eri.10.102

11. Wilson LR, Tripkovic L, Hart KH, Lanham-New SA. Vitamin D deficiency as a public health issue: using vitamin D2 or vitamin D3 in future fortification strategies. *Proc Nutr Soc* (2017) 76:392–9. doi: 10.1017/S0029665117000349

12. Pojednic RM, Trussler EM, Navon JD, Lemire SC, Siu EC, Metallinos-Katsaras ES. Vitamin D deficiency associated with risk of prediabetes among older adults: Data from the National Health and Nutrition Examination Survey (NHANES), 2007-2012. *Diabetes/Metabol Res Rev* (2022) 38:e3499. doi: 10.1002/dmrr.3499

13. Yuan P, Wang T, Li H, Lan R, Li M, Liu J. Systematic review and meta-analysis of the association between vitamin D status and lower urinary tract symptoms. *J Urol* (2021) 205:1584–94. doi: 10.1097/JU.0000000000001441

14. Baer R, Tene L, Weintraub AY, Kalichman L. The effect of vitamin D deficiency and supplementation on urinary incontinence: scoping review. *Int Rogynecol J* (2022) 33:1083–90. doi: 10.1007/s00192-021-04963-z

15. Shahraki SK, Emadi SF, Salarfard M, Chenari Z, Tadayyonfar F, AlikaMali M. Effect of vitamin D supplementation on the severity of stress urinary incontinence in premenopausal women with vitamin D insufficiency: a randomized controlled clinical trial. *BMC Women's Health* (2022) 22:431. doi: 10.1186/s12905-022-02024-1

16. Markland AD, Vaughan C, Huang A, Kim E, Bubes VY, Tangpricha V, et al. Effect of vitamin D supplementation on urinary incontinence in older women: ancillary findings from a randomized trial. *Am J Obstetrics Gynecol* (2022) 226:535.e1–535.e12. doi: 10.1016/j.ajog.2021.10.017

17. Vaughan CP, Johnson TM 2nd, Goode PS, Redden DT, Burgio KL, Markland AD. Vitamin D and lower urinary tract symptoms among US men: results from the 2005-2006 National Health and Nutrition Examination Survey. *Urol* (2011) 78:1292–7. doi: 10.1016/j.urology.2011.07.1415

18. Harms LR, Burne TH, Eyles DW, McGrath JJ. Vitamin D and the brain. Best practice & research. *Clin Endocrinol Metab* (2011) 25:657–69. doi: 10.1016/j.beem.2011.05.009

19. Seyedi M, Gholami F, Samadi M, Djalali M, Effatpanah M, Yekaninejad MS, et al. The effect of vitamin D3 supplementation on serum BDNF, dopamine, and serotonin in children with attention-deficit/hyperactivity disorder. *CNS Neurol Disord Drug Targets* (2019) 18:496–501. doi: 10.2174/1871527318666190703103709

20. Huang YN, Ho YJ, Lai CC, Chiu CT, Wang JY. 1,25-Dihydroxyvitamin D3 attenuates endotoxin-induced production of inflammatory mediators by inhibiting MAPK activation in primary cortical neuron-glia cultures. *J Neuroinflamm* (2015) 12:147. doi: 10.1186/s12974-015-0370-0

21. Shirazi HA, Rasouli J, Ciric B, Rostami A, Zhang GX. 1,25-Dihydroxyvitamin D3 enhances neural stem cell proliferation and oligodendrocyte differentiation. *Exp Mol Pathol* (2015) 98:240–5. doi: 10.1016/j.yexmp.2015.02.004

22. Comeglio P, Chavalmane AK, Fibbi B, Filippi S, Marchetta M, Marini M, et al. Human prostatic urethra expresses vitamin D receptor and responds to vitamin D receptor ligation. *J Endocrinol Invest* (2010) 33:730–8. doi: 10.1007/BF03346679

23. Montenegro KR, Cruzat V, Carlessi R, Newsholme P. Mechanisms of vitamin D action in skeletal muscle. *Nutr Res Rev* (2019) 32:192–204. doi: 10.1017/S0954422419000064

24. Li WX, Qin XH, Poon CC, Wong MS, Feng R, Wang J, et al. Vitamin D/vitamin D receptor signaling attenuates skeletal muscle atrophy by suppressing renin-angiotensin system. *J Bone Mineral Res* (2022) 37:121–36. doi: 10.1002/jbmr.4441

25. Monti M, Fischetti M, D.I.P. A, Santangelo G, Giannini A, D'Oria O, et al. Update on surgical treatment of female stress urinary incontinence. *Minerva Obstetrics Gynecol* (2021) 73:140–4. doi: 10.23736/S2724-606X.20.04658-4

26. Schiavi MC, D'Oria O, Aleksa N, Vena F, Prata G, Di Tucci C, et al. Usefulness of Ospemifene in the treatment of urgency in menopausal patients affected by mixed urinary incontinence underwent mid-urethral slings surgery. *Gynecol Endocrinol* (2019) 35:155–9. doi: 10.1080/09513590.2018.1500534

27. Parker-Autry CY, Burgio KL, Richter HE. Vitamin D status: a review with implications for the pelvic floor. *Int Rogynecol J* (2012) 23:1517–26. doi: 10.1007/s00192-012-1710-6

28. Haghsheno MA, Mellström D, Behre CJ, Damber JE, Johansson H, Karlsson M, et al. Low 25-OH vitamin D is associated with benign prostatic hyperplasia. *J Urol* (2013) 190:608–14. doi: 10.1016/j.juro.2013.01.104

29. Adorini L, Penna G, Fibbi B, Maggi M. Vitamin D receptor agonists target static, dynamic, and inflammatory components of benign prostatic hyperplasia. *Ann New York Acad Sci* (2010) 1193:146–52. doi: 10.1111/j.1749-6632.2009.05299.x

30. Alavi NM, Khademalhoseini S, Vakili Z, Assarian F. Effect of vitamin D supplementation on depression in elderly patients: A randomized clinical trial. *Clin Nutr (Edinburgh Scotland)* (2019) 38:2065–70. doi: 10.1016/j.clnu.2018.09.011

31. Anglin RE, Samaan Z, Walter SD, McDonald SD. Vitamin D deficiency and depression in adults: systematic review and meta-analysis. *Br J Psychiatry* (2013) 202:100–7. doi: 10.1192/bjp.bp.111.106666

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 05 October 2023
DOI 10.3389/fendo.2023.1221442



OPEN ACCESS

EDITED BY
Youxin Wang,
Capital Medical University, China

REVIEWED BY
Eric Morris Bomberg,
University of Minnesota, United States
Diego Fernández Lázaro,
University of Valladolid, Spain
Worradorn Phairuang,
Kanazawa University, Japan

*CORRESPONDENCE
Xin Wan
403677@csu.edu.cn
Peng Luo
luopeng@smu.edu.cn

†These authors have contributed equally to this work

RECEIVED 12 May 2023
ACCEPTED 10 August 2023
PUBLISHED 05 October 2023

CITATION
Zhang J, Wen J, Wan X and Luo P (2023) The causal relationship between air pollution, obesity, and COVID-19 risk: a large-scale genetic correlation study. *Front. Endocrinol.* 14:1221442. doi: 10.3389/fendo.2023.1221442



# The causal relationship between air pollution, obesity, and COVID-19 risk: a large-scale genetic correlation study

Jingwei Zhang[1,2,3†], Jie Wen[1,2,3†], Xin Wan[1,2,3*] and Peng Luo[4*]

[1]Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha, China, [2]Hypothalamic Pituitary Research Centre, Xiangya Hospital, Central South University, Changsha, China, [3]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China, [4]Department of Oncology, Zhujiang Hospital, Southern Medical University, Guangzhou, China

**Objective:** Observational evidence reported that air pollution is a significant risk element for numerous health problems, such as obesity and coronavirus disease 2019 (COVID-19), but their causal relationship is currently unknown. Our objective was to probe the causal relationship between air pollution, obesity, and COVID-19 and to explore whether obesity mediates this association.

**Methods:** We obtained instrumental variables strongly correlated to air pollutants [PM2.5, nitrogen dioxide ($NO_2$) and nitrogen oxides (NOx)], 9 obesity-related traits (abdominal subcutaneous adipose tissue volume, waist-to-hip ratio, body mass index, hip circumference, waist circumference, obesity class 1-3, visceral adipose tissue volume), and COVID-19 phenotypes (susceptibility, hospitalization, severity) from public genome-wide association studies. We used clinical and genetic data from different public biological databases and performed analysis by two-sample and two-step Mendelian randomization.

**Results:** PM2.5 genetically correlated with 5 obesity-related traits, which obesity class 1 was most affected (beta = 0.38, 95% CI = 0.11 - 0.65, $p$ = 6.31E-3). $NO_2$ genetically correlated with 3 obesity-related traits, which obesity class 1 was also most affected (beta = 0.33, 95% CI = 0.055 - 0.61, $p$ = 1.90E-2). NOx genetically correlated with 7 obesity-related traits, which obesity class 3 was most affected (beta = 1.16, 95% CI = 0.42-1.90, $p$ = 2.10E-3). Almost all the obesity-related traits genetically increased the risks for COVID-19 phenotypes. Among them, body mass index, waist circumference, hip circumference, waist-to-hip ratio, and obesity class 1 and 2 mediated the effects of air pollutants on COVID-19 risks ($p$ < 0.05). However, no direct causal relationship was observed between air pollution and COVID-19.

**Conclusion:** Our study suggested that exposure to heavy air pollutants causally increased risks for obesity. Besides, obesity causally increased the risks for COVID-19 phenotypes. Attention needs to be paid to weight status for the population who suffer from heavy air pollution, as they are more likely to be susceptible and vulnerable to COVID-19.

KEYWORDS

COVID-19, air pollution, obesity, Mendelian randomization, mediation

# Introduction

First reported in late 2019, coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a pandemic affecting people's health worldwide (1). The latest epidemiological data from World Health Organization shows that COVID-19 has caused nearly 757 million infections and more than 6.85 million deaths worldwide as of Feb. 23, 2023. Over 40% of COVID-19 survivors suffered from unresolved symptoms at four months, regardless of hospitalization status (2). Over 10% of people survived with long-term impacts on multiple organ systems, known as long COVID-19 (3). Although vaccination has reduced the incidence of severe COVID-19 to some extent, no specific treatment can target SARS-CoV-2 infection until now other than hormonal drug therapy for oxygen-dependent COVID-19 patients (4). Many exposures increase the susceptibility and severity of COVID-19, such as cardiovascular and metabolic disorders, high BMI, C-reactive protein (CRP), and smoking (5). Numerous genome-wide association studies (GWASs) in healthy populations of patients have allowed us to begin identifying the genetic correlation between exposure and disease at the genetic level. Identifying and uncovering novel factors influencing COVID-19 is essential for understanding this pandemic and enhancing its treatment.

With the rapid development of socioeconomic, air pollution remains a global health threat. Air pollution contributes to many acute and chronic diseases, such as respiratory tumors, pneumonia, chronic obstructive pulmonary disease (COPD), stroke, and heart and mental health disease (6, 7). Air pollution, including particulate matter with a diameter smaller than 2.5 μm (PM2.5), nitrogen oxides (NOx), nitrogen dioxides ($NO_2$), and ozone ($O_3$), are common and highly concentrated substances in modern cities and are relevant to people's daily lives (8). Air pollution molecules entering the respiratory tract can cause respiratory tract damage through pathological mechanisms such as inflammation and oxidative stress, thereby increasing the susceptibility and severity of respiratory diseases (9). Recent observational research indicates that PM2.5 and carbon monoxide can increase the number of daily cases, cumulative cases, and cumulative deaths of COVID-19 (10). However, the causal relationship between these components (PM2.5, NOx, and $NO_2$) and COVID-19 risk (susceptibility, hospitalization, and severity) remains largely unclear.

Due to multiple factors (genetics, epigenetics, environment, socioeconomic status, etc.), obesity has become another health problem that plagues a large number of young people. It is a medical problem that increases the risk for certain illnesses, such as cardiovascular disease, metabolic disease, neurodegenerative disease, and certain tumors (11, 12). Obesity generally cannot often be prevented through just eating a healthier diet, increasing activity, and behavioral change as evidenced by the fact that most obesity prevention strategies geared towards healthy diets, increasing activity, and other behavioral changes have been ineffective or at best only minimally effective (13). Obesity often plays a very important role as a mediator in the influence of many environmental factors on various diseases. In particular, BMI, an important obesity-related trait, directly contributed to COVID-19 Susceptibility (14).

Genome-wide association study (GWAS) is a genetics research methodology used to identify genomic variants that are statistically associated with the risk of a disease or a specific trait. However, the relationship between other obesity traits and COVID-19 still needs to be further investigated, which is made possible by the increasing availability of GWAS data related to these traits. Moreover, studies suggest that prolonged environmental exposure is strongly associated with obesity and/or metabolic disease. A multicenter study found that positive association between chronic exposure to PM2.5 during working and fasting plasma glucose among asymptomatic adults (15). A meta-analysis approach indicated that PM2.5 increase obesity (OR = 1.96, 95% CI = 1.21-3.18) among adolescents in Latin American cities (16). Considering the close connection between air pollution, obesity, and COVID-19, it is vital to explore their causal relationship and mediating role based on GWAS and two-step Mendelian randomization (MR).

Mendelian randomization (MR) is a novel epidemiologic method that uses a genetic variation to infer a causal correlation between exposure and outcome based on genetic variation closely related to exposure as potentially unconstrained instrumental variables (IVs). First proposed by Katan in 1986 to disclose whether low LDL cholesterol levels increase cancer risk, MR has become increasingly popular as genetic information on health and disease has expanded with data from genome-wide association studies and genome sequencing (17). The cardinal principle of MR assumes that genetic variants are randomly allocated at conception, mimicking the randomized controlled studies and operating independently of potential confounding variables such as environmental and lifestyle factors. MR also avoids the bias from reverse causality because diseases cannot affect genotypes. It provides a way to answer questions of causality without the typical errors that affect conclusions prevalent in many traditional epidemiological methods (18, 19). Based on the fact that the prevalence of obesity and long-COVID and the threat of air pollution have not yet been fully controlled, in order to reduce the morbidity and mortality of COVID-19, to better detect and prevent the occurrence of related diseases in key populations, and to advocate the importance of environmental protection, we discussed in detail the relationship between the three. In this paper, we applied an initial MR to explore the causal role of air pollution on COVID-19 and then explored whether obesity plays an intermediary role using a two-step MR. In step one, genetic IVs robustly associated with air pollution (PM2.5, NOx, $NO_2$) were used to assess the causal relationship with obesity. In step two, genetic IVs robustly associated with obesity were used to assess the causal relationship with COVID-19 risk (susceptibility, hospitalization, severity).

# Methods

## Data sources for air pollution

All data used for analysis in our paper were obtained from publicly available GWAS datasets and therefore do not require

ethical approval or informed consent. Summary statistics of GWAS data for the participants exposed in different levels of air pollution (PM2.5, NOx, $NO_2$) were obtained from the UK Biobank (UKB) (20). The UK Biobank is a large biomedical database and research resource, an organization that collected in-depth genetic and health information on approximately 500,000 UK participants between 2006 and 2010 through questionnaires, medical tests and other methods. New data are added regularly to the database, which is accessible to all researchers worldwide. The extents of air pollution were estimated in different sites in the UK by a land use regression for annual average 2010 (21). The mean PM2.5 was 9.99 ± 1.06 micro-g/$m^3$, ranging from 8.17 - 21.31 micro-g/$m^3$, in a GWAS including 423,796 individuals and a total of 9,851,867 single-nucleotide polymorphisms (SNPs) (22). The mean $NO_2$ was 26.71 ± 7.58 micro-g/$m^3$, ranging from 12.93 - 108.49 micro-g/$m^3$, in a GWAS including 456,380 individuals and a total of 9,851,867 SNPs (23). The mean NOx was 44.11 ± 15.53 micro-g/$m^3$, ranging from 19.74 - 265.94 micro-g/$m^3$, in a GWAS including 456,380 individuals and a total of 9,851,867 SNPs (23).

## Data sources for obesity

Summary statistics of obesity were obtained from the GIANT consortium (https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files) (24, 25) and Liu et al. GWAS meta-analyses (26). The GIANT Alliance is an international collaboration of researchers from different groups, institutions, countries and research organizations. The consortium aims to identify genetic loci that regulate human size and shape (including obesity-related traits such as height, BMI, waist circumference, etc.), primarily through meta-analysis of genome-wide association data and other large-scale genetic datasets. The GWAS of the volume of abdominal subcutaneous adipose tissue (ASAT) and visceral adipose tissue (VAT) included 32,860 individuals and a total of 9,275,407 SNPs, respectively. The GWAS of body mass index (BMI) included 681,275 individuals and a total of 2,336,260 SNPs. The GWAS of hip circumference (HC) included 213,038 individuals and a total of 2,559,739 SNPs. The GWAS of obesity class 1 (OB1) included 98,697 individuals and a total of 2,380,428 SNPs. The GWAS of obesity class 2 (OB2) included 72,546 individuals and a total of 2,331,456 SNPs. The GWAS of obesity class 3 (OB3) included 50,364 individuals and a total of 2,250,779 SNPs. The GWAS of waist circumference (WC) included 232,101 individuals and a total of 2,565,408 SNPs. The GWAS of waist-to-hip ratio (WHR) included 212,244 individuals and a total of 2,560,782 SNPs.

## Data sources for COVID-19

Summary statistics of COVID-19 were obtained from the COVID-19 host genetic websites released on April 8, 2022 (round 7, GRCh38, https://www.covid19hg.org/results/r7/) (27). The GWAS of COVID-19 susceptibility included 2,597,856 cases and 14,496,978 SNPs (C2_ALL_eur_leave_23andme). The GWAS of COVID-19 hospitalization included 2,095,324 cases and 12,469,431 SNPs (B2_ALL_eur_leave_23andme). The GWAS of COVID-19 severity included 1,086,211 cases and 12,174,527 SNPs (A2_ALL_eur_leave_23andme).

## Mendelian randomization

Three principles genetic tools were followed for MR analysis: a. genetic tools were strongly correlated with corresponding exposures ($p < 5\times10^{-5}$), which could avoid the possibility of insufficient powered instrumental variables (IVs) and has been applied on previous studies (28, 29); b. genetic tools were independent of outcomes and could only influence outcome through exposure; and c. when conducting MRs between air pollutions and COVID-19 risks, the genetic tools were independent of the mediators (30). The IVs of SNPs were conjugated using the PLINK algorithm (LD < 0.001 and < 10 MB from the index variant) to select independent IVs. The F-statistic was calculated by the $(R^2/K)/[(1-R^2)\ (N-K-1)]$, where K is the number of SNP, N is the sample size, $R^2$ is the variance explained by SNPs calculated by $2*EAF*(1-EAF) * (Beta/SE)^2$. The IVs with F < 10 were excluded to retain the reliable SNPs which robustly represented the exposures. The random effects inverse variance weighting (IVW) was used as the main analysis method, which combines the Wald ratios of the causal effect of each SNP on the outcome and provides the most accurate estimates (31). Meanwhile, MR-Egger regression method and weighted median method were used as supplements to IVW. Moreover, MR-Egger intercept test, Cochran's Q test, MR-Egger intercept test and leave-one-out analysis were used to determine the presence of pleiotropy and to assess the reliability of the results.

## Mediated effects analysis

Three beta values would be gained through two-step MR, namely beta0 (initial MR of exposures on outcomes), beta1 (step one MR of exposures on mediators), and beta2 (step two MR of mediators on outcomes). The results are interpreted as follows: 1. If beta0, beta1 and beta2 are all significant, this indicates that there is a causal association from exposure to outcome and that this association may be partially mediated by the mediating variable; 2. If beta0 is not significant but both beta1 and beta2 are significant, meanwhile the quantified indirect effects are significant, this indicates that the causal association from exposure to outcome is indirect and mediated by this variable; 3. If beta0 is significant, at least one of beta1 and beta2 is insignificant, indicating that there is no mediating effect mediated by this mediating variable in the causal association from exposure to outcome (32).

The indirect effects were recognized as the effects of exposures on outcomes mediated through the causal mediators, which was quantified by the product of coefficients method (32, 33).

## Statistical analysis

Results of Mendelian analysis were presented using beta, 95% confidence interval (95%CI) and *p* values. $P < 0.05$ was considered as statistical significance. R (version 4.0.5) packages (TwoSampleMR, version 0.5.6) was applied to perform statistical analysis.

# Results

We used graphical figures to demonstrate the entire analytical process of Mendelian randomization (Figure 1). In summary, 251, 295 and 254 index SNPs were obtained to demonstrate the genetic characteristics of PM2.5, $NO_2$, and NOx, respectively (Supplementary Tables S2–4); 155, 837, 178, 96, 77, 38, 152, 165, 136 index SNPs were obtained to demonstrate the genetic characteristics of ASAT, BMI, HC, OB1, OB2, OB3, VAT, WC, and WHR, respectively (Supplementary Tables S5-13). First, we performed two sample MR to calculated the casual relationship between air pollution and obesity traits (Figure 2A and Table 1). IVW analysis indicated a positive causal relationship between PM2.5 exposure and ASAT ($p$ = 1.49E-02), BMI ($p$ = 5.73E-03), OB1 ($p$ = 6.31E-03), VAT ($p$ = 4.38E-02), WC ($p$ = 2.85E-02); a positive causal relationship between $NO_2$ exposure and HC ($p$ = 3.77E-02), OB1 ($p$ = 1.90E-02), WC ($p$ = 2.90E-02); a positive causal relationship between NOx exposure and BMI ($p$ = 4.84E-02), HC ($p$ = 1.74E-03), OB1 ($p$ = 3.60E-02), OB2 ($p$ = 1.85E-03), OB3 ($p$ = 2.10E-03), WC ($p$ = 1.61E-03), WHR ($p$ = 6.37E-03).

Second, we performed two sample MR to calculated the casual relationship between air pollution and COVID-19 (Figure 2B and Table 2). IVW analysis suggested that there is no direct causal relationship between them.

Third, we performed two sample MR to calculated the casual correlation between obesity traits and COVID-19 (Figure 2C and Table 3). IVW analysis indicated a positive causal relationship between ASAT and COVID-19 susceptibility ($p$ = 7.35E-03), COVID-19 hospitalization ($p$ = 5.88E-03). IVW analysis indicated a positive causal relationship between BMI and COVID-19 susceptibility ($p$ = 1.74E-27), COVID-19 hospitalization ($p$ = 2.46E-40), COVID-19 severity ($p$ = 1.44E-40). IVW analysis indicated a positive causal relationship between HC and COVID-19 susceptibility ($p$ = 1.95E-07), COVID-19 hospitalization ($p$ = 1.09E-09), COVID-19 severity ($p$ = 1.64E-09). We also found a positive causal relationship between OB1 and COVID-19 susceptibility ($p$ = 1.20E-05), COVID-19 hospitalization ($p$ = 7.97E-08), COVID-19 severity ($p$ = 1.75E-07). There was a positive causal relationship between OB2 and COVID-19 susceptibility ($p$ = 1.51E-04), COVID-19 hospitalization ($p$ = 5.77E-10), COVID-19 severity ($p$ = 1.20E-07).

Meanwhile, our paper revealed a positive causal relationship between VAT and COVID-19 susceptibility ($p$ = 3.33E-03), COVID-19 hospitalization ($p$ = 1.23E-04), COVID-19 severity ($p$ = 2.24E-2). WC was positively related to the risks of COVID-19 susceptibility ($p$ = 1.44E-08), COVID-19 hospitalization ($p$ = 3.23E-16), COVID-19 severity ($p$ = 5.99E-13). WHR was positively related to the risks of COVID-19 susceptibility ($p$ = 3.03E-04), COVID-19 hospitalization ($p$ = 4.37E-05), COVID-19 severity ($p$ = 3.32E-03). In addition, MR Egger and Weighted median were used as supplementary analysis methods for IVW, and detailed results are presented in Supplementary Table S14. The above results suggested that air pollution may indirectly increase the risk of COVID-19 by affecting obesity, with obesity traits playing a mediating role.

Next, we to calculate the indirect effect played by obesity traits in air pollution affecting COVID-19 (Figure 3 and Supplementary Table S16). Our results found that PM2.5 indirectly increased the risk of COVID-19 susceptibility by affecting BMI (OR = 1.01, 95% CI = 1.00-1.02, $p$ = 7.41E-03), OB1 (OR = 1.01, 95% CI = 1.00-1.02, $p$ = 2.05E-02), and WC (OR = 1.01, 95% CI = 1.00-1.02, $p$ = 4.10E-02). PM2.5 also indirectly increased the risk of COVID-19 hospitalization by affecting BMI (OR = 1.03, 95% CI = 1.01-1.06, $p$ = 6.82E-03), OB1 (OR = 1.04, 95% CI = 1.01-1.08, $p$ = 1.49E-02), and WC (OR = 1.04, 95% CI = 1.00-1.07, $p$ = 3.44E-02). Moreover, PM2.5 indirectly increased the risk of COVID-19 severity by affecting BMI (OR = 1.05, 95% CI = 1.01-1.09, $p$ = 6.82E-03), OB1 (OR = 1.06, 95% CI = 1.01-1.10, $p$ = 1.55E-02), and WC (OR = 1.05, 95% CI = 1.00-1.09, $p$ = 3.61E-02) (Figure 3A and Table 4). Moreover, $NO_2$ indirectly increased the risk of COVID-19 susceptibility by affecting OB1 (OR = 1.01, 95% CI = 1.00-1.02, $p$ = 3.87E-02) and WC (OR = 1.01, 95% CI = 1.00-1.02, $p$ = 4.16E-02). Our data also indicated that $NO_2$ indirectly increased the risk of COVID-19 hospitalization by affecting HC (OR = 1.03, 95% CI = 1.00-1.05, $p$ = 4.92E-02), OB1 (OR = 1.04, 95% CI = 1.00-1.07, $p$ = 3.16E-02), and WC (OR = 1.04, 95% CI = 1.00-1.07, $p$ = 3.49E-02) (Figure 3B and Table 4).

Furthermore, NOx indirectly increased the risk of COVID-19 susceptibility by affecting HC (OR = 1.02, 95% CI = 1.00-1.03, $p$ = 7.29E-03), OB2 (OR = 1.02, 95% CI = 1.00-1.03, $p$ = 1.62E-03), WC (OR = 1.02, 95% CI = 1.01-1.03, $p$ = 5.85E-03), and WHR (OR = 1.01, 95% CI = 1.00-1.02, $p$ = 2.95E-02). NOx indirectly increased the risk of COVID-19 hospitalization by affecting HC (OR = 1.04, 95% CI = 1.01-1.07, $p$ = 5.35E-03), OB2 (OR = 1.05, 95% CI = 1.02-1.09, $p$ = 5.41E-03), WC (OR = 1.06, 95% CI = 1.02-1.10, $p$ = 3.26E-03), and WHR (OR = 1.03, 95% CI = 1.00-1.06, $p$ = 2.33E-02). NOx indirectly increased the risk of COVID-19 severity by affecting HC (OR = 1.06, 95% CI = 1.02-1.10, $p$ = 5.45E-03), OB2 (OR = 1.07, 95% CI = 1.02-1.12, $p$ = 7.30E-03), WC (OR = 1.07, 95% CI = 1.02-1.13, $p$ = 3.86E-03), and WHR (OR = 1.03, 95% CI = 1.00-1.07, $p$ = 4.56E-02) (Figure 3C and Table 4).

Finally, to enhance the reliability of our results, we used MR-Egger-intercept test, Cochran's Q test, and leave-one-out analysis to perform sensitivity analysis on our results (Supplementary Figures S1–3 and Supplementary Table S15). The results of Cochran's Q test in IVW showed that there is basically no heterogeneity, and the MR-Egger-intercept test and leave-one-out analysis showed that our results are quite reliable. The F-statistic for the instrumental variables were all greater than 10, also indicating the reliability of the results.

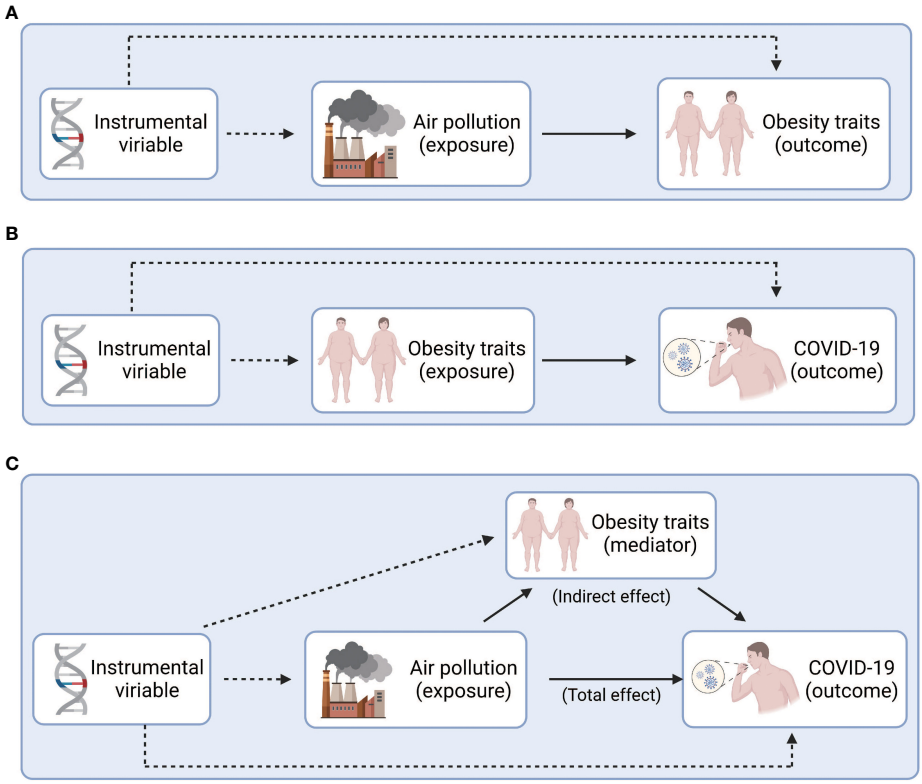



FIGURE 1
Study design overview. **(A)** Explore the causal relationship between air pollution and obesity. **(B)** Explore the causal relationship between obesity and COVID-19. **(C)** Explore the intermediary role of obesity between air pollution and COVID-19. Figure built by the Biorender.

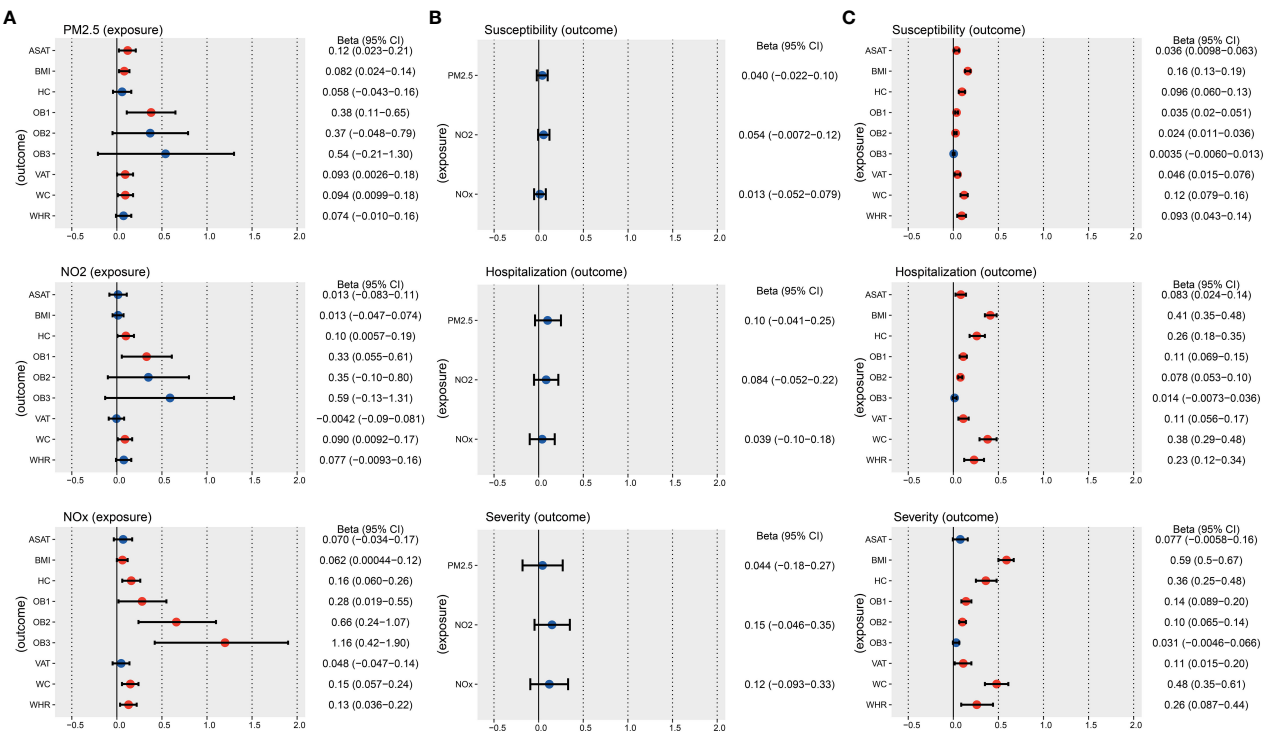



FIGURE 2
IVW results of the causal relationship between air pollution, obesity, and COVID-19 risk. **(A)** IVW results of the causal relationship between air pollution and obesity. **(B)** IVW results of the causal relationship between air pollution and COVID-19. **(C)** IVW results of the causal relationship between obesity and COVID-19. The red color means the p-value is less than 0.05.

10.3389/fendo.2023.1221442

TABLE 1 MR results of air pollution effects on obesity traits by IVW.

| Exposure | Outcome | nSNP | Beta | LCI | UCI | p |
|---|---|---|---|---|---|---|
| PM2.5 | ASAT | 240 | 0.12 | 0.023 | 0.21 | **1.49E-02** |
| | BMI | 99 | 0.082 | 0.024 | 0.14 | **5.73E-03** |
| | HC | 119 | 0.058 | -0.043 | 0.16 | 2.63E-01 |
| | OB1 | 113 | 0.38 | 0.11 | 0.65 | **6.31E-03** |
| | OB2 | 113 | 0.37 | -0.048 | 0.79 | 8.27E-02 |
| | OB3 | 113 | 0.54 | -0.21 | 1.30 | 1.59E-01 |
| | VAT | 240 | 0.093 | 0.0026 | 0.18 | **4.38E-02** |
| | WC | 118 | 0.094 | 0.0099 | 0.18 | **2.85E-02** |
| | WHR | 119 | 0.074 | -0.010 | 0.16 | 8.44E-02 |
| $NO_2$ | ASAT | 283 | 0.013 | -0.083 | 0.11 | 7.87E-01 |
| | BMI | 110 | 0.013 | -0.047 | 0.074 | 6.69E-01 |
| | HC | 133 | 0.10 | 0.0057 | 0.19 | **3.77E-02** |
| | OB1 | 127 | 0.33 | 0.055 | 0.61 | **1.90E-02** |
| | OB2 | 127 | 0.35 | -0.10 | 0.80 | 1.27E-01 |
| | OB3 | 123 | 0.59 | -0.13 | 1.31 | 1.10E-01 |
| | VAT | 283 | -0.0042 | -0.090 | 0.081 | 9.23E-01 |
| | WC | 131 | 0.090 | 0.0092 | 0.17 | **2.90E-02** |
| | WHR | 132 | 0.077 | -0.0093 | 0.16 | 8.04E-02 |
| NOx | ASAT | 240 | 0.070 | -0.034 | 0.17 | 1.87E-01 |
| | BMI | 99 | 0.062 | 0.00044 | 0.12 | **4.84E-02** |
| | HC | 119 | 0.16 | 0.060 | 0.26 | **1.74E-03** |
| | OB1 | 113 | 0.28 | 0.019 | 0.55 | **3.60E-02** |
| | OB2 | 113 | 0.66 | 0.24 | 1.07 | **1.85E-03** |
| | OB3 | 113 | 1.16 | 0.42 | 1.90 | **2.10E-03** |
| | VAT | 240 | 0.048 | -0.047 | 0.14 | 3.20E-01 |
| | WC | 118 | 0.15 | 0.057 | 0.24 | **1.61E-03** |
| | WHR | 119 | 0.13 | 0.036 | 0.22 | **6.37E-03** |

Beta = log (OR). P < 0.05 were bolded.
PM2.5, Particulate matter air pollution; $NO_2$, Nitrogen dioxide; $NO_X$, Nitrogen oxides; ASAT, abdominal subcutaneous adipose tissue; BMI, body mass index; HC, hip circumference; OB1, obesity class 1; OB2, obesity class 2; OB3, obesity class 3; VAT, visceral adipose tissue; WC, waist circumference; WHR, waist-to-hip ratio; UCI, Upper confidence interval; LCI, Lower confidence interval.

## Discussion

To date, several epidemiological studies have found that certain airborne pollutants are risk factors for obesity and COVID-19 (34), but the limitations of traditional observational study methods make it difficult to establish a causal relationship between them. In this paper, we conducted two-sample and two-step MR to assess the role of air pollution exposure on obesity traits and COVID-19 based on large-scale GWAS datasets. We found that prolonged exposure to three air pollutant molecules (PM2.5, $NO_2$, and NOx) increased the risk of obesity, suggesting a causal relationship between them. There is also a causal relationship between obesity traits and COVID-19 susceptibility, hospitalization and severity. Chronic exposure to three air pollution molecules (PM2.5, $NO_2$, and NOx) did not directly contribute to COVID-19 risk, but rather increased COVID-19 susceptibility, hospitalization and severity by affecting obesity. Given these findings, we believe that among those living in areas with heavy air pollution, maintaining a healthy weight may help prevent COVID-19 infections.

In recent years, numerous epidemiological studies have explored the relationship between long-term exposure to air pollution and obesity in different regions and populations. However, the findings of these observational studies are controversial. To date, most of the current evidence supports that air pollution can contribute to the development of obesity in children and adults, but there is also a small amount of evidence

10.3389/fendo.2023.1221442

TABLE 2 MR results of air pollution effects on COVID-19 by IVW.

| Exposure | Outcome | nSNP | Beta | LCI | UCI | p |
|---|---|---|---|---|---|---|
| PM2.5 | Susceptibility | 249 | 0.040 | -0.022 | 0.10 | 2.06E-01 |
| NO2 | | 289 | 0.054 | -0.0072 | 0.12 | 8.35E-02 |
| NOx | | 247 | 0.013 | -0.052 | 0.079 | 6.92E-01 |
| PM2.5 | Hospitalization | 247 | 0.10 | -0.041 | 0.25 | 1.59E-01 |
| NO2 | | 287 | 0.084 | -0.052 | 0.22 | 2.24E-01 |
| NOx | | 247 | 0.039 | -0.10 | 0.18 | 5.90E-01 |
| PM2.5 | Severity | 247 | 0.044 | -0.18 | 0.27 | 6.97E-01 |
| NO2 | | 288 | 0.15 | -0.046 | 0.35 | 1.31E-01 |
| NOx | | 247 | 0.12 | -0.093 | 0.33 | 2.70E-01 |

Beta = log (OR).
PM2.5, Particulate matter air pollution; NO2, Nitrogen dioxide; $NO_X$, Nitrogen oxides; UCI, Upper confidence interval; LCI, Lower confidence interval.

suggesting no relationship or a negative association between the two. For example, Qian Guo et al. (35) found that the risk of childhood obesity elevated by 10.0% (95% CI = 3.0-16.0%) for each 10 μg/$m^3$ increment in PM2.5 exposure. Meanwhile, the risk associated with PM2.5 was significantly higher in groups that were older or lived in urban areas. Another prospective cohort study suggested a negative correlation between decreasing PM2.5 concentrations and the prevalence of obesity in children and

TABLE 3 MR results of obesity trait effects on COVID-19 by IVW.

| Exposure | Outcome | nSNP | Beta | LCI | UCI | p |
|---|---|---|---|---|---|---|
| ASAT | Susceptibility | 146 | 0.036 | 0.0098 | 0.063 | **7.35E-03** |
| BMI | | 812 | 0.16 | 0.13 | 0.19 | **1.74E-27** |
| HC | | 174 | 0.096 | 0.060 | 0.13 | **1.95E-07** |
| OB1 | | 88 | 0.035 | 0.020 | 0.051 | **1.20E-05** |
| OB2 | | 75 | 0.024 | 0.011 | 0.036 | **1.51E-04** |
| OB3 | | 37 | 0.0035 | -0.0060 | 0.013 | 4.75E-01 |
| VAT | | 140 | 0.046 | 0.015 | 0.076 | **3.33E-03** |
| WC | | 160 | 0.12 | 0.079 | 0.16 | **1.44E-08** |
| WHR | | 130 | 0.093 | 0.043 | 0.14 | **3.03E-04** |
| ASAT | Hospitalization | 140 | 0.083 | 0.024 | 0.14 | **5.88E-03** |
| BMI | | 806 | 0.41 | 0.35 | 0.48 | **2.46E-40** |
| HC | | 174 | 0.26 | 0.18 | 0.35 | **1.09E-09** |
| OB1 | | 88 | 0.11 | 0.069 | 0.15 | **7.97E-08** |
| OB2 | | 74 | 0.078 | 0.053 | 0.10 | **5.77E-10** |
| OB3 | | 37 | 0.014 | -0.0073 | 0.036 | 1.95E-01 |
| VAT | | 134 | 0.11 | 0.056 | 0.17 | **1.23E-04** |
| WC | | 160 | 0.38 | 0.29 | 0.48 | **3.23E-16** |
| WHR | | 130 | 0.23 | 0.12 | 0.34 | **4.37E-05** |
| ASAT | Severity | 141 | 0.077 | -0.0058 | 0.16 | 6.84E-02 |
| BMI | | 811 | 0.59 | 0.50 | 0.67 | **1.44E-40** |
| HC | | 174 | 0.36 | 0.25 | 0.48 | **1.64E-09** |

*(Continued)*

TABLE 3 Continued

| Exposure | Outcome | nSNP | Beta | LCI | UCI | p |
|---|---|---|---|---|---|---|
| OB1 | | 89 | 0.14 | 0.089 | 0.20 | **1.75E-07** |
| OB2 | | 75 | 0.10 | 0.065 | 0.14 | **1.20E-07** |
| OB3 | | 37 | 0.031 | -0.0046 | 0.066 | 8.87E-02 |
| VAT | | 134 | 0.11 | 0.015 | 0.20 | **2.24E-02** |
| WC | | 160 | 0.48 | 0.35 | 0.61 | **5.99E-13** |
| WHR | | 131 | 0.26 | 0.087 | 0.44 | **3.32E-03** |

Beta = log (OR). P < 0.05 were bolded.
Abdominal subcutaneous adipose tissue, ASAT; Body mass index, BMI; Hip circumference, HC; OB1, obesity class 1; OB2, obesity class 2; OB3, obesity class 3; Visceral adipose tissue, VAT; Waist circumference, WC; Waist-to-hip ratio, WHR; Upper confidence interval, UCI, Lower confidence interval, LCI.

adolescents, suggesting that cleaning up airborne pollutants could prevent the development of obesity in these populations (36). Sara Fioravanti et al. (37) found no association between exposure to vehicle traffic-related air pollutants and obesity-related indicators such as BMI and abdominal fat during childhood. Jian V Huang et al. (38) found that high air pollutants in childhood were related to a lower BMI at age 13 to 15 years.

Limited reports suggested that air pollution may contribute to obesity by affecting adipocyte function through mechanisms such as cellular inflammation or oxidative stress. For example, animal experiments have shown that PM2.5 exposure may cause metabolic disorders of lipid synthases and fatty acid transporter proteins in adipose tissue and liver through the Nrf2/PPAR pathway, leading to adipose tissue overgrowth (39). Cellular experiments have shown that acute or chronic exposure to PM2.5 can lead to the overproduction of cytoplasmic reactive oxygen species (ROS), induce oxidative damage and activate the oxygen-sensitive NRF2 and NF-kB signaling pathways (40). In current study, we confirmed, using Mendelian randomization analysis, PM2.5 as a direct cause of various obesity-related parameters such as ASAT, BMI, OB1, VAT and WC. In addition, $NO_2$ is a direct cause of elevated risk for HC, OB1 and WC; and NOx is a direct cause of elevated risk for BMI, HC, OB1, OB2, OB3, WC and WHR.

The prevalence of overweight/obesity has continued to increase worldwide over the past half century, currently affecting 2 billion adults, with 770 million having obesity (41). Obesity is a major health challenge because it greatly increases the risk of many chronic diseases, which leads to reduced quality of life and life expectancy (42). In particular, with the focus on COVID-19 from 2019, more and more people are focusing on the correlation between obesity and COVID-19. Studies have shown that obesity may influence the response and prognosis of COVID-19 through a variety of mechanisms such as immune response, metabolic abnormalities and the gut-lung axis (43). An observational study that included 5,279 participants showed that COVID-19 patients with a BMI ≥40 kg/m$^2$ had a more than 2-fold increased risk of hospitalization compared with patients of normal weight (OR = 2.5; 95% CI = 1.8-3.4), after excluding the effects of age, gender, and race (44). Similarly, in another study conducted by Norbert Stefan

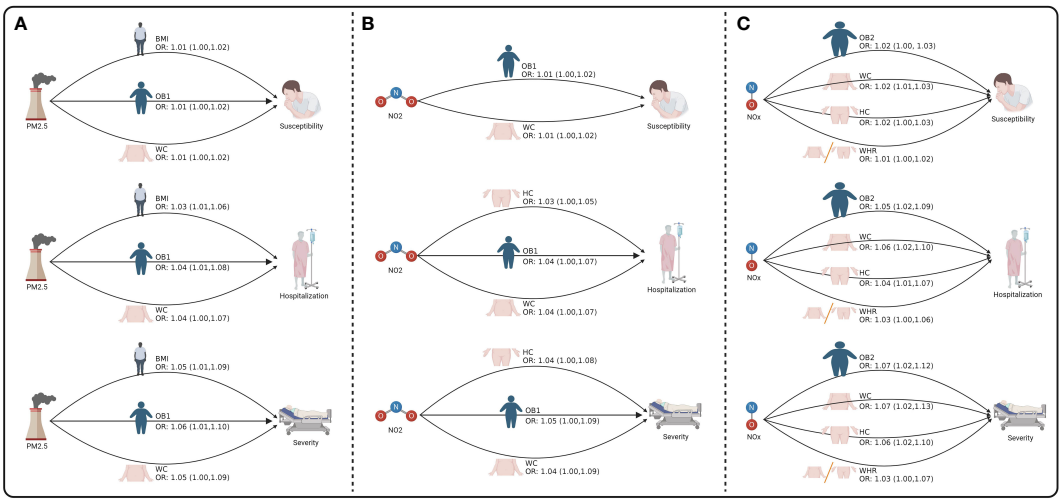



FIGURE 3
The mediating role of obesity between air pollution and COVID-19. **(A)** The mediating role of obesity between PM2.5 and COVID-19. **(B)** The mediating role of obesity between NO2 and COVID-19. **(C)** The mediating role of obesity between NOX and COVID-19. Figure built by the Biorender.

10.3389/fendo.2023.1221442

TABLE 4 Significant indirect effects of air pollution on COVID-19 mediated by obesity traits.

| Exposure | Mediate | Outcome | OR | LCI | UCI | p |
|---|---|---|---|---|---|---|
| PM2.5 | BMI | Susceptibility | 1.01 | 1.00 | 1.02 | 7.41E-03 |
| | OB1 | | 1.01 | 1.00 | 1.02 | 2.05E-02 |
| | WC | | 1.01 | 1.00 | 1.02 | 4.10E-02 |
| | BMI | Hospitalization | 1.03 | 1.01 | 1.06 | 6.82E-03 |
| | OB1 | | 1.04 | 1.01 | 1.08 | 1.49E-02 |
| | WC | | 1.04 | 1.00 | 1.07 | 3.44E-02 |
| | BMI | Severity | 1.05 | 1.01 | 1.09 | 6.82E-03 |
| | OB1 | | 1.06 | 1.01 | 1.10 | 1.55E-02 |
| | WC | | 1.05 | 1.00 | 1.09 | 3.61E-02 |
| NO2 | OB1 | Susceptibility | 1.01 | 1.00 | 1.02 | 3.87E-02 |
| | WC | | 1.01 | 1.00 | 1.02 | 4.16E-02 |
| | HC | Hospitalization | 1.03 | 1.00 | 1.05 | 4.92E-02 |
| | OB1 | | 1.04 | 1.00 | 1.07 | 3.16E-02 |
| | WC | | 1.04 | 1.00 | 1.07 | 3.49E-02 |
| | HC | Severity | 1.04 | 1.00 | 1.08 | 4.95E-02 |
| | OB1 | | 1.05 | 1.00 | 1.09 | 3.24E-02 |
| | WC | | 1.04 | 1.00 | 1.09 | 3.67E-02 |
| NOx | HC | Susceptibility | 1.02 | 1.00 | 1.03 | 7.29E-03 |
| | OB2 | | 1.02 | 1.00 | 1.03 | 1.62E-02 |
| | WC | | 1.02 | 1.01 | 1.03 | 5.85E-03 |
| | WHR | | 1.01 | 1.00 | 1.02 | 2.95E-02 |
| | HC | Hospitalization | 1.04 | 1.01 | 1.07 | 5.35E-03 |
| | OB2 | | 1.05 | 1.02 | 1.09 | 5.41E-03 |
| | WC | | 1.06 | 1.02 | 1.10 | 3.26E-03 |
| | WHR | | 1.03 | 1.00 | 1.06 | 2.33E-02 |
| | HC | Severity | 1.06 | 1.02 | 1.10 | 5.45E-03 |
| | OB2 | | 1.07 | 1.02 | 1.12 | 7.30E-03 |
| | WC | | 1.07 | 1.02 | 1.13 | 3.86E-03 |
| | WHR | | 1.03 | 1.00 | 1.07 | 4.56E-02 |

Beta = log (OR).
PM2.5, Particulate matter air pollution; $NO_2$, Nitrogen dioxide; $NO_X$, Nitrogen oxides; ASAT, abdominal subcutaneous adipose tissue; BMI, body mass index; HC, hip circumference; OB1, obesity class 1; OB2, obesity class 2; OB3, obesity class 3; VAT, visceral adipose tissue; WC, waist circumference; WHR, waist-to-hip ratio; UCI, Upper confidence interval; LCI, Lower confidence interval.

et al. (45), the adjusted OR for admission of COVID-19 inpatients with currently states obesity or BMI >= 30 kg/m$^2$ in the past 12 months was 1.43. Meanwhile, many evidence support that obesity is a key indicator for the severity of COVID-19 inpatients (46, 47). These studies generally concluded that obesity prolongs the time to intensive care unit admission, intubation, and mechanical ventilation in COVID-19 patients (48).

Another observational study including more than 140,000 COVID-19 patients showed that the adjusted risk ratio for patients with BMI >=45 kg/m$^2$ admitted to intensive care unit (ICU) was 1.16 (95% CI = 1.11-1.20). And, the adjusted risk ratio for patients on mechanical ventilation increased from 1.12 (25-29.9 kg/m$^2$) to 2.08 (BMI >= 45 kg/m$^2$) (49). Further studies suggested that obesity may increase the susceptibility and severity of COVID-19 by upregulating the expression of angiotensin-converting enzyme 2 receptors that bind to SARS-CoV-2 (50). Recent studies suggested that obesity may potentially reduce the long-term efficacy of COVID-19 vaccine by affecting the collective immune system, suggesting that we should closely monitor the efficacy of COVID-19 vaccination in this vulnerable group of obesity (51). It has been

shown that the adipocyte membrane receptors ACE2, DPP4 and CD147 as well as the expression of SARS-CoV-2 entry protease-furin are upregulated in patients with obesity (52). These receptors and proteins may therefore be potential targets for SARS-CoV-2 attack and contribute to the severe consequences of COVID-19 in patients with obesity by enhancing systemic inflammation and immune responses. However, these observational studies do not provide powerful evidence for a causal correlation between obesity and COVID-19 risk. In our MR study, we found that most obesity traits, including BMI, HC, OB1, OB2, VAT, WC and WHR, directly increase COVID-19 risk. Our findings are generally consistent with previous observational studies and will provide theoretical support for future prevention of COVID-19 and improved prognosis of patients with COVID-19.

Long-term exposure to air pollution can damage the body's immune system to defend against external pathogens, which can cause a lot of diseases. In recent years, several studies have shown that exposure to air pollution, such as PM2.5, $NO_2$, and $O_3$, can increase the susceptibility and severity of COVID-19 (9). Although the molecular mechanisms by which pollutant exposure affects the pathogenesis of COVID-19 remain unknown. Studies suggested that air pollutants may contribute to virus transmission by modulating mucociliary clearance, altered proteases required for viruses, interferon production, mediated autophagy, immune presenting cell activation, and epithelial cell permeability (34). An epidemiological study from the United Kingdom found that PM2.5 was a major contributor to COVID-19 hospitalization in England, with a 12% increase in COVID-19 cases for every 1 cubic meter increase in the long-term mean PM2.5 (53). Moreover, the relationship between air pollution and COVID-19 mortality remained significant after adjusting for other relevant variables.

An observational study that collected data on COVD-19 from 3,087 countries in the United States showed that a 1 $\mu g/m^3$ elevation in PM2.5 increased COVID-19 mortality by 8% (95% CI: 2%-15%) (54). In this present study, we used Mendelian randomization and found no evidence that air pollution directly increased COVID-19 risk. Interestingly, we found that air pollution can indirectly increase hospitalization, susceptibility and severity of COVID-19 by contributing to obesity. We found that PM2.5 and NOx increased COVID-19 risk (hospitalization and susceptibility) through BMI. PM2.5, NO2 and NOx increased COVID-19 susceptibility through WC. $NO_2$ and NOx increased COVID-19 hospitalization through HC. NOx increased COVID-19 hospitalization through WC and WHR, and increased COVID-19 severity through WC and HC. More attention should be paid to those with obesity living in heavy air pollution in terms of COVID-19 prevention and protection, because obesity caused by air pollution might mediate increasing COVID-19 susceptibility, hospitalization and severity. Encouraging weight loss for this population is needed.

According to our understanding, this is the first systematic exploration of the causal correlation between air pollution and COVID-19 and whether obesity traits play a possible mediating role between them, using an MR approach. We used latest and comprehensive GWAS data (exposure, mediators, and outcomes) to systematically explore the relationship between the three, and will contribute in part to reducing the prevalence of obesity and COVID-19 in the future, as well as raising awareness of environmental protection. However, our study has several limitations. First, our studies were based on online public databases and, therefore, we could not validate them in our own or other databases. Second, obesity may be only one of many mediators of the risk of air pollution affecting COVID-19, and there may be other mediators between the two. Third, this paper only explored the causal correlation between air pollution, obesity and COVID-19 using Mendelian randomization, and the exact molecular mechanisms of the interactions still need to be explored in future studies.

## Conclusion

To summarize, this study exposes a causal relationship between air pollution, obesity and COVID-19. Our results suggested that air pollution can increase the risk of obesity and indirectly increase COVID-19 susceptibility and severity through mediating factors such as obesity. However, the specific mechanism of action between the three has not been clarified, and the detailed pathological mechanisms and molecular pathways need to be further explored in future studies.

## Ethics approval and consent to participate

All data used by this study were publicly available from participant studies with the approvement of the ethical standards committee related to human experimentation.

## Data availability statement

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

## Author contributions

Writing -Original Draft, Methodology, Validation, Visualization: JW and JW-Z. Data Collection and Validation: JW and JW-Z. Conceptualization, Methodology, Supervision, Project Administration and Funding Acquisition: JW, JW-Z, XW, PL. All authors contributed to the article and approved the submitted version.

## Funding

This research was funded by the Natural Science Foundation of Hunan Province (NO. 2023JJ30971, NO.2019JJ80086).

## Acknowledgments

The authors express gratitude to the public databases, websites, and software used in the paper. Especially, we would like to express our gratitude to the UK Biobank, the GIANT consortium, the COVID-19 host gene, and the biorender website. The authors are grateful to the High Performance Computing Center of Central South University for partial support of this work.

## 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.2023.1221442/full#supplementary-material

## References

1. Thompson MG, Burgess JL, Naleway AL, Tyner H, Yoon SK, Meece J, et al. Prevention and attenuation of covid-19 with the BNT162b2 and mRNA-1273 vaccines. *N Engl J Med* (2021) 385:320–9. doi: 10.1056/NEJMoa2107058

2. Poyraz BC, Poyraz CA, Olgun Y, Gurel O, Alkan S, Ozdemir YE, et al. Psychiatric morbidity and protracted symptoms after COVID-19. *Psychiatry Res* (2021) 295:113604. doi: 10.1016/j.psychres.2020.113604

3. O'Mahoney LL, Routen A, Gillies C, Ekezie W, Welford A, Zhang A, et al. The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: A systematic review and meta-analysis. *EClinicalMedicine* (2023) 55:101762. doi: 10.1016/j.eclinm.2022.101762

4. Zhou S, Butler-Laporte G, Nakanishi T, Morrison DR, Afilalo J, Afilalo M, et al. A Neanderthal OAS1 isoform protects individuals of European ancestry against COVID-19 susceptibility and severity. *Nat Med* (2021) 27:659–67. doi: 10.1038/s41591-021-01281-1

5. Solomon MD, Escobar GJ, Lu Y, Schlessinger D, Steinman JB, Steinman L, et al. Risk of severe COVID-19 infection among adults with prior exposure to children. *Proc Natl Acad Sci U.S.A.* (2022) 119:e2204141119. doi: 10.1073/pnas.2204141119

6. Buoli M, Grassi S, Caldiroli A, Carnevali GS, Mucci F, Iodice S, et al. Is there a link between air pollution and mental disorders? *Environ Int* (2018) 118:154–68. doi: 10.1016/j.envint.2018.05.044

7. Shahrbaf MA, Akbarzadeh MA, Tabary M, Khaheshi I. Air pollution and cardiac arrhythmias: A comprehensive review. *Curr Probl Cardiol* (2021) 46:100649. doi: 10.1016/j.cpcardiol.2020.100649

8. Meo SA, Abukhalaf AA, Alessa OM, Alarifi AS, Sami W, Klonoff DC. Effect of environmental pollutants PM2.5, CO, NO(2), and O(3) on the incidence and mortality of SARS-coV-2 infection in five regions of the USA. *Int J Environ Res Public Health 18* (2021) 18(15):7810. doi: 10.3390/ijerph18157810

9. Copat C, Cristaldi A, Fiore M, Grasso A, Zuccarello P, Signorelli SS, et al. The role of air pollution (PM and NO(2)) in COVID-19 spread and lethality: A systematic review. *Environ Res* (2020) 191:110129. doi: 10.1016/j.envres.2020.110129

10. Meo SA, Abukhalaf AA, Alomar AA, Alessa OM. Wildfire and COVID-19 pandemic: effect of environmental pollution PM-2.5 and carbon monoxide on the dynamics of daily cases and deaths due to SARS-COV-2 infection in San-Francisco USA. *Eur Rev Med Pharmacol Sci* (2020) 24:10286–92. doi: 10.26355/eurrev_202010_23253

11. Brombach C, Tong W, Giussani DA. Maternal obesity: new placental paradigms unfolded. *Trends Mol Med* (2022) 28:823–35. doi: 10.1016/j.molmed.2022.05.013

12. Cheng Y, Han J, Li Q, Shi Y, Zhong F, Wu Y, et al. Metabolic obesity phenotypes: a friend or foe of digestive polyps?-An observational study based on National Inpatient Database. *Metabolism* (2022) 132:155201. doi: 10.1016/j.metabol.2022.155201

13. Malik VS, Hu FB. The role of sugar-sweetened beverages in the global epidemics of obesity and chronic diseases. *Nat Rev Endocrinol* (2022) 18:205–18. doi: 10.1038/s41574-021-00627-6

14. Aung N, Khanji MY, Munroe PB, Petersen SE. Causal inference for genetic obesity, cardiometabolic profile and COVID-19 susceptibility: A mendelian randomization study. *Front Genet* (2020) 11:586308. doi: 10.3389/fgene.2020.586308

15. He J, Hu S, Xu X, Guo P, Niu Y, Zhang J, et al. Association of long-term exposure to PM(2.5) in workplace with fasting plasma glucose among asymptomatic adults: A multicenter study in North China. *Environ Int* (2022) 166:107353. doi: 10.1016/j.envint.2022.107353

16. Tamayo-Ortiz M, Tellez-Rojo MM, Rothenberg SJ, Gutierrez-Avila I, Just AC, Kloog I, et al. Exposure to PM(2.5) and obesity prevalence in the greater Mexico City area. *Int J Environ Res Public Health* (2021) 18(5):2301. doi: 10.3390/ijerph18052301

17. Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet* (1986) 1:507–8. doi: 10.1016/S0140-6736(86)92972-7

18. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *JAMA* (2017) 318:1925–6. doi: 10.1001/jama.2017.17219

19. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: The STROBE-MR statement. *JAMA* (2021) 326:1614–21. doi: 10.1001/jama.2021.18236

20. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* (2018) 562:203–9. doi: 10.1038/s41586-018-0579-z

21. Bahcall OG. UK Biobank - a new era in genomic medicine. *Nat Rev Genet* (2018) 19(12):737. doi: 10.1038/s41576-018-0065-3

22. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM (10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. *Environ. Sci. Technol.* (2012) 46(20):11195–205. doi: 10.1021/es301948k

23. Eeftens M, Meier R, Schindler C, Aguilera I, Phuleria H, Ineichen A, et al. (2016). Development of land use regression models for nitrogen dioxide, ultrafine particles, lung deposited surface area, and four other markers of particulate matter pollution in the Swiss SAPALDIA regions. *Environ Health* (2016) 15:53. doi: 10.1186/s12940-016-0137-9

24. 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) 27:3641–9. doi: 10.1093/hmg/ddy271

25. 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

26. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* (2015) 518:187–96. doi: 10.1038/nature14132

27. Initiative C-HG. A first update on mapping the human genetic architecture of COVID-19. *Nature* (2022) 608:E1–E10. doi: 10.1038/s41586-022-04826-7

28. Pei Y-F, Liu Y-Z, Yang X-L, Zhang H, Feng G-J, Wei X-T, et al. The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. *Commun Biol* (2020) 3:608. doi: 10.1038/s42003-020-01334-0

29. Liu C, Zheng S, Gao H, Yuan X, Zhang Z, Xie J, et al. Causal relationship of sugar-sweetened and sweet beverages with colorectal cancer: a Mendelian randomization study. *Eur J Nutr* (2023) 62:379–83. doi: 10.1007/s00394-022-02993-x

30. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* (2018) 362:k601. doi: 10.1136/bmj.k601

31. Wu F, Huang Y, Hu J, Shao Z. Mendelian randomization study of inflammatory bowel disease and bone mineral density. *BMC Med* (2020) 18:312. doi: 10.1186/s12916-020-01778-5

32. 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

33. Zhao SS, Holmes MV, Zheng J, Sanderson E, Carter AR. The impact of education inequality on rheumatoid arthritis risk is mediated by smoking and body mass index: Mendelian randomization study. *Rheumatol (Oxford)* (2022) 61:2167–75. doi: 10.1093/rheumatology/keab654

34. Woodby B, Arnold MM, Valacchi G. SARS-CoV-2 infection, COVID-19 pathogenesis, and exposure to air pollution: What is the connection? *Ann N Y Acad Sci* (2021) 1486:15–38. doi: 10.1111/nyas.14512

35. Guo Q, Xue T, Jia C, Wang B, Cao S, Zhao X, et al. Association between exposure to fine particulate matter and obesity in children: A national representative cross-sectional study in China. *Environ Int* (2020) 143:105950. doi: 10.1016/j.envint.2020.105950

36. Liang X, Liu F, Liang F, Ren Y, Tang X, Luo S, et al. Association of decreases in PM2.5 levels due to the implementation of environmental protection policies with the incidence of obesity in adolescents: A prospective cohort study. *Ecotoxicol Environ Saf* (2022) 247:114211. doi: 10.1016/j.ecoenv.2022.114211

37. Fioravanti S, Cesaroni G, Badaloni C, Michelozzi P, Forastiere F, Porta D. Traffic-related air pollution and childhood obesity in an Italian birth cohort. *Environ Res* (2018) 160:479–86. doi: 10.1016/j.envres.2017.10.003

38. Huang JV, Leung GM, Schooling CM. The association of air pollution with body mass index: evidence from Hong Kong's "Children of 1997" birth cohort. *Int J Obes (Lond)* (2019) 43:62–72. doi: 10.1038/s41366-018-0070-9

39. Jiang M, Li D, Piao J, Li J, Sun H, Chen L, et al. Real-ambient exposure to air pollution exaggerates excessive growth of adipose tissue modulated by Nrf2 signal. *Sci Total Environ* (2020) 730:138652. doi: 10.1016/j.scitotenv.2020.138652

40. Leclercq B, Kluza J, Antherieu S, Sotty J, Alleman LY, Perdrix E, et al. Air pollution-derived PM(2.5) impairs mitochondrial function in healthy and chronic obstructive pulmonary diseased human bronchial epithelial cells. *Environ pollut* (2018) 243:1434–49. doi: 10.1016/j.envpol.2018.09.062

41. Bluher M. Obesity: global epidemiology and pathogenesis. *Nat Rev Endocrinol* (2019) 15:288–98. doi: 10.1038/s41574-019-0176-8

42. Jiang C, Cifu AS, Sam S. Obesity and weight management for prevention and treatment of type 2 diabetes. *JAMA* (2022) 328:389–90. doi: 10.1001/jama.2022.10338

43. Alberca RW, Oliveira LM, Branco A, Pereira NZ, Sato MN. Obesity as a risk factor for COVID-19: an overview. *Crit Rev Food Sci Nutr* (2021) 61:2262–76. doi: 10.1080/10408398.2020.1775546

44. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* (2020) 369:m1966. doi: 10.1136/bmj.m1966

45. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among black patients and white patients with covid-19. *N Engl J Med* (2020) 382:2534–43. doi: 10.1056/NEJMsa2011686

46. Grasselli G, Greco M, Zanella A, Albano G, Antonelli M, Bellani G, et al. Risk factors associated with mortality among patients with COVID-19 in intensive care units in lombardy, Italy. *JAMA Intern Med* (2020) 180:1345–55. doi: 10.1001/jamainternmed.2020.3539

47. Tartof SY, Qian L, Hong V, Wei R, Nadjafi RF, Fischer H, et al. Obesity and mortality among patients diagnosed with COVID-19: Results from an integrated health care organization. *Ann Intern Med* (2020) 173:773–81. doi: 10.7326/M20-3742

48. Stefan N, Birkenfeld AL, Schulze MB. Global pandemics interconnected - obesity, impaired metabolic health and COVID-19. *Nat Rev Endocrinol* (2021) 17:135–49. doi: 10.1038/s41574-020-00462-1

49. Kompaniyets L, Goodman AB, Belay B, Freedman DS, Sucosky MS, Lange SJ, et al. Body mass index and risk for COVID-19-related hospitalization, intensive care unit admission, invasive mechanical ventilation, and death - United States, march-december 2020. *MMWR Morb Mortal Wkly Rep* (2021) 70:355–61. doi: 10.15585/mmwr.mm7010e4

50. Sanchis-Gomar F, Lavie CJ, Mehra MR, Henry BM, Lippi G. Obesity and outcomes in COVID-19: When an epidemic and pandemic collide. *Mayo Clin Proc* (2020) 95:1445–53. doi: 10.1016/j.mayocp.2020.05.006

51. Westheim AJF, Bitorina AV, Theys J, Shiri-Sverdlov R. COVID-19 infection, progression, and vaccination: Focus on obesity and related metabolic disturbances. *Obes Rev* (2021) 22:e13313. doi: 10.1111/obr.13313

52. Ritter A, Kreis NN, Louwen F, Yuan J. Obesity and COVID-19: Molecular mechanisms linking both pandemics. *Int J Mol Sci* (2020) 21(16):5793. doi: 10.3390/ijms21165793

53. Travaglio M, Yu Y, Popovic R, Selley L, Leal NS, Martins LM. Links between air pollution and COVID-19 in England. *Environ pollut* (2021) 268:115859. doi: 10.1016/j.envpol.2020.115859

54. Brandt EB, Beck AF, Mersha TB. Air pollution, racial disparities, and COVID-19 mortality. *J Allergy Clin Immunol* (2020) 146:61–3. doi: 10.1016/j.jaci.2020.04.035

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 23 October 2023
DOI 10.3389/fendo.2023.1236878



OPEN ACCESS

EDITED BY
Xiao Wang,
Lund University, Sweden

REVIEWED BY
Marco Centanni,
Sapienza University of Rome, Italy
Rosario Le Moli,
University of Catania, Italy

*CORRESPONDENCE
Carlo Cappelli
carlo.cappelli@unibs.it

RECEIVED 08 June 2023
ACCEPTED 11 October 2023
PUBLISHED 23 October 2023

CITATION
Gatta E, Anelli V, Cimino E, Di Lodovico E, Piovani E, Zammarchi I, Gozzoli G, Maltese V, Cavadini M, Agosti B, Delbarba A, Pirola I, Girelli A, Buoso C, Bambini F, Alfieri D, Bremi W, Facondo P, Lupo R, Bezzi F, Fredi M, Mazzola AM, Gandossi E, Saullo M, Marini F, Licini M, Pezzaioli LC, Pini L, Franceschini F, Ricci C and Cappelli C (2023) Autoimmune polyglandular syndrome type 4: experience from a single reference center. *Front. Endocrinol.* 14:1236878. doi: 10.3389/fendo.2023.1236878



# Autoimmune polyglandular syndrome type 4: experience from a single reference center

Elisa Gatta[1], Valentina Anelli[1], Elena Cimino[2], Elena Di Lodovico[3], Elda Piovani[4], Irene Zammarchi[5], Giorgia Gozzoli[4], Virginia Maltese[1], Maria Cavadini[1], Barbara Agosti[3], Andrea Delbarba[3], Ilenia Pirola[1], Angela Girelli[2], Caterina Buoso[1], Francesca Bambini[1], Daniele Alfieri[5], Walter Bremi[5], Paolo Facondo[1], Roberto Lupo[4], Francesco Bezzi[4], Micaela Fredi[4], Anna Maria Mazzola[5], Elena Gandossi[3], Maura Saullo[3], Fiorella Marini[3], Massimo Licini[1], Letizia Chiara Pezzaioli[1], Laura Pini[6], Franco Franceschini[4], Chiara Ricci[5] and Carlo Cappelli[1]*

[1]Department of Clinical and Experimental Sciences, SSD Endocrinologia, University of Brescia, ASST Spedali Civili of Brescia, Brescia, Italy, [2]UOC Medicina Generale ad indirizzo Metabolico e Diabetologico, ASST Spedali Civili of Brescia, Brescia, Italy, [3]Sindacato Unico Medicina Ambulatoriale Italiana e Professionalità dell'Area Sanitaria – SUMAI, Brescia, Italy, [4]Department of Clinical and Experimental Sciences, Rheumatology and Clinical Immunology, University of Brescia, ASST Spedali Civili of Brescia, Brescia, Italy, [5]Department of Clinical and Experimental Sciences, Gastroenterology Unit, University of Brescia, ASST Spedali Civili of Brescia, Brescia, Italy, [6]Department of Clinical and Experimental Sciences, Respiratory Medicine Unit, University of Brescia, ASST Spedali Civili of Brescia, Brescia, Italy

**Purpose:** To characterize patients with APS type 4 among those affected by APS diagnosed and monitored at our local Reference Center for Autoimmune Polyglandular Syndromes.

**Methods:** Monocentric observational retrospective study enrolling patients affected by APS diagnosed and monitored in a Reference Center. Clinical records were retrieved and analyzed.

**Results:** 111 subjects (51 males) were affected by APS type 4, mean age at the onset was 23.1 ± 15.1 years. In 15 patients the diagnosis of APS was performed during the first clinical evaluation, in the other 96 after a latency of 11 years (range 1-46). The most frequent diseases were type I diabetes mellitus and celiac disease, equally distributed among sexes.

**Conclusions:** The prevalence of APS type 4 is 9:100,000 people. Type I diabetes mellitus was the leading indicator of APS type 4 in 78% subjects and in 9% permitted the diagnosis occurring as second manifestation of the syndrome. Our

data, showing that 50% of patients developed APS type 4 within the first ten years, don't suggest any particular follow-up time and, more importantly, don't specify any particular disease. It is important to emphasize that 5% of women developed premature ovarian failure.

KEYWORDS

autoimmune polyglandular syndrome, type I diabetes mellitus, autoimmune diseases, polyendocrinopathies, autoimmunity

# 1 Introduction

Autoimmune polyglandular syndromes (APS) are rare and registered orphan diseases (ORPHAcode ORPHA:282196) characterized by insidious presentation and circulating autoantibodies and lymphocytic infiltration of one or more endocrine glands, with possible additional involvement of non-endocrine organs, eventually leading to organ failure (1–3).

There is a broad heterogeneity of APS and these manifest sequentially with a variable time interval between the occurrence of the diseases (4). The original classification into 4 types by Neufeld et al. in 1980 was revised by Betterle and Zanchetta in 2003, sub-classifying APS type 3 into four different sub-groups (5, 6). Indeed, some authors considered only APS-1 and APS-2 and did not consider APS-3 and APS-4 as independent entities (2, 7, 8).

APS type 1 (ORPHA:3453), due to different mutations of the autoimmune regulator (AIRE) gene on chromosome 21, is characterized by the presence of chronic candidiasis, chronic hypoparathyroidism, and Addison's disease (1, 6). It has onset in childhood with an estimated prevalence of 1:80,000 live births (9) and a slight female predominance (10).

APS type 2 (ORPHA:3143), associated with a genetic pattern of human leukocyte antigen (HLA) DR3/DR4, is characterized by the presence of Addison's disease and autoimmune thyroid diseases and/or type 1 diabetes mellitus (T1DM) (1, 6). The onset is predominantly in young adulthood (2), with a prevalence of 1:20,000 and a sex ratio male/female 1:3 (11).

Autoimmune thyroid diseases associated with other autoimmune diseases (excluding Addison's disease and/or hypoparathyroidism) fall under APS type 3 (ORPHA:227982) (1, 5, 12). This syndrome is subdivided into type 3A if associated with other endocrine diseases, type 3B if associated with gastrointestinal diseases, type 3C with skin, haemopoietic system or nervous system diseases and type 3D with rheumatic diseases (6). The actual incidence is estimated at 1:20,000 and it is three times more frequent among women (11, 13). APS type 4 (ORPHA:227990) includes all the different clinical combinations of autoimmune diseases not included in the previous groups and affecting an endocrine organ (with the exception of Addison's disease, thyroid diseases, or hypoparathyroidism) in combination with at least one more endocrine or non-endocrine organs (1, 6). To the best of our knowledge, there is scarce clinical data and no epidemiological data on this category globally.

The aim of the present study was to characterize patients with APS type 4 among those affected by APS diagnosed and monitored at our local Reference Center for Autoimmune Polyglandular Syndromes.

# 2 Methods

## 2.1 Subjects

All the medical records of patients referred for autoimmune disease to the Units of Endocrinology, Diabetology, Gastroenterology, Rheumatology and Clinical Immunology at the ASST Spedali Civili in Brescia were retrospectively reviewed from January 2000 up to 30 November 2022. All the patients were screened for the most frequent autoimmune endocrinopathies annually, as well as for Addison's disease and all autoimmune pathologies when clinically suspected. All patients affected by Autoimmune Polyglandular Syndrome, according to the ORPHAcode (ORPHA 282196, 3453, 3143, 227982, 227990), were included in this study. The study (ASST_BS_CLIN_PZ_SPA-BS) was approved by the local Ethics Committee (no 5517).

## 2.2 Clinical data collection

Clinical manifestations of APS type 4 [including type 1 diabetes mellitus or latent autoimmune diabetes in adults (LADA), premature ovarian failure, celiac disease, atrophic gastritis, inflammatory bowel disease, rheumatoid arthritis, systemic lupus erythematosus, scleroderma, Sjogren syndrome, mixed connective tissue disease, vasculitis, antiphospholipid syndrome, primary biliary cirrhosis, autoimmune hepatitis, alopecia areata, autoimmune urticaria, myasthenia gravis, multiple sclerosis, pernicious anemia, immune thrombocytopenia, vitiligo, seronegative arthritis, ankylosing spondylitis, psoriasis, pemphigoid], as well as related patient information such as sex, onset age or age at diagnosis of the first and subsequent APS manifestations, were retrospectively retrieved from medical records. Diagnosis was performed in accordance with good clinical practice by antibody serology tests and, where required, histopathological analysis (i.e., celiac disease, atrophic gastritis, systemic lupus erythematosus, scleroderma, vasculitis, autoimmune hepatitis). In agreement with ORPHANET classification all the patients affected

or showing antibodies suggesting Addison's disease, thyroid diseases, or hypoparathyroidism were excluded from this study.

## 2.3 Statistical analysis

All data were collected in an electronic case report database. Normal distribution was checked using the Shapiro-Wilk test. Latency results were nonnormally distributed and were not normalized by the usual procedures of data transformation; in these cases, the results are presented as a median, with minimum and maximum values. Comparison between groups and differences in proportion were calculated using the $\chi^2$ test for categorical variables and ANOVA for quantitative variables, as appropriate. Between-group comparison was performed using the Student's T-test for unpaired data or Kruskal-Wallis test, as appropriate. The Kaplan-Meier curve was fitted to determine the APS type 4 diagnosis time. A $p$-value < 0.05 was considered statistically significant. The statistical analyses were performed using SPSS 20.0 software (SPSS, Inc., Evanston, IL, USA). The results are reported in compliance with the STROBE reporting guidelines for cross-sectional studies; the checklist is reported in Supplementary File 1.

# 3 Results

A total of 9164 patients were referred to the Units of Endocrinology, Diabetology, Gastroenterology, Rheumatology and Clinical Immunology for autoimmune diseases. Among these, 1161 (12.7%) were diagnosed with any autoimmune polyglandular syndromes in accordance with good clinical practice by antibody serology tests and, where required, histopathological analysis, such as for all patients with positive antibodies for celiac disease, atrophic gastritis and/or vasculitis. Among the 1161 patients with APS, 111 (9.6%) subjects (51 males) were affected by APS type 4 (47.8 ± 17.1 years old, range 20-85) and were enrolled in the present study.

The mean age at the onset of APS was 23.1 ± 15.1 years, with no significant difference between sexes (22.3 ± 15.2 vs. 23.8 ± 15.0 yrs, M/F, p = .611). APS type 4 was diagnosed during first clinical evaluation in 15/111 (13.5%) patients (Group 1): celiac disease and multiple sclerosis were concomitantly diagnosed during T1DM evaluation in 13 and 2 subjects, respectively (Table 1). These patients did not develop further diseases during follow-up (14.3 ± 8.6, range 1-33 yrs).

APS type 4 was diagnosed in 96 patients in the years following the first disease (range 1-46 yrs) (Group 2); in detail, the most frequent first clinical manifestations were T1DM in 72/96 (75%) patients, celiac disease in 9 (9%) and vitiligo in 4 (4%) (Table 2). Groups 1 and 2 were superimposable for age of disease onset (23.5 ± 13.3 vs. 23.0 ± 15.4 yrs, p=.292) and sex (8/7 vs 43/53, M/F, p=.537).

The development of APS over the years (Group 2) is shown in Figure 1. The diagnosis was reached after a median latency of 11 years (range 1-46) [10 (range 1-46) vs. 11 (range 1.38), M/F, p=.198]. 50% of subjects developed APS within ten years (Figure 1). No difference was found after subdividing the first clinical manifestation into the different diseases to which they referred [11 (range 1-46) vs. 15.5 (range 1-29) vs. 9 (range 2-22) for endocrine, gastroenterological, and rheumatologic diseases, respectively, p=0.643] (Figure 2). Five patients developed two subsequent concomitant manifestations (2 subjects had both connective tissue disease and inflammatory bowel disease, 1 vitiligo and T1DM, 1 vitiligo and celiac disease and 1 celiac disease and rheumatoid arthritis); 3/96 (3%) showed a third disease with latency from the onset of the second of 34 years (range 14-36). The developing features of APS are shown in Figure 3 and Table 3. The most frequent pathway was T1DM followed by celiac disease occurring in 35 (48.6%) patients (15 males) with a latency of 4 years (range 1-28); conversely, all the celiac patients developed T1DM with a latency of 14 years (range 1-29) (Table 3).

The demographic characteristics of all patients affected by autoimmune polyglandular syndrome Type 4 are reported in Table 4. In detail, 108/111 (97%) subjects (51males) showed T1DM. In 87/108 (81%) (40 males) it occurred as the first clinical feature of APS with a mean age at the onset of 22.3 ± 14.8 years old. Celiac disease was diagnosed in 57 (51%) patients (27 males), occurring as first disease in 22 (39%) subjects (12 males) at the age of 23.7 ± 16.3 years old. Rheumatoid arthritis was diagnosed in 12 (11%) patients (2 males). It occurred as first disease in two females (28.5 ± 3.5 yrs) followed by premature ovarian failure (POF) or T1DM after 17 and 22 years, respectively. Vitiligo was diagnosed in 11 (10%) patients, and in 4 (2 males) as first manifestation followed by three cases of diabetes mellitus and one case of POF. This last condition was diagnosed in 3/60 women (5%), showing anti-21-hydroxylase antibodies.

# 4 Discussion

The present study describes for the first time the prevalence of APS type 4 among a large series of patients affected by autoimmune diseases.

Autoimmune polyglandular syndromes are rare orphan diseases encompassing a wide spectrum of autoimmune disease, with the involvement of endocrine and non-endocrine organs (1,

TABLE 1 Demographic and clinical characteristics of patients with diagnosis of APS type 4 at the first clinical evaluation.

| Disease | Number of patients (%) | Sex M/F | Overall mean age of diagnosis (years) | Latency before second manifestation (years) | Total follow-up from first manifestation (years) |
|---|---|---|---|---|---|
| Type I diabetes mellitus and celiac disease | 13 (86.7%) | 8/5 | 23.5 ± 14.3 | NA | 13.7 ± 6.9 |
| Type I diabetes mellitus and multiple sclerosis | 2 (13.3%) | 0/2 | 23.5 ± 4.9 | NA | 18.5 ± 20.5 |

NA, not applicable.

10.3389/fendo.2023.1236878

TABLE 2 Demographic and clinical characteristics of patients developing APS type 4.

| Disease | Number of patients (%) | Sex M/F | Overall mean age of diagnosis (years) | Median latency before second manifestation (years) | Total follow-up from first manifestation (years) |
|---|---|---|---|---|---|
| Type I diabetes mellitus | 72 (75.0%) | 32/40 | 22.0 ± 15.1 | 11.5 (1–46) | 26.7 ± 12.6 |
| Celiac disease | 9 (9.4%) | 4/5 | 24.0 ± 19.8 | 13 (1-29) | 24.2 ± 11.0 |
| Vitiligo | 4 (4.2%) | 2/2 | 19.5 ± 15.2 | 9 (2-20) | 24.3 ± 12.3 |
| Psoriasis | 3 (3.1%) | 2/1 | 18.3 ± 6.5 | 9 (6-16) | 31.3 ± 21.1 |
| Inflammatory bowel disease | 2 (2.1%) | 1/1 | 34.0 ± 26.9 | 19.5 (19-20) | 25.5 ± 0.7 |
| Rheumatoid arthritis | 2 (2.1%) | 0/2 | 28.5 ± 3.5 | 14.5 (7-22) | 23.5 ± 9.2 |
| Premature ovarian failure | 1 (1.0%) | 0/1 | 30 | 8 | 22 |
| Seronegative arthritis | 1 (1.0%) | 0/1 | 42 | 2 | 13 |
| Ankylosing spondylitis | 1 (1.0%) | 1/0 | 42 | 10 | 28 |
| Primary biliary cirrhosis | 1 (1.0%) | 1/0 | 41 | 20 | 32 |

10). There are many studies that describe these syndromes (2, 3, 6–8, 11), but few articles, mainly series of case reports, have focused on APS type 4. This can also be due to the fact that some Authors consider APS type 2, 3 and 4 as different phenotypes of the same underlying disease mechanism classifying them as a single entity (2, 7). However, taking in account that each APS type is uniquely characterized by a unique endocrinopathy, we recognized them as separate entity in agreement with the Resource of Rare Disease Co-founded by the Health Programme of European Union (1).

To the best of our knowledge, there is no data on its prevalence reported worldwide (1, 14). One possible explanation could be the large heterogeneity of conditions characterizing the syndrome, which can lead to over- or under-diagnosis. In fact, the few studies about APS type 4 include patients affected by autoimmune thyroid disease, hypoparathyroidism and/or adrenal insufficiency that, according to the Orphanet definition, exclude *a priori* APS type 4 (1). Another reason could be the scarcity of studies performed in a large series of patients affected by "trigger" diseases referred to a single center; we

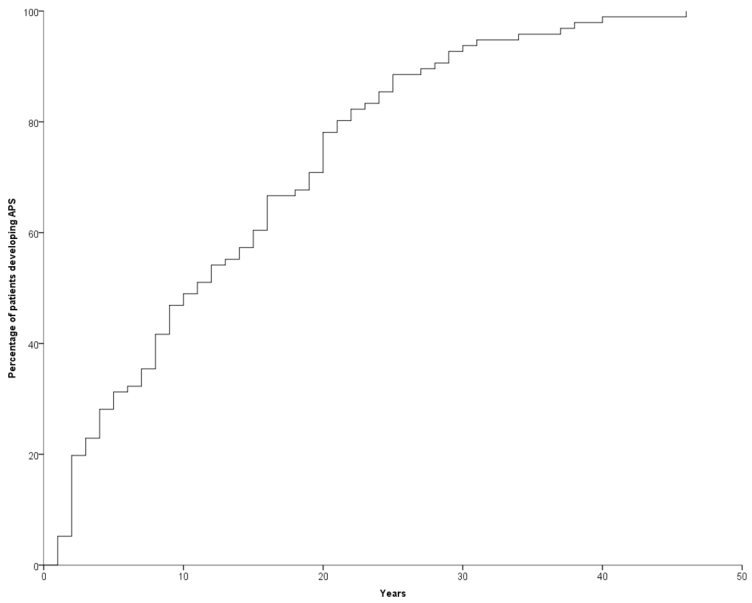



FIGURE 1
Temporal trend of manifestation of the second disease.

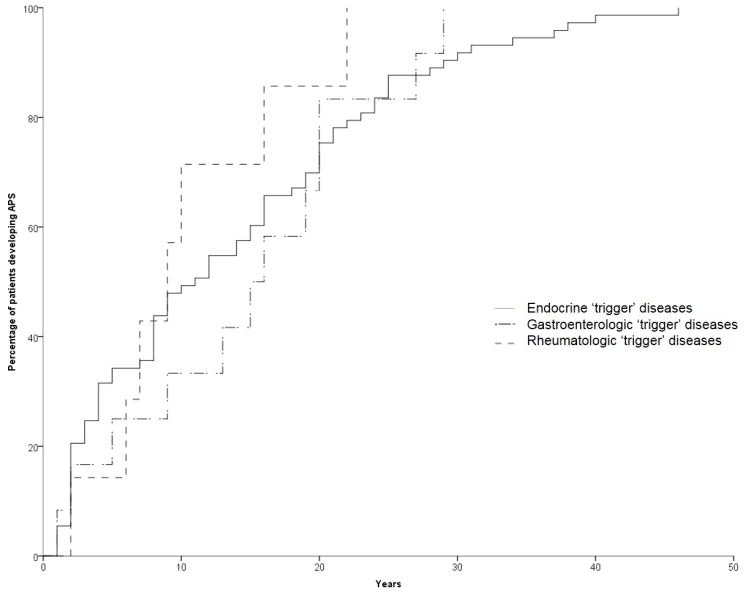



FIGURE 2
Temporal trend of manifestation of the second disease according to the first "trigger" disease.

believe that this is the key point of the present study. As reported above, we carefully selected patients affected by APS type 4, in accordance with the ORPHAcode, among those referred to our units for autoimmune diseases. Keeping in mind our data and taking into account the population of our province (1,253,993 inhabitants) (15), the estimated prevalence of APS type 4 is 9 cases per 100,000, thus classifying it as a rare disease as defined by the European Union Regulation on Orphan Medicinal Products (1).

Again, no data about sex distribution are reported. However, Frommer and Kahaly showed a sex ratio (M:F) of 1:3 in adult patients affected by APS type 2, 3 and 4 (11). On the contrary, we found a male to female ratio of 1:1 in our sample. Although most autoimmune diseases are more common in females, no sex difference in the overall incidence of youth T1DM is demonstrated (16, 17). Our ratio appears to be in keeping with literature, since T1DM is present in almost all patients (97%). T1DM is a key element *in* and *for* the diagnosis of APS type 4. In fact, T1DM was the leading indicator of APS type 4 in 87/111 (78%) subjects and in 21/111 (19%) permitted the diagnosis as the second manifestation of the syndrome.

About ten years ago, Van den Driessche et al. suggested a flowchart for screening and follow-up of few associated autoimmune disorders (autoimmune thyroid diseases, atrophic gastritis, celiac disease, Addison's disease, and vitiligo) in patients affected by T1DM. With the exception of thyroid diseases, the authors proposed an annual screening for the first three years and then once every 5 years; the thyroid should be checked annually. However, this recommendation covered all the autoimmune syndromes except APS type 1 (18). Our data, showing that 50% of patients developed APS type 4 within the first ten years, don't suggest any particular follow-up time and, more importantly, don't specify any particular disease (Figure 1). In other words, these data suggest a lifelong follow-up, although the cost-effectiveness of this is yet to be proven. However, in 68% of patients with T1DM developing celiac disease this occurred within 10 years. For this reason, the screening of celiac disease should be done very early since therapy of T1DM is very difficult in patients with unknown celiac disease (19). In addition, it is important to emphasize that 5% of women in our series developed POF. These data reinforce what Li et al. previously reported on the prevalence of autoimmune disorders in women affected by POF (20). As is well known, this condition severely affects women's lives (21–25). For this reason, we suggest and recommend that gynecologists perform regular check-ups with complete blood exams during childbearing age for these patients.

Celiac disease was the second most frequent disease among our patients (57/111, 51%). Literature data show that females are affected approximately twice as often as males, although the ratio varies depending on the strategy used to identify cases (26). Our male to female ratio was instead 1:1 among our patients. As reported above, this could be because all these patients were also affected by T1DM, and celiac disease is known to be closely associated with type 1 diabetes mellitus (27–31). These conditions share the same HLA susceptibility alleles, specifically DR3/DQ2 and DR4/DQ8 molecules (32). However, the co-occurrence of the disorders is not fully explained by shared genetic risk loci (30). Few studies in animal models (33, 34) and in humans (35) have shown that celiac disease may trigger autoimmune processes leading to diabetes. On the other hand, some authors have reported the development of celiac autoantibodies after the onset of diabetes (30, 36, 37). In line with this finding, 61% of our patients developed celiac disease 1-28 years after T1DM diagnosis (Table 3).

As is widely known, it is the pathologic response to self- or autoantigens that characterize autoimmune diseases. It is generically categorized as autoimmunity or autoreactivity, which covers a wide range of clinical disorders (38). The pathogenesis of autoimmune disease is still largely unknown: familial or genetic, infectious, immunologic, and psychological factors have all been implicated as

10.3389/fendo.2023.1236878

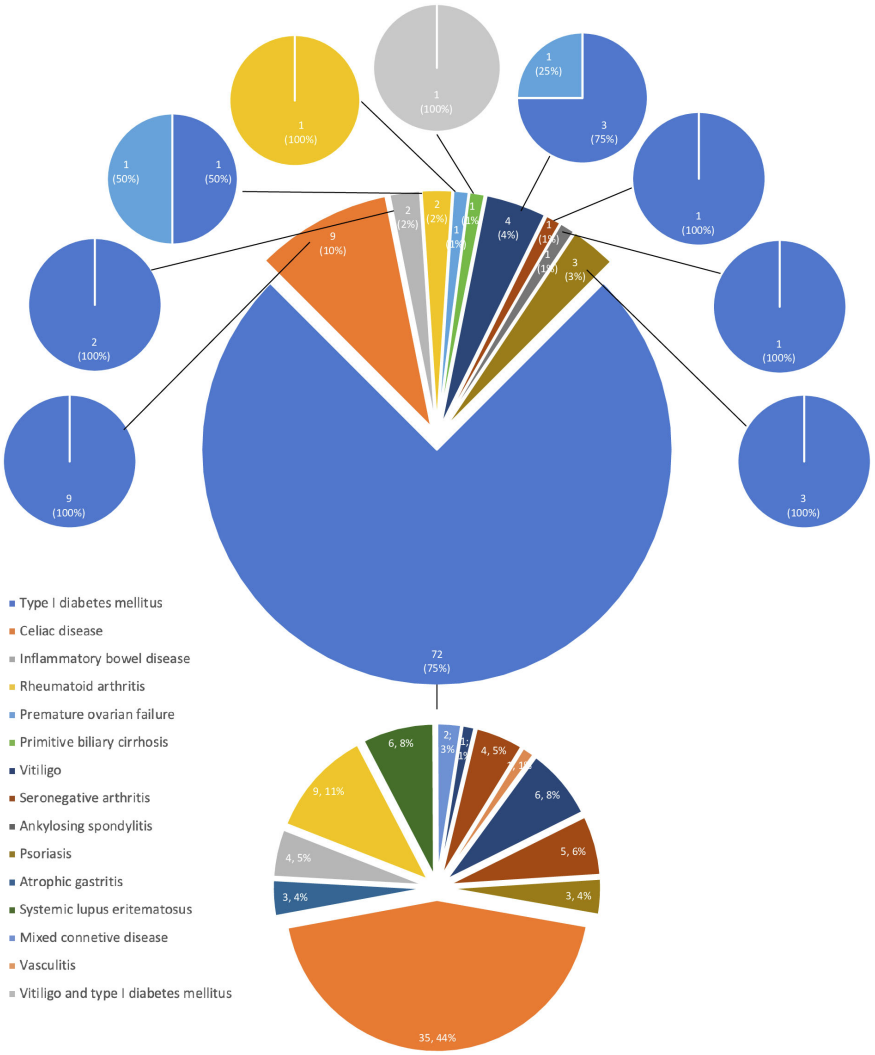



FIGURE 3
Developing features of APS in accordance with the first disease.

TABLE 3 Pathway of subsequent APS manifestation according to the first clinical feature.

| Second disease | Number of patients (%) | Sex M/F | Latency and range (years) | Latency M/F (years) | *p* |
|---|---|---|---|---|---|
| Type I diabetes mellitus as first disease | | | | | |
| Celiac disease | 35 (48.6%) | 15/20 | 4 (1 – 28) | M 4 (1-28)<br>F 4.5 (1-25) | .198 |
| Rheumatoid arthritis | 9 (12.5%) | 2/7 | 20 (3 – 38) | M 19.5 (14-25)<br>F 20 (3 – 38) | .397 |
| Systemic lupus erythematosus | 6 (8.3%) | 2/4 | 11 (2 – 37) | M 22.5 (8-37)<br>F 9 (2-20) | .009 |
| Vitiligo | 6 (8.3%) | 4/2 | 20 (2 – 40) | M 20 (2-40)<br>F 15.5 (9-22) | .696 |
| Seronegative arthritis | 5 (6.9%) | 2/3 | 20 (12 – 36) | M 23 (12-34)<br>F 20 (15-36) | .512 |

*(Continued)*

10.3389/fendo.2023.1236878

TABLE 3 Continued

| Second disease | Number of patients (%) | Sex M/F | Latency and range (years) | Latency M/F (years) | *p* |
|---|---|---|---|---|---|
| Inflammatory bowel disease | 4 (5.6%) | 2/2 | 25 (16 – 34) | M 31.5 (34-29)<br>F 18.5 (16-21) | .054 |
| Multiple sclerosis | 4 (5.6%) | 1/3 | 24 (14 – 30) | M 24<br>F 24 (14-30) | NA |
| Atrophic gastritis | 3 (4.2%) | 1/2 | 8 (4 – 16) | M 4<br>F 12.0 (8-16) | NA |
| Psoriasis | 3 (4.2%) | 3/0 | 23 (12 – 46) | M 23 (12-46) | NA |
| Mixed connective tissue disease | 2 (2.8%) | 1/1 | 25 (21 – 29) | M 29<br>F 21 | NA |
| Vasculitis | 1 (1.4%) | 1/0 | 8 | M 8 | NA |
| Immune thrombocytopenia | 1 (1.4%) | 1/0 | 3 | M 3 | NA |
| **Celiac disease as first disease** | | | | | |
| Type I diabetes mellitus | 9 (100%) | 4/5 | 13 (1 – 29) | M 15.5 (2-27)<br>F 9 (1-29) | .684 |
| **Vitiligo as first disease** | | | | | |
| Type I diabetes mellitus | 3 (75%) | 2/1 | 7 (2 – 20) | M 11 (2-20)<br>F 7 | NA |
| Premature ovarian failure | 1 (25%) | 0/1 | 11 | F 11 | NA |
| **Psoriasis as first disease** | | | | | |
| Type I diabetes mellitus | 3 (100%) | 2/1 | 9 (6 – 16) | M 7.5 (6-9)<br>F 16 | NA |
| **Inflammatory bowel disease as first disease** | | | | | |
| Type I diabetes mellitus | 2 (100%) | 1/1 | 19.5 (19 – 20) | M 19<br>F 20 | NA |
| **Rheumatoid arthritis as first disease** | | | | | |
| Type I diabetes mellitus | 1 (50%) | 0/1 | 22 | F 22 | NA |
| Premature ovarian failure | 1 (50%) | 0/1 | 17 | F 17 | NA |
| **Premature ovarian failure as first disease** | | | | | |
| Rheumatoid arthritis | 1 (100%) | 0/1 | 7 | F 7 | NA |
| **Primary biliary cirrhosis as first disease** | | | | | |
| Vitiligo and Type I diabetes mellitus | 1 (100%) | 1/0 | 20 | M 20 | NA |
| **Seronegative arthritis as first disease** | | | | | |
| Type I diabetes mellitus | 1 (100%) | 0/1 | 2 | F 2 | NA |

NA, not applicable.

triggers (39, 40). Consequently, it is reasonable to believe that once the trigger, if any, activates the process, then it can be amplified. Moreover, a recent observational study by Bechi Genzano et al. showed that the circulating immune profile was similar in patients diagnosed with T1DM and those affected by other autoimmune diseases. The authors demonstrated an increase in CD4 T-cells and a consensual reduction in natural killer (NK) cells and CD8 T-cells, underlying a similar pathogenetic pathway (31). In addition, major autoimmune disorders share much of their molecular background, including class II HLA haplotypes (41–45). Houcken et al. demonstrated that protein tyrosine phosphatase non-receptor type 22 (*PTPN22*) and cytotoxic T-lymphocyte associated protein 4

10.3389/fendo.2023.1236878

TABLE 4 Overall demographic and clinical characteristics of APS patients.

| Disease | Number of patients (%) | Sex M/F | Overall mean age of diagnosis (years) | Mean age of diagnosis M/F (years) | *p* | Disease $1^{st}$ – $2^{nd}$ – $3^{rd}$ |
|---|---|---|---|---|---|---|
| Type I diabetes mellitus | 108 (97.3%) | 51/57 | 25.4 ± 16.4 | M 25.2 ± 16.8<br>F 25.6 ± 16.2 | .828 | 87 – 20 – 1 |
| Celiac disease | 57 (51.4%) | 27/30 | 25.7 ± 15.3 | M 23.6 ± 14.2<br>F 25.7 ± 15.3 | .767 | 22 – 35 – 0 |
| Rheumatoid arthritis | 12 (10.8%) | 2/10 | 40.5 ± 15.0 | M 36.7 ± 27.4<br>F 41.7 ± 12.0 | .032 | 2 – 10 – 0 |
| Vitiligo | 11 (9.9%) | 7/4 | 37.1 ± 18.8 | M 37.8 ± 24.8<br>F 36.0 ± 5.4 | .009 | 4 – 7 – 0 |
| Inflammatory bowel diseases | 6 (5.4%) | 3/3 | 47.2 ± 19.7 | M 45.3 ± 7.5<br>F 49.9 ± 30.0 | .068 | 2 – 4 – 0 |
| Systemic lupus erythematosus | 6 (5.4%) | 2/4 | 14.2 ± 13.0 | M 22.5 ± 20.5<br>F 10.0 ± 8.5 | .018 | 0 – 6 – 0 |
| Multiple sclerosis | 6 (5.4%) | 1/5 | 29.7 ± 8.6 | M 42<br>F 27.2 ± 6.8 | NA | 2 – 3 – 1 |
| Psoriasis | 6 (5.4%) | 5/1 | 30.0 ± 18.3 | M 33.6 ± 17.9<br>F 12 | NA | 3 – 3 – 0 |
| Seronegative arthritis | 5 (4.5%) | 1/4 | 27.8 ± 17.1 | M 12<br>F 31.8 ± 16.9 | NA | 1 – 3 – 1 |
| Atrophic gastritis | 3 (2.7%) | 1/2 | 39.0 ± 19.5 | M 59<br>F 20.0 ± 12.7 | NA | 0 – 3 – 0 |
| Premature ovarian failure | 3 (5.0%) | 0/3 | 36.3 ± 5.7 | F 36.3 ± 5.7 | NA | 1 – 2 – 0 |
| Mixed connective tissue disease | 2 (1.8%) | 1/1 | 58.5 ± 19.1 | M 45<br>F 72 | NA | 0 – 2 – 0 |
| Immune thrombocytopenia | 1 (0.9%) | 1/0 | 17 | M 17 | NA | 0 – 1 – 0 |
| Vasculitis | 1 (0.9%) | 1/0 | 63 | M 63 | NA | 0 – 1 – 0 |
| Primary biliary cirrhosis | 1 (0.9%) | 1/0 | 41 | M 41 | NA | 1 – 0 – 0 |
| Ankylosing spondylitis | 1 (0.9%) | 1/0 | 42 | M 42 | NA | 1 – 0 – 0 |

NA, not applicable.

(*CTLA-4*) polymorphisms are also associated with autoimmune polyglandular syndromes (46). Most recently, evidence by Fichna et al. appears to show that BTB domain and CNC homolog 2 (*BACH2*) gene polymorphism, implicated in lymphocyte differentiation and function, may also promote multitarget autoimmunity (47). Genetic screening is growing popular to identify patients at risk of autoimmune disorders, though this remains too expensive to be included in routine clinical management and is not readily available. Furthermore, there is still a lack of evidence as to the usefulness of this process in daily clinical practice. Therefore, we have no genetic screening data for the patients in our study. It is also widely accepted that autoantibodies play a crucial role in the diagnosis of autoimmune disease, especially in the early phases when the patient is still asymptomatic and biochemical markers are normal (7). In a minority of cases, however, patients may not show any autoantibodies, a condition that is referred to as seronegative autoimmune diseases, as recently reviewed by Lenti et al. (48). In these cases, the diagnosis is more challenging and must rely on clinical features and other available tests, often including histopathological evaluation and radiological diagnostic tests. In all our patients, the diagnosis was confirmed by the presence of serum autoantibodies and/or histopathological specimens.

The main limitations of the present study are its retrospective nature and the possible patient drop-out during follow-up after the diagnosis of the first "trigger" disease. The latter is a key point, as it could reduce the prevalence of APS in our cohort of patients. Unfortunately, we have no data on patient drop-out precisely due to the retrospective nature of the study. Indeed, it is unlikely that all these patients were affected by APS type 4. Finally, we must state that a possible bias of case selection is possible since the data collection has been performed in a reference center for autoimmune diseases. However, the large set of patients, the careful selection procedure and detailed analysis of patient clinical records strengthen our results.

In conclusion, the prevalence of APS type 4 is 9:100,000 people. Type 1 diabetes mellitus, for its high prevalence among our patients, could be the clinical "driver" of this syndrome: for this reason, diabetologists should pay particular attention during clinical examinations of T1DM patients. Our data don't suggest any particular follow-up time and, more importantly, don't specify any particular disease, but only indicate a lifelong follow-up. Finally, we recommend regular gynecological evaluations with complete blood exams during childbearing age due to the non-negligible risk of developing premature ovarian failure.

## 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 author.

## Ethics statement

The studies involving humans were approved by Brescia Ethics Committee (no 5517). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

Conceptualization and Methodology: CC, CR, FF, and AG. Data curation: EliG VA, EC, VM, IZ, EP, and GG. Formal analysis: EliG and CC. Investigation: All the authors. Project administration: CC. Supervision, validation, and visualization: IP and LaP. Writing - original manuscript: EliG and CC. Writing – review and editing: CC, CR, FF, VA, IZ, EP, GG, EC, and AG. CC, CR, FF, and AG had full access to and verified all the study data and were responsible for the decision to submit for publication. All authors contributed to the article and approved the submitted version. All authors affirm the accuracy and completeness of the data and attest to the fidelity of the study to the protocol.

## 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.2023.1236878/full#supplementary-material

## References

1. Orphanet. *The portal for rare diseases and orphan drugs*. Available at: https://www.orpha.net/.

2. Husebye ES, MS A, Kämpe O. Autoimmune polyendocrine syndromes. *N Engl J Med* (2018) 378(12):1132–41. doi: 10.1056/NEJMra1713301

3. Betterle C, Sabbadin C, Scaroni C, Presotto F. Autoimmune polyendocrine syndromes (APS) or multiple autoimmune syndromes (MAS). In: Colao A, Jaffrain-Rea M-L, Beckers A, editors. *Polyendocrine disorders and endocrine neoplastic syndromes*. Cham: Springer International Publishing (2019). p. 1–50.

4. Kahaly GJ, Frommer L. Polyglandular autoimmune syndromes. *J Endocrinol Invest* (2018) 41(1):91–8. doi: 10.1007/s40618-017-0740-9

5. Neufeld M, Maclaren N, Blizzard R. Autoimmune polyglandular syndromes. *Pediatr Ann* (1980) 9(4):154–62. doi: 10.3928/0090-4481-19800401-07

6. Betterle C, Zanchetta R. Update on autoimmune polyendocrine syndromes (APS). *Acta BioMed* (2003) 74(1):9–33.

7. Eisenbarth GS, Gottlieb PA. Autoimmune polyendocrine syndromes. *N Engl J Med* (2004) 350(20):2068–79. doi: 10.1056/NEJMra030158.è

8. Kahaly GJ, Frommer L. Autoimmune polyglandular diseases. *Best Pract Res Clin Endocrinol Metab* (2019) 33(6):101344. doi: 10.1016/j.beem.2019.101344

9. Husebye ES, Perheentupa J, Rautemaa R, Kämpe O. Clinical manifestations and management of patients with autoimmune polyendocrine syndrome type I. *J Intern Med* (2009) 265(5):514–29. doi: 10.1111/j.1365-2796.2009.02090.x

10. Hansen MP, Matheis N, Kahaly GJ. Type 1 diabetes and polyglandular autoimmune syndrome: A review. *World J Diabetes* (2015) 6(1):67–79. doi: 10.4239/wjd.v6.i1.67

11. Frommer L, Kahaly GJ. Autoimmune polyendocrinopathy. *J Clin Endocrinol Metab* (2019) 104(10):4769–82. doi: 10.1210/jc.2019-00602

12. Betterle C, Furmaniak J, Sabbadin C, Scaroni C, Presotto F. Type 3 autoimmune polyglandular syndrome (APS-3) or type 3 multiple autoimmune syndrome (MAS-3): an expanding galaxy. *J Endocrinol Invest.* (2023) 46(4):643–65. doi: 10.1007/s40618-022-01994-1

13. Hodé AK, Dédjan H. Autoimmune thyroiditis - track towards autoimmune polyendocrinopathy type III. *Arch Clin cases* (2019) 6(4):109–11. doi: 10.22551/2019.25.0604.10163

14. Crafa A, Calogero AE, Cannarella R, Mongioì LM, Condorelli RA, Greco EA, et al. The burden of hormonal disorders: A worldwide overview with a particular look in Italy. *Front Endocrinol (Lausanne)* (2021) 12:694325. doi: 10.3389/fendo.2021.694325

15. ISTAT. *Resident population at January 1st 2022 in province of Brescia* (2023). ISTAT: Italian National Institute of Statistics. Available at: https://www.istat.it/.

16. Mayer-Davis EJ, Dabelea D, Lawrence JM. Incidence trends of type 1 and type 2 diabetes among youths, 2002-2012. *N Engl J Med* (2017) 376(15):1419–29. doi: 10.1056/NEJMoa1610187

17. Lawrence JM, Divers J, Isom S, Saydah S, Imperatore G, Pihoker C, et al. Trends in prevalence of type 1 and type 2 diabetes in children and adolescents in the US, 2001-2017. *Jama* (2021) 326(8):717–27. doi: 10.1001/jama.2021.11165

18. Van den Driessche A, Eenkhoorn V, Van Gaal L, De Block C. Type 1 diabetes and autoimmune polyglandular syndrome: a clinical review. *Neth J Med* (2009) 67 (11):376–87.

19. Akirov A, Pinhas-Hamiel O. Co-occurrence of type 1 diabetes mellitus and celiac disease. *World J Diabetes.* (2015) 6(5):707–14. doi: 10.4239/wjd.v6.i5.707

20. Li M, Zhu Y, Wei J, Chen L, Chen S, Lai D. The global prevalence of premature ovarian insufficiency: a systematic review and meta-analysis. *Climacteric* (2022), 1–8. doi: 10.1080/13697137.2022.2153033

21. Leite-Silva P, Bedone A, AM P-N, JV C, Costa-Paiva L. Factors associated with bone density in young women with karyotypically normal spontaneous premature ovarian failure. *Arch Gynecol Obstet* (2009) 280(2):177–81. doi: 10.1007/s00404-008-0881-3

22. Pouillès JM, Trémollières F, Bonneu M, Ribot C. Influence of early age at menopause on vertebral bone mass. *J Bone Miner Res* (1994) 9(3):311–5. doi: 10.1002/jbmr.5650090304

23. Popat VB, Calis KA, Vanderhoof VH, Cizza G, Reynolds JC, Sebring N, et al. Bone mineral density in estrogen-deficient young women. *J Clin Endocrinol Metab* (2009) 94(7):2277–83. doi: 10.1210/jc.2008-1878

24. Atsma F, ML B, DE G, van der Schouw YT. Postmenopausal status and early menopause as independent risk factors for cardiovascular disease: a meta-analysis. *Menopause* (2006) 13(2):265–79. doi: 10.1097/01.gme.0000218683.97338.ea

25. Groff AA, Covington SN, Halverson LR, Fitzgerald OR, Vanderhoof V, Calis K, et al. Assessing the emotional needs of women with spontaneous premature ovarian failure. *Fertil Steril* (2005) 83(6):1734–41. doi: 10.1016/j.fertnstert.2004.11.067

26. Jansson-Knodell CL, Hujoel IA, West CP, Taneja V, Prokop LJ, Rubio-Tapia A, et al. Sex difference in celiac disease in undiagnosed populations: A systematic review and meta-analysis. *Clin Gastroenterol Hepatol* (2019) 17(10):1954–1968.e13. doi: 10.1016/j.cgh.2018.11.013

27. Crone J, Rami B, WD H, Granditsch G, Schober E. Prevalence of celiac disease and follow-up of EMA in children and adolescents with type 1 diabetes mellitus. *J Pediatr Gastroenterol Nutr* (2003) 37(1):67–71. doi: 10.1097/00005176-200307000-00011

28. Schuppan D, Hahn EG. Celiac disease and its link to type 1 diabetes mellitus. *J Pediatr Endocrinol Metab* (2001) 14 Suppl 1:597–605. doi: 10.1515/jpem.2001.14.s1.597

29. Cronin CC, Feighery A, JB F, Liddy C, Shanahan F, Feighery C. High prevalence of celiac disease among patients with insulin-dependent (type I) diabetes mellitus. *Am J Gastroenterol* (1997) 92(12):2210–2.

30. Hagopian W, Lee HS, Liu E, Rewers M, She JX, Ziegler AG, et al. Co-occurrence of type 1 diabetes and celiac disease autoimmunity. *Pediatrics* (2017) 140(5). doi: 10.1542/peds.2017-1305

31. Bechi Genzano C, Bezzecchi E, Carnovale D, Mandelli A, Morotti E, Castorani V, et al. Combined unsupervised and semi-automated supervised analysis of flow cytometry data reveals cellular fingerprint associated with newly diagnosed pediatric type 1 diabetes. *Front Immunol* (2022) 13:1026416. doi: 10.3389/fimmu.2022.1026416

32. Rostom A, JA M, Kagnoff MF. American Gastroenterological Association (AGA) Institute technical review on the diagnosis and management of celiac disease. *Gastroenterology* (2006) 131(6):1981–2002. doi: 10.1053/j.gastro.2006.10.004

33. Meddings JB, Jarand J, SJ U, Hardin J, Gall DG. Increased gastrointestinal permeability is an early lesion in the spontaneously diabetic BB rat. *Am J Physiol* (1999) 276(4):G951–7. doi: 10.1152/ajpgi.1999.276.4.G951

34. Funda DP, Kaas A, Bock T, Tlaskalová-Hogenová H, Buschard K. Gluten-free diet prevents diabetes in NOD mice. *Diabetes Metab Res Rev* (1999) 15(5):323–7. doi: 10.1002/(sici)1520-7560(199909/10)15:5<323::aid-dmrr53>3.0.co;2-p

35. Galli-Tsinopoulou A, Nousia-Arvanitakis S, Dracoulacos D, Xefteri M, Karamouzis M. Autoantibodies predicting diabetes mellitus type I in celiac disease. *Horm Res* (1999) 52(3):119–24. doi: 10.1159/000023447

36. Seissler J, Schott M, Boms S, Wohlrab U, Ostendorf B, Morgenthaler NG, et al. Autoantibodies to human tissue transgutaminase identify silent coeliac disease in Type I diabetes. *Diabetologia* (1999) 42(12):1440–1. doi: 10.1007/s001250051317

37. Saukkonen T, Savilahti E, Reijonen H, Ilonen J, Tuomilehto-Wolf E, Akerblom HK. Coeliac disease: frequent occurrence after clinical onset of insulin-dependent diabetes mellitus. *Childhood Diabetes Finland Study Group Diabetes Med* (1996) 13 (5):464–70. doi: 10.1002/(sici)1096-9136(199605)13:5<464::Aid-dia101>3.0.Co;2-r

38. Davidson A, Diamond B. Autoimmune diseases. *N Engl J Med* (2001) 345 (5):340–50. doi: 10.1056/nejm200108023450506

39. Sloka S. Observations on recent studies showing increased co-occurrence of autoimmune diseases. *J Autoimmun* (2002) 18(3):251–7. doi: 10.1006/jaut.2002.0588

40. Wielosz E, Majdan M, Zychowska I, Jeleniewicz R. Coexistence of five autoimmune diseases: diagnostic and therapeutic difficulties. *Rheumatol Int* (2008) 28(9):919–23. doi: 10.1007/s00296-008-0544-5

41. Myhre AG, Undlien DE, Løvås K, Uhlving S, Nedrebø BG, Fougner KJ, et al. Autoimmune adrenocortical failure in Norway autoantibodies and human leukocyte antigen class II associations related to clinical features. *J Clin Endocrinol Metab* (2002) 87(2):618–23. doi: 10.1210/jcem.87.2.8192

42. Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, Concannon P, et al. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. *Diabetes* (2008) 57(4):1084–92. doi: 10.2337/db07-1331

43. Smigoc Schweiger D, Mendez A, Kunilo Jamnik S, Bratanic N, Bratina N, Battelino T, et al. High-risk genotypes HLA-DR3-DQ2/DR3-DQ2 and DR3-DQ2/DR4-DQ8 in co-occurrence of type 1 diabetes and celiac disease. *Autoimmunity* (2016) 49(4):240–7. doi: 10.3109/08916934.2016.1164144

44. Weinstock C, Matheis N, Barkia S, Haager MC, Janson A, Marković A, et al. Autoimmune polyglandular syndrome type 2 shows the same HLA class II pattern as type 1 diabetes. *Tissue Antigens* (2011) 77(4):317–24. doi: 10.1111/j.1399-0039.2011.01634.x

45. Badenhoop K, Walfish PG, Rau H, Fischer S, Nicolay A, Bogner U, et al. Susceptibility and resistance alleles of human leukocyte antigen (HLA) DQA1 and HLA DQB1 are shared in endocrine autoimmune disease. *J Clin Endocrinol Metab* (1995) 80 (7):2112–7. doi: 10.1210/jcem.80.7.7608264

46. Houcken J, Degenhart C, Bender K, König J, Frommer L, Kahaly GJ. PTPN22 and CTLA-4 polymorphisms are associated with polyglandular autoimmunity. *J Clin Endocrinol Metab* (2018) 103(5):1977–84. doi: 10.1210/jc.2017-02577

47. Fichna M, Żurawek M, Słomiński B, Sumińska M, Czarnywojtek A, Rozwadowska N, et al. Polymorphism in BACH2 gene is a marker of polyglandular autoimmunity. *Endocrine* (2021) 74(1):72–9. doi: 10.1007/s12020-021-02743-9

48. Lenti MV, Rossi CM, Melazzini F, Gastaldi M, Bugatti S, Rotondi M, et al. Seronegative autoimmune diseases: A challenging diagnosis. *Autoimmun Rev* (2022) 21 (9):103143. doi: 10.1016/j.autrev.2022.103143

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 28 November 2023
DOI 10.3389/fendo.2023.1229659



OPEN ACCESS

EDITED BY
Di Liu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Chunheng Mo,
Sichuan University, China
Xiaoni Meng,
Capital Medical University, China
Alessandro Maloberti
University of Milano Bicocca, Italy

*CORRESPONDENCE
Xue-Wen Ren
rxw2007@126.com
Qing-Hong Zhang
z_qinghong@aliyun.com

†These authors have contributed equally to this work

RECEIVED 26 May 2023
ACCEPTED 23 October 2023
PUBLISHED 28 November 2023

CITATION
Ren X-W, Chen K, Wu J, Yang Z-L, Ji T and Zhang Q-H (2023) Distinctive biochemistry profiles associated with hyperuricemia between Tibetans and Hans in China. *Front. Endocrinol.* 14:1229659. doi: 10.3389/fendo.2023.1229659



# Distinctive biochemistry profiles associated with hyperuricemia between Tibetans and Hans in China

Xue-Wen Ren[1*†], Kang Chen[2†], Jue Wu[3], Zhang-Lin Yang[1], Tao Ji[1] and Qing-Hong Zhang[4*]

[1]Department of Emergency Medicine, First Medical Center of General Hospital of Chinese People's Liberation Army, Beijing, China, [2]Department of Endocrinology, First Medical Center of General Hospital of Chinese People's Liberation Army, Beijing, China, [3]Translational Medicine Research Center, General Hospital of Chinese People's Liberation Army, Beijing, China, [4]Trauma Repair and Tissue Regeneration Center, Department of Medical Innovation Study, General Hospital of Chinese People's Liberation Army, Beijing, China

**Purpose:** We sought to identify distinct risk factors for hyperuricemia in native Tibetan and immigrant Han populations in Tibet, China.

**Methods:** Three cohorts of male participants aged between 20 and 40 years were enrolled in this study. Biochemical parameters including serum uric acid (UA), fasting plasma glucose, insulin, lactate dehydrogenase (LDH), thyroxin, blood cell count, aminotransferase, and lipid profiles were analyzed. The association of risk factors with UA levels was evaluated using a multivariable line regression model. The effect of UA level on the biochemical parameters between the Hans and Tibetans was evaluated by two-way ANOVA.

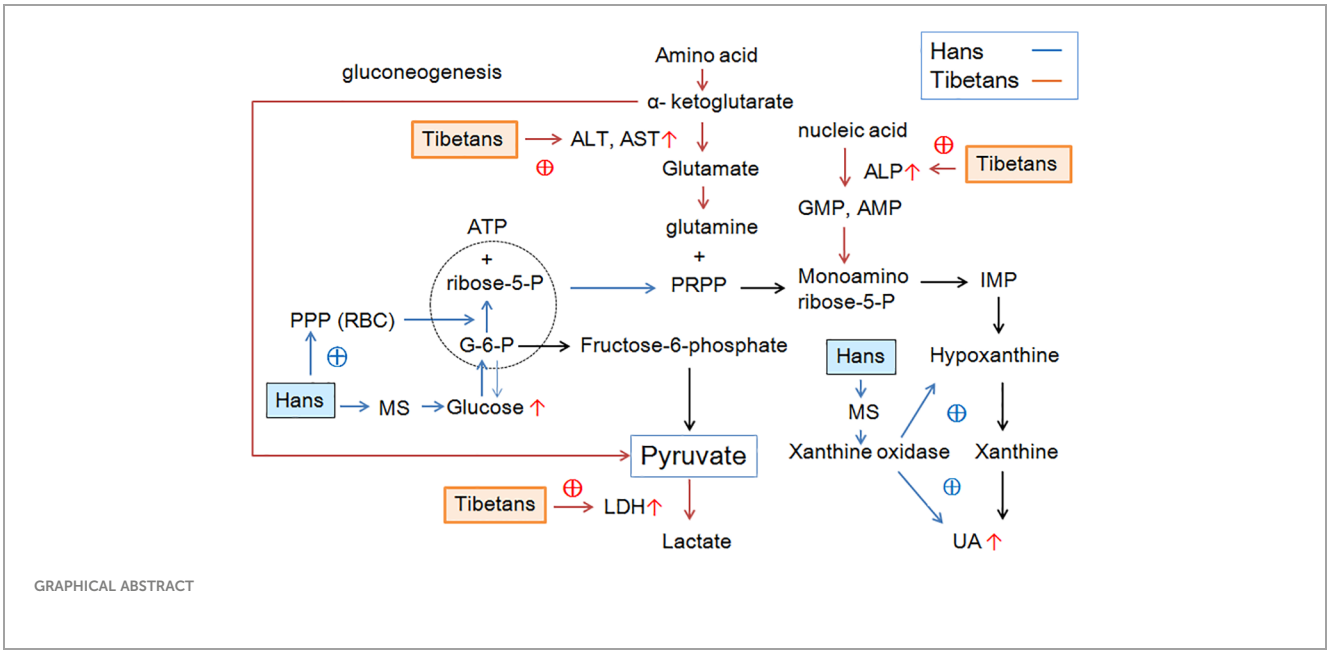
**Results:** The prevalence of hyperuricemia (≥420 μmol/L) was 24.8% (62/250) in the Hans, similar to 23.8% (29/136) in the Tibetans. In the regression analysis, the risk factors that were significantly associated with UA in Hans did not apply to Tibetans. Tibetans had higher fasting insulin ($P<0.05$) and LDH ($P<0.01$) levels, in contrast with lower levels of triglycerides ($P<0.05$), total cholesterol ($P<0.01$), and low-density lipoprotein-cholesterol ($P<0.01$) than Hans in normal UA populations. Biochemistry analysis revealed lower albumin levels ($P<0.001$) and higher levels of all aminotransaminase and especially alkaline phosphatase ($P<0.01$) in Tibetans than in Hans in both populations. Compared with Hans, Tibetans had lower serum levels of urea, creatinine, and electrolytes in the normal UA population, which were further exacerbated in the high UA population. Tibetans had comparable white blood cell counts as Hans in both normal and high UA populations. In contrast, the red blood cell count and hemoglobin concentration were much lower in Tibetans than in Hans under high UA conditions.

10.3389/fendo.2023.1229659

**Conclusions:** The distinctive biochemistry between Tibetans and Hans may underlie the different etiologies of hyperuricemia in Tibet, China.

KEYWORDS

Tibetan, hyperuricemia, hypoxia, glycolysis, aminotransferase, lipid metabolism, red blood cell




GRAPHICAL ABSTRACT

## Introduction

Uric acid (UA) is the end product of purine metabolism and plays a key role in the pathogenesis of gout and other diseases including diabetes, hypertension, and chronic kidney disease. In addition, hyperuricemia is significantly associated with the prevalence of metabolic syndrome (MS) (1, 2). The prevailing view is that the prevalence of hyperuricemia in Chinese Tibet is much higher than that in other parts of China, except for one study that reported a relatively lower prevalence of gout (0.30%) and hyperuricemia (1.83%) in Naqu, Chinese Tibet (3). However, the prevalence of hyperuricemia in immigrants in Chinese Tibetan region was 37.2% in Ganzi (4), 54.2% in Shannan (5), and 40.7% in general Tibetans region (1). In contrast, the prevalence of hyperuricemia was 13.3% in inland China (6), 6.4% in Chinese middle-aged and older adults (7), 10.2% in Chinese rural areas (8), and 15.4% in Hans from northwest China (9). In two Italian population, the prevalence of hyperuricemia (>7.0 mg/dL) in men was 12.9% (56/435) (10) and 7.3%, respectively (11). Regarding ethnicity, the prevalence of hyperuricemia was much higher in Tibetans than in Hans (58.8% *vs.* 28.4%, $P<0.001$) in the same Tibetan region (4). Several factors have been suggested to contribute to the high prevalence of hyperuricemia in the Tibetan Plateau, including MS components (1), ethnicity, dietary habits, hypoxic environment, and gene polymorphisms (12).

Nowadays, with the development of economy and tourism, more inlanders would settle in Chinese Tibet area. The above surveys raised a great concern that Han inlanders accumulated increasing levels of UA if they assimilated into the high altitude as Tibetan highlanders. It has been reported that the clinical indices of Hans were increasingly similar to those of Tibetans with their plateau living (13). Tibetans have undergone natural selection toward a phenotype of exceptional tolerance to hypobaric hypoxia compared to non-Tibetans living at the same altitude (14). Adaptation to acute hypoxia consists of a variety of physiological, metabolic, and molecular changes, such as increased uptake and oxidation of circulatory glucose during exercise (15), increased circulating nitric oxide metabolites (16) to increase vasodilatation and blood flow (16), a shift from aerobic to anaerobic metabolism that favors glycolytic over fatty acid energy supply (17, 18), an increase in muscle glucose toward the pentose phosphate pathway (PPP) (19) to improve muscle energetic performance (18), and protection against oxidative stress (18) with decreased hemoglobin concentration [Hb] (20). The advantageous haplotype of peroxisome proliferator-activated receptor alpha (PPARA) is associated with a lower capacity for fatty acid oxidation in

skeletal muscle in Tibetans (21) and in high-altitude Sherpas (18). Adaptation to chronic hypoxia also involves relative hypometabolism in the brain to minimize the impact of oxygen limitation (22). Therefore, it is necessary to elucidate the different adaptive mechanisms underlying hyperuricemia between immigrant Hans and native Tibetans on the Chinese Tibetan Plateau.

A high-altitude environment features sustained hypobaric hypoxia (23). Acute hypoxia directly enhances UA production and secretion. The phosphorylation of critical enzymes for UA production, xanthine dehydrogenase/xanthine oxidase (XO), was greatly increased (50-fold) in response to acute hypoxia in rat pulmonary microvascular endothelial cells (24). Adipose tissue has abundant expression and activity of xanthine oxidoreductase (XOR). Adipocyte UA secretion increases under hypoxia (25). Furthermore, it was demonstrated that hypoxia diminishes adenosine triphosphate (ATP) utilization by downregulating the activity of Na-K-ATPase in proximal renal tubular epithelial cells (26), limiting renal filtration and excretion ability in rats (27). These findings emphasize that exposure to a hypobaric hypoxic environment may play a crucial role in the pathogenesis of hyperuricemia in the Tibetan Plateau. The discrepancy in evolutionary adaptation toward the cruel environment between the Hans and Tibetans may underlie the unique etiology of hyperuricemia in Tibetans.

However, most of the related studies in Tibetans were on middle-aged or old populations who had already assimilated to high altitudes for many years and may have developed other confounding diseases. Considering that hypoxic research on healthy individuals at high altitudes may be translated into hypoxemic critically ill patients in a hospital setting (14), we carried out a thorough survey of three cohorts of populations aged 20-40 years and compared the biochemical profiles between native Tibetans and immigrant Hans on the Chinese Tibetan Plateau. We aimed to identify the distinctive mechanism underlying the high incidence of hyperuricemia in highland Tibetans and inland Hans acclimatizing to high altitudes.

# Methods

## Study participants

All study protocols were approved by the ethics committee of the Chinese PLA General Hospital (approval identifier S2021-016-01) and were in accordance with established national and institutional ethical guidelines. The study was clearly described to all participants who signed informed consent forms before the collection of blood and personal information. Healthy adults aged 18-60 years old who had lived in the Tibetan region for > 1 year were included in the study. Participants with major operation, tumor, severe lung, heart, digestive, or endocrine diseases were excluded.

## Survey method and data collection

A comprehensive questionnaire, including questions on demographics, medical history, and lifestyle risk factors, was administered by the staff at local health stations according to a standard protocol. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured using a standardized automatic electronic sphygmomanometer (Omron HEM-770A). Body weight (BW) and height were measured, and body mass index (BMI) was calculated as weight/height2 (kg/m2).

## Sample collection

From April 2018 to October 2022, we recruited three groups of healthy male adults (18-60 years old) from both the Han and Tibetan populations in the Chinese Tibetan region (Table 1). Routine physical examinations were conducted in three counties at three altitudes: Lhasa (altitude: 3670-3835 m), Nyingchi (altitude: approximately 2900 m), and Naqu (altitude:4298-4352 m). Group A (n=149) covered suburban Lhasa and Nyingchi, including teachers, students, soldiers, workers, and a few farmers. A standard questionnaire, including information on diet, exercise, altitude, and smoking, was administered. Plasma samples from Group A transported and subjected to thorough biochemical analysis in Beijing, China. Group B (n=226) covered only urban Lhasa and mainly consisted of civil servants, with a few items of biochemistry and routine blood cell counts analyzed immediately on site. Group C (n=111) recruited young adults in Tibetan Naqu, with routine blood cell counts and a few biochemical parameters analyzed immediately at the local station (Table 1). Sampling was conducted between April and October to avoid seasonal variation. To avoid possible systematic errors, doctors (XWR) instructed local medical staff and monitored the data collection and detection procedures.

## Collection of blood and measurements

Fasting venous blood (8 ml) was collected in EDTA-K2 tubes. Hematological parameters were determined immediately using an automated hematology analyzer (Sysmex KX-21, Japan). Blood samples were separated by centrifugation at 4,000 rpm for 10 min within 4 h and then stored at -80°C (Group A) or subjected to automatic analysis of a few biochemical indicators (Hitachi 7180, Japan) immediately in Tibetan Peoples' Hospital (Group B) or Naqu Peoples' Hospital (Group C).

Plasma samples in Group A were transported and assayed in the Department of Laboratory Medicine of Chinese PLA General Hospital (Beijing) (Group A), including UA, glucose, lactate dehydrogenase (LDH), albumin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), γ-glutamyl transferase (GGT), total bilirubin (TBIL), and direct bilirubin (DBIL); renal

TABLE 1 Charactteristic of the three populations in high-altitude regions of Tibet, China.

| Population | Time | Demography | Ethnicity | | *P* value |
|---|---|---|---|---|---|
| | | | Hans | Tibetan | |
| Group A (n=149) | Apr, 2018 | n | 80 | 69 | |
| | | Age (year), median (IQR) | 25 (21-30) | 29 (18-41) | 0.032 |
| | | Location | Lshsa and Nyingchi,Tibetan | | |
| | | Livelihood | teachers, student, soldiers, worker,farmer,herdsman | | |
| | | Altitude | 3670-3835 m, 2900 m | | |
| | | Sample analysis | PLA General Hospital, Beijing | | |
| | | Hyperuricemia, n (%) | 50 (62.5) | 24 (34.8) | |
| Group B (n=226) | Jan-Dec, 2018 | n | 70 | 156 | |
| | | Age, median (IQR) | 33 (24-47) | 42 (33-52) | 0.004 |
| | | Location | Lshsa,Tibetan | | |
| | | Livelihood | civil servant | | |
| | | Altitude | 3670-3835 m | | |
| | | Sample analysis | Lshsa Peoples' Hospital, Lshsa, Tibet | | |
| | | Hyperuricemia, n (%) | 31 (44.3) | 51 (32.7) | |
| Group C (n=111) | Oct, 2018 | n | 46 | 65 | |
| | | Age, median (IQR) | 20 (19-22) | 19 (18-20) | 0.020 |
| | | Location | Naqu,Tibetan | | |
| | | Livelihood | student,worker,farmer,herdsman | | |
| | | Altitude | 4298-4352 m | | |
| | | Sample analysis | Naqu Peoples' Hospital, Naqu, Tibet | | |
| | | Hyperuricemia, n (%) | 16 (34.8) | 13 (20.0) | |

IQR, interquartile range; PLA, Peoples Republic of Chima.

parameters, including creatinine, urea, and cystatin C; and lipid parameters, including triglyceride (TRIG), total cholesterol (TCHOL), high-density lipoprotein cholesterol (HDL-CH), low-density lipoprotein- cholesterol (LDL-CH), and electrolytes, using an automatic biochemical immunity analyzer (Cobas 8000, Roche, USA). The remaining plasma (1.5 ml) was used for hormone assays. The thyroxin concentration was determined using an automated chemiluminescence immunoassay analyzer (ADVIA Centaur®XP, Siemens, Germany). Insulin was detected using a chemiluminescence immunoassay analyzer (Cobas e601; Roche, Switzerland). Tissue nonspecific alkaline phosphatase (ALP) was detected using colorimetry (Roche Diagnostic GMBH, Mannheim, Germany).

## Definitions of hyperuricemia

Considering that the males predominate over the females in the incidence of hyperuricemia (1, 4), because of the limitation of the page space and concision of the study, we only present the male data this time. Hyperuricemia in males was defined as a fasting serum UA level≥420 μmol/L according to Chinese guidelines (28).

## Statistical analyses

The normality of the variables was analyzed using the "Shapiro-Wilk test." The mean and standard deviation were used to describe variables that met the normal distribution, and the median and interquartile (IQR) distance were used to describe variables that did not meet the normal distribution. GraphPad Prism 9.0 (GraphPad Software, USA) was used for data analysis. The two ethnic populations in Group A (Lhasa and Nyingchi) were subjected to multiple line regression analysis to estimate the associations between the major indices and UA levels, including age, altitude, BMI, and biochemical parameters. Where the dependent variable was affected by UA level and ethnicity, the data were analyzed using two-way ANOVA. When the main effect or interaction was significant, *post-hoc* analyses using the Bonferroni correction were performed. Differences between subjects were analyzed using the Chi-squared test or Fisher's exact test for categorical data.

Group B (Lhasa) was subjected to the calculation of the correlation coefficients ($r$) between the red blood cell (RBC) profile and UA level using Pearson's analysis. A two-sided $P<0.05$ was considered statistically significant. Multivariate line regression

was used to analyze the associations between RBCs indices and UA levels in the two populations.

Group C (Naqu) was used for the comparison of demography, RBCs profile, and blood cell count between Hans and Tibetans. The Mann–Whitney U test or unpaired *t* test was used for intergroup comparisons.

# Results

## Characteristics of the three cohorts of populations

The three groups of populations covered the ages 20-30 (Group A), 30-40 (Group B), and 19-20 (Group C) years old. Tibetans were older than Hans in the former two populations, whereas they were younger than Hans in Group C (Table 1).

The survey of Group A was based on a sample drawn from a larger population. We first enrolled 250 Hans and 136 native Tibetans for general physical examinations. The incidence of hyperuricemia was 24.8% in the total Han population (62/250) with a median (IQR) age (years) of 24 (21-28) and 23.8% in the total native Tibetan population (29/136) with a median (IQR) age of 31 (19-46). From the above population, we matched samples with similar ages and backgrounds into Group A for the following thorough biochemistry analysis, in which the Hans (n=85) were still younger than the Tibetans (years) (n=64) [25 (21-30) *vs.* 29 (18-41), *P*=0.032]. However, there was no difference in age between the subgroups (Table 2). The median (IQR) length of residence (years) for the Hans in Tibet was similar between the subgroups with high UA and normal UA levels [3.3 (1.3-7.8) *vs.* 3.7 (1.3-6.3), *P*=0.53].

Intriguingly, in the normal UA populations of Group A, the Tibetans had heavier BW in contrast with shorter height than the Hans, leading to higher BMI than the Hans (Table 2). The Tibetans had different professional compositions and dietary habits, but more alcohol consumption and were more likely to develop hypertension in both normal and high UA populations, and less exercise than Hans in normal UA populations.

In Group B, the Hans (n=70) were also younger than Tibetans (n=156) [33 (24-47) *vs.* 42 (33-52) (year), *P*=0.004]. The incidence of hyperuricemia was 44.3% in the Hans and 32.7% in the Tibetan.

In Group C, the Hans (n=46) were slightly older than native Tibetans (n=65) (*P*=0.019). The Hans had a higher height and heavier BW than the Tibetans, resulting in an identical BMI. The incidence of hyperuricemia was 40.0% in the Hans and 18.2% in the Tibetan. Both populations had similar blood pressure and heart

TABLE 2 Comparison of the characteristic between Tibetans and Hans in Tibet, China (Group A).

| Variates | Hans | | | Tibetan | | | P value (Hans vs.Tibetan) | |
|---|---|---|---|---|---|---|---|---|
| | normal (n=30) | high (n=55) | P value | normal (n=37) | high (n=27) | P value | Normal | High |
| age (year; median, IQR) | 23.0 (21.0-35.0) | 26.0 (21.0-30.0) | 0.11 | 20.0 (17.0-41.5) | 32.0 (25.0-46.0) | 0.61 | 0.09 | 0.71 |
| Height (cm) | 173.5 (168.0-177.0) | 173.0 (169.0-176.0) | 0.60 | 168.5(160.0-173.8) | 175.0(169.3-180.0) | 0.002 | 0.011 | 0.31 |
| Body weight (kg) | 67.0 (58.8-122.5) | 70.0 (62.0-88.5) | 0.59 | 100.0 (92.00-140.0) | 90.0 (80.0-103.0) | 0.21 | 0.017 | 0.11 |
| BMI (median, IQR) | 22.50 (20.0-44.0) | 23.5 (21.0-29.0) | 0.44 | 37.5(32.0-47.8) | 29.5 (25.3-35.8) | 0.13 | 0.014 | 0.10 |
| Length of stay in Tibetan | 3.7 (1.3-6.3) | 3.3 (1.3-7.8) | 0.57 | 18.8 (16.5-40.6) | 31.3 (17.4-42.0) | 0.16 | <0.0001 | <0.0001 |
| altitude (1<3000m,2:3000-4000m,3≥4000m) | | | 0.38 | | | 0.53 | 0.06 | 0.45 |
| 1 | 3 | 2 | | 2 | 1 | | | |
| 2 | 27 | 52 | | 29 | 24 | | | |
| 3 | 0 | 1 | | 6 | 2 | | | |
| Profession (1:teacher,2:student,3:army,4: civil servant,5:worker, 6:farmer) | | | 0.17 | | | 0.005 | <0.0001 | <0.0001 |
| 1 | 6 | 7 | | 5 | 4 | | | |
| 2 | 0 | 0 | | 23 | 7 | | | |
| 3 | 11 | 37 | | 0 | 2 | | | |
| 4 | 2 | 4 | | 4 | 4 | | | |
| 5 | 7 | 5 | | 0 | 7 | | | |

*(Continued)*

TABLE 2 Continued

| Variates | Hans | | | Tibetan | | | P value (Hans vs.Tibetan) | |
|---|---|---|---|---|---|---|---|---|
| | normal (n=30) | high (n=55) | P value | normal (n=37) | high (n=27) | P value | Normal | High |
| 6 | 4 | 2 | | 5 | 3 | | | |
| diet (1:meat,2:vegetable,3:balanced) | | | 0.41 | | | 0.08 | 0.0001 | 0.016 |
| 1 | 4 | 10 | | 18 | 12 | | | |
| 2 | 4 | 3 | | 11 | 3 | | | |
| 3 | 22 | 42 | | 8 | 12 | | | |
| Alcohol(1:never,2:seldom,3:often) | | | 0.73 | | | 0.63 | 0.015 | 0.14 |
| 1 | 13 | 19 | | 5 | 6 | | | |
| 2 | 16 | 34 | | 27 | 17 | | | |
| 3 | 1 | 2 | | 5 | 4 | | | |
| water intake(1<1000ml,2:1000-2000ml;3≥2000m) | | | 0.45 | | | 0.40 | 0.09 | 0.13 |
| 1 | 10 | 20 | | 5 | 5 | | | |
| 2 | 13 | 28 | | 25 | 20 | | | |
| 3 | 7 | 7 | | 7 | 2 | | | |
| exercise (1:never,2:seldom,3:often) | | | 0.16 | | | 0.34 | 0.003 | 0.06 |
| 1 | 2 | 1 | | 1 | 3 | | | |
| 2 | 11 | 31 | | 29 | 18 | | | |
| 3 | 17 | 23 | | 7 | 6 | | | |
| Gout (n,%) | 0 | 5 | 0.16 | 2 | 2 | 0.74 | 0.20 | 0.80 |
| Hypertension (n,%) | 0 | 1 | 0.46 | 6 | 4 | 0.88 | 0.021 | 0.021 |
| Pulmonary arterial hypertension(n,%) | 0 | 0 | | 0 | 0 | | | |
| Biliary/renal calculus(n,%) | 0 | 1 | 0.46 | 1 | 0 | 0.39 | 0.36 | 0.48 |
| Gastropathy (n,%) | 5 | 3 | 0.09 | 8 | 2 | 0.12 | 0.61 | 0.73 |
| Hepatitis (n,%) | 0 | 1 | 0.46 | 1 | 0 | 0.39 | 0.36 | 0.48 |
| Tuberculosis (n,%) | 0 | 1 | 0.46 | 0 | 2 | 0.09 | - | 0.21 |

rates (Table 3). The Tibetans differed from the Hans in an extremely long Tibetan dwelling time (year) [19.0 (18.0-20.0) *vs.* 2.0 (1.0-4.0), $P<0.001$], more meat in diet and water intake, more tap water over purified water, and less exercise (all $P<0.001$). Nonetheless, Tibetans had fewer signs of acute and chronic altitude sickness.

## Distinct risk factors associated with serum UA between the Hans and the Tibetans

Consistent with previous reports (1, 29), almost all risk factors associated with MS were significantly associated with UA levels in the Han population, including age, altitude, BMI, aminotransaminase, the product of heme catabolism (bilirubin), cholesterol transporter (HDL), a biomarker for glomerular filtration rate (cystatin C), glucose, and insulin levels (Table 4). Meanwhile, factors that were not significantly associated with serum UA levels in Hans were lipid metabolism and transportation (total TRIG, TCHOL, and LDL-C3), renal function (creatinine and urea), metabolism of extracellular nucleotides (ALP), and glycolysis (LDH) (Table 4). Unexpectedly, none of the above factors contributed to UA levels in Tibetans, suggesting a distinct mechanism underlying hyperuricemia in native Tibetans.

## The Tibetans showed heightened insulin resistance and glycolysis compared to the Hans

The normality of biochemical indicators in the three groups was tested (Supplementary 1). We then investigated glycolysis, which may allow ATP to be rapidly generated in hypoxic cells. In Group A, although the glucose levels were similar between the Hans and Tibetans in the normal UA populations, they were increased by

TABLE 3 Comparison of the biochemistry between the Hans and the Tibetans in Naqu, Chinese Tibetan region (Group C).

| Variate | Age (y) | Dwelling time (y) | Height (cm) | Body weight (kg) | BMI | WBC ($10^9$/L) | Eosinophil ($10^9$/L) | Lymphocyte ($10^9$/L) | Neutrophil ($10^9$/L) | Platelet ($10^9$/L) | MPV (fL) | P-LCR% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tibetan, median (IQR) | 19 (18-20) | 19 (18-20) | 164 (161-168) | 56.0 (50.5-22.0) | 20 (19-22) | 7.65 (5.8-9.5) | 0.5 (0.4-0.7) | 1.60 (1.38-2.00) | 5.5 (3.8-7.0) | 239.5 (184.3-259.8) | 9.6 (8.8-10.2) | 0.22 (0.17-0.26) |
| Hans, median (IQR) | 20 (19-22) | 2 (1-4) | 170 (168.3-175) | 60.5 (55.0-67.0) | 21 (19-22) | 8.15 (5.95-10.8) | 0.6 (0.5-0.8) | 1.65 (1.48-2.10) | 5.7 (3.8-7.7) | 215.5 (166.8-256.8) | 9.8 (9.3-11.3) | 0.24 (0.19-0.35) |
| *P* value (*t* test) | 0.019 | <0.001 | <0.001 | <0.001 | 0.40 | 0.57 | 0.55 | 0.59 | 0.65 | 0.32 | 0.036 | 0.038 |
| Variate | RBC ($10^{12}$/L) | SPO2 | Hematocrit (%) | Hemoglobin (g/L) | MCV (fL) | MCH (pg) | MCHC (g/L) | RDW-SD (fL) | RDW-CV (%) | RDW % | Uric acid (μmol/L) | Glucose (mmol/L) |
| Tibetan | 5.57 (0.7) | 86 (84-88) | 155 (139-200) | 0.49 (0.47-0.50) | 89.5 (87.3-91.5) | 32.5 (1.5) | 357 (352-364) | 43.2 (3.5) | 0.140 (0.134-0.143) | 11.2 (10.2-12.6) | 373 (362-382) | 3.70 (0.10) |
| Hans | 6.0 (0.6) | 88 (83-90) | 201 (182.5-220.5) | 0.54 (0.50-0.58) | 89.0 (85.9-92.1) | 32.9 (2.6) | 369 (352-378) | 43.2 (4.1) | 0.138 (0.131-0.143) | 12.1 (10.7-14.2) | 400 (361-489) | 3.31 (0.44) |
| *P* value (*t* test) | 0.010 | 0.71 | 0.018 | 0.004 | 0.73 | 0.38 | 0.76 | 0.94 | 0.93 | 0.40 | 0.37 | 0.22 |
| Variate | Total protein (g/L) | Albumin (g/L) | globulin (g/L) | TBIL (μmol/L) | DBIL (μmol/L) | IBIL (μmol/L) | ALT (U/L) | ALP (U/L) | GGT (U/L) | Urea (mmol/L) | Creatinine (μmol/L) | |
| Tibetan | 82.3 (3.8) | 48.8 (47.6-49.9) | 33.0 (29.6-35.5) | 11.1 (9.4-15.4) | 4.4 | 6.5 (5.4-8.3) | 23 (17-30) | 102 (79-141) | 23 (21-26) | 8.85 (1.95) | 94.0 (14.0) | |
| Hans | 80.2 (4.7) | 49.2 (47.6-50.8) | 30.8 (28.9-33.8) | 14.8 (11.0-22.2) | 4.95 | 9.8 (7.1-14.8) | 24.5 (20.0-40.0) | 93 (73-119) | 19 (17-24) | 8.21 (1.17) | 93.3 (10.2) | |
| *P* value (*t* test) | 0.18 | 0.20 | 0.004 | 0.003 | 0.052 | <0.001 | 0.20 | 0.27 | 0.52 | 0.36 | 0.16 | |

Data are expressed as median (IQR) or median (SD). SBP, Systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; RDW, red cell distribution width; MPV, mean platelet volume; BIL, bilirubin; WBC, white blood cell; RBC, red blood cell; SPO2, blood oxygen saturation; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red distribution width; SD, standard deviation; CV, coefficient of variation; MPV, mean platelet volume; P-LCR, platelet-larger cell ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-glutamyl transferase; TBIL, total bilirubin; DBIL, direct bilirubin.





10.3389/fendo.2023.1229659

TABLE 4 Multivariate line regression analysis of the risk factors associated with serum uric acid levels between Hans and Tibetans in Tibet, China (Group A).

| Ethnic | | Chinese-Tibetans | | Chinese-Hans | | |
|---|---|---|---|---|---|---|
| Variable | OR | 95% CI | P value | OR | 95% CI | P value |
| Age | 1.29 | -10.22 to 2.659 | 0.22 | 4.78 | -8.316 to -3.333 | <0.001 |
| Altitude | 0.15 | -91.17 to 79.90 | 0.89 | 2.54 | 16.02 to 148.2 | 0.017 |
| BMI | 0.28 | -8.118 to 6.295 | 0.79 | 2.84 | 1.912 to 11.79 | 0.008 |
| Albumin | 0.80 | -17.93 to 8.383 | 0.44 | 2.26 | 0.6394 to 13.01 | 0.032 |
| ALT | 0.14 | -3.176 to 2.794 | 0.89 | 2.68 | -6.590 to -0.8812 | 0.012 |
| AST | 0.23 | -6.670 to 8.198 | 0.83 | 2.63 | 1.446 to 11.52 | 0.013 |
| GGT | 0.58 | -2.516 to 1.460 | 0.57 | 2.61 | 0.5106 to 4.235 | 0.014 |
| ALP | 1.21 | -0.7793 to 0.2275 | 0.25 | 0.91 | -1.145 to 0.4411 | 0.37 |
| TBIL | 1.01 | -16.19 to 43.56 | 0.34 | 4.34 | -26.58 to -9.549 | <0.001 |
| DBIL | 0.71 | -104.9 to 53.82 | 0.49 | 4.35 | 28.98 to 80.32 | <0.001 |
| TRIG | 0.21 | -135.4 to 164.4 | 0.83 | 0.56 | -27.31 to 47.94 | 0.58 |
| CHOL | 0.65 | -462.1 to 251.2 | 0.53 | 0.06 | -113.4 to 107.5 | 0.96 |
| HDL-CH | 0.54 | -393.1 to 238.7 | 0.60 | 2.37 | -321.4 to -23.80 | 0.025 |
| LDL-CH | 0.53 | -288.0 to 469.4 | 0.61 | 0.40 | -91.44 to 136.0 | 0.69 |
| Creatinine | 0.44 | -4.761 to 7.131 | 0.67 | 1.40 | -0.4563 to 2.441 | 0.17 |
| Cystatin C | 1.28 | -156.6 to 589.1 | 0.23 | 4.80 | 241.2 to 599.7 | <0.001 |
| UREA | 0.05 | -33.79 to 35.23 | 0.96 | 0.74 | -13.81 to 6.488 | 0.47 |
| Glucose | 0.30 | -44.12 to 33.40 | 0.77 | 2.26 | 3.002 to 59.01 | 0.031 |
| LDH | 0.54 | -1.273 to 0.7681 | 0.60 | 0.33 | -0.3414 to 0.4751 | 0.74 |
| Insulin | 1.09 | -2.001 to 5.896 | 0.30 | 3.14 | -3.128 to -0.6621 | 0.004 |

OR, odds ratio; CI, confidence interval; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-glutamyl transferase; ALP, alkaline phosphatase; TBIL, total bilirubin; DBIL, direct bilirubin; TRIG, triglyceride; TCHOL, total cholesterol; HDL-CH, high-density lipoprotein cholesterol; LDL-CH, low-density lipoprotein cholesterol; LDH, lactate dehydrogenase.

high UA exclusively in the Tibetans. Conversely, both fasting insulin ($P$<0.05) and LDH ($P$<0.01) levels were higher in Tibetans than in Hans in normal UA populations (Figure 1). These results indicate an increased tendency of insulin resistance and subsequent glycolysis in Tibetans relative to Hans.

At higher altitudes, as in Group C, the glucose level was further decreased in the young adults of Hans [3.3 (3.1-3.6) *vs.* 4.2 (3.9-4.5) mmol/L, $P$<0.001] but not in the Tibetans [3.70 (3.6-3.8) *vs.* 4.1 (3.9-4.4) mmol/L, $P$=0.45] compared with the respective ethnic groups in Group A with normal UA levels. However, the glucose levels were comparable between the two ethnicities in Group C) (Table 3).

## The Tibetans had reduced lipid and biliary metabolism compared with the Hans

In both groups of our populations, the level of total cholesterol was 3.21-4.48mmol/L (124-173mg/dl), LDL cholesterol was between 1.85-2.51mmol/L (77-97mg/dl), both were much lower than the total cholesterol (224.0 ± 42.9 mg/dl) and LDL cholesterol (145.3 ± 39.3 mg/dl) from an Italian population, respectively. However, the level of HDL-cholesterol was between 1.00-1.18mmol/L (37-44 mg/dl) and TG was between 0.86-2.24 mmol/L (55-143mg/dl), similar to those from the same Italian population with HDL (55.6 ± 15.6 mg/dl) and TG [96 (69-137) mg/dl], respectively (30).

Surprisingly, in Group A (Figure 2I), the lipid parameters, including TRIG, TCHOL, and LDL-CH, were significantly lower in Tibetans than in Hans, even in the normal UA setting. Both TRIG and TCHOL levels were significantly increased in the high UA population. TBIL and HDL levels were identical between the Hans and Tibetans, except that TBIL was only increased by high UA levels in the Hans. In Group B, at higher altitudes, it was noteworthy that TRIG, total, direct, and indirect bilirubin levels were significantly higher in the Hans than in the Tibetans (Figure 2II). Likewise, in Group C from even higher altitudes of Naqu, Tibetans showed consistently lower levels of total, direct, and indirect bilirubin than Hans (Table 3). Bilirubin is a breakdown product of heme released during RBCs lysis (31). In normal UA populations,

10.3389/fendo.2023.1229659

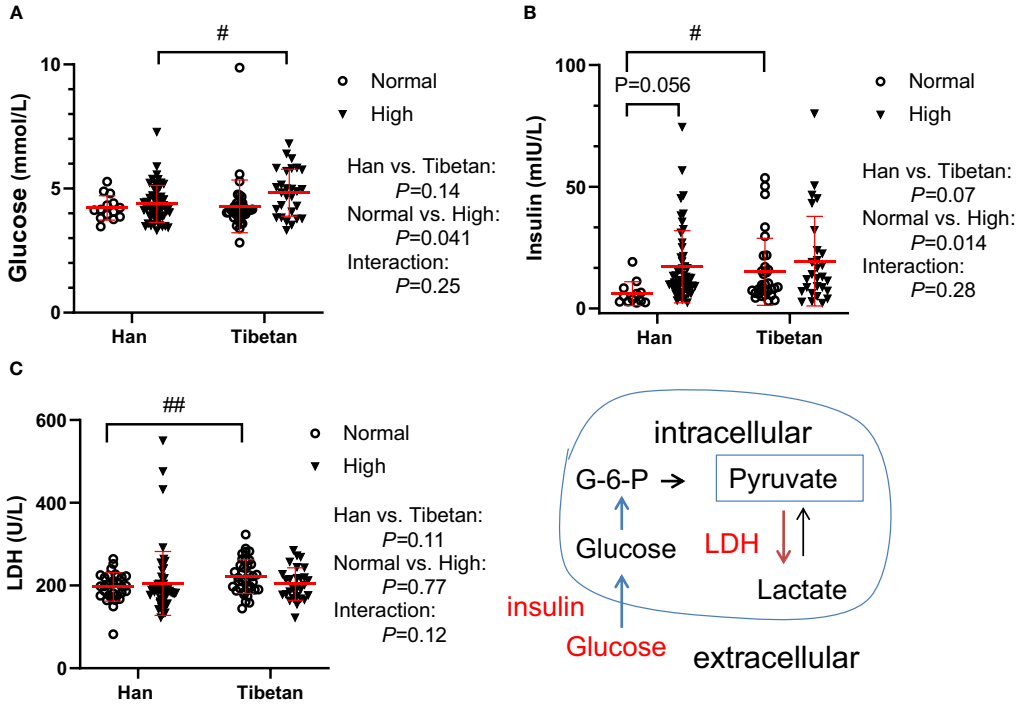



FIGURE 1
Comparison of glucose metabolism between Hans and Tibetans with normal or high uric acid (UA) levels in Group A. LDH, lactate dehydrogenase. #$P<0.05$, ##$P<0.01$ Tibetans *vs.* Hans.

the relatively lower levels of bilirubin in Tibetans relative to Hans indicated reduced RBCs hemolysis in Tibetans (Table 3; Figure 2II).

## Distinct biochemistry biomarkers for protein metabolism between Hans and Tibetans

In Groups A and B, although the total protein levels were similar between the Han and Tibetan populations, they were increased by high UA levels in both populations (Figures 3I, II). Conversely, the albumin levels were elevated by high UA in both ethnic groups in Group A (Figure 3I), but were still lower in the Tibetans than in the Hans, regardless of UA level in both Groups A and B (Figures 3I, II). Biochemistry analysis revealed higher levels of all aminotransaminase and ALP in Tibetans than in Hans in both Group A (ALT, AST, GGT, and ALP) and Group B (ALT and AST). Intriguingly, the levels of aminotransaminase were elevated by high UA exclusively in the Hans (Figures 3IC–E, 3IIC, D). In Group C, Tibetans had higher total protein and globulin levels than Hans, but identical levels of aminotransaminases (Table 3).

## The Tibetans had less severe kidney injury than the Hans

In Group A, UA levels were comparable between the Hans and Tibetans in both the normal and high UA populations (Figure 4IA). Tibetans had lower levels of creatinine and urea than Hans, but cystatin C levels were similar to those of Hans (Figures 4IB–D). All the levels of the above biomarkers were greatly increased by high UA levels in Tibetans. However, the ethnic differences, especially the differences between the normal and high UA groups, were less significant in Group B than in Group A (Figure 4II). Although the levels of urea and creatinine UA were higher in high UA than in normal UA populations in Groups A and B, they were identical between Tibetans and Hans in young adults in Group C (Table 3). As the levels of biomarkers for kidney injury increased with altitude in the Han population, the difference between the two ethnic populations seemed to be attenuated with the increase in altitude.

Creatine kinase (CK) and its MB isoenzyme (CK-MB) are the most commonly used serological biomarkers for the diagnosis of myocardial infarction. In Group A, CK did not differ between Tibetans and Hans regardless of the UA level (Figure S2A). CK-MB was increased by high UA levels exclusively in Tibetans (Figure S2B). N-terminal fragment B-type natriuretic peptide (NT-pro-BNP) is frequently used for the diagnosis of congestive heart failure. NT-proBNP levels are affected by age or the presence of one or several comorbidities, such as chronic renal failure, type 2 diabetes, and acute coronary syndrome (32). In the Hans of Group A, we detected decreased NT-pro-BNP levels in the high UA group compared to the normal UA group (Figure S2C), its significance remains unclear.

## The Tibetans had lower serum levels of thyroxin than the Hans

Regardless of the UA level, thyroid-stimulating hormone (TSH), triiodothyronine (T3), and free T3 levels were not altered

10.3389/fendo.2023.1229659

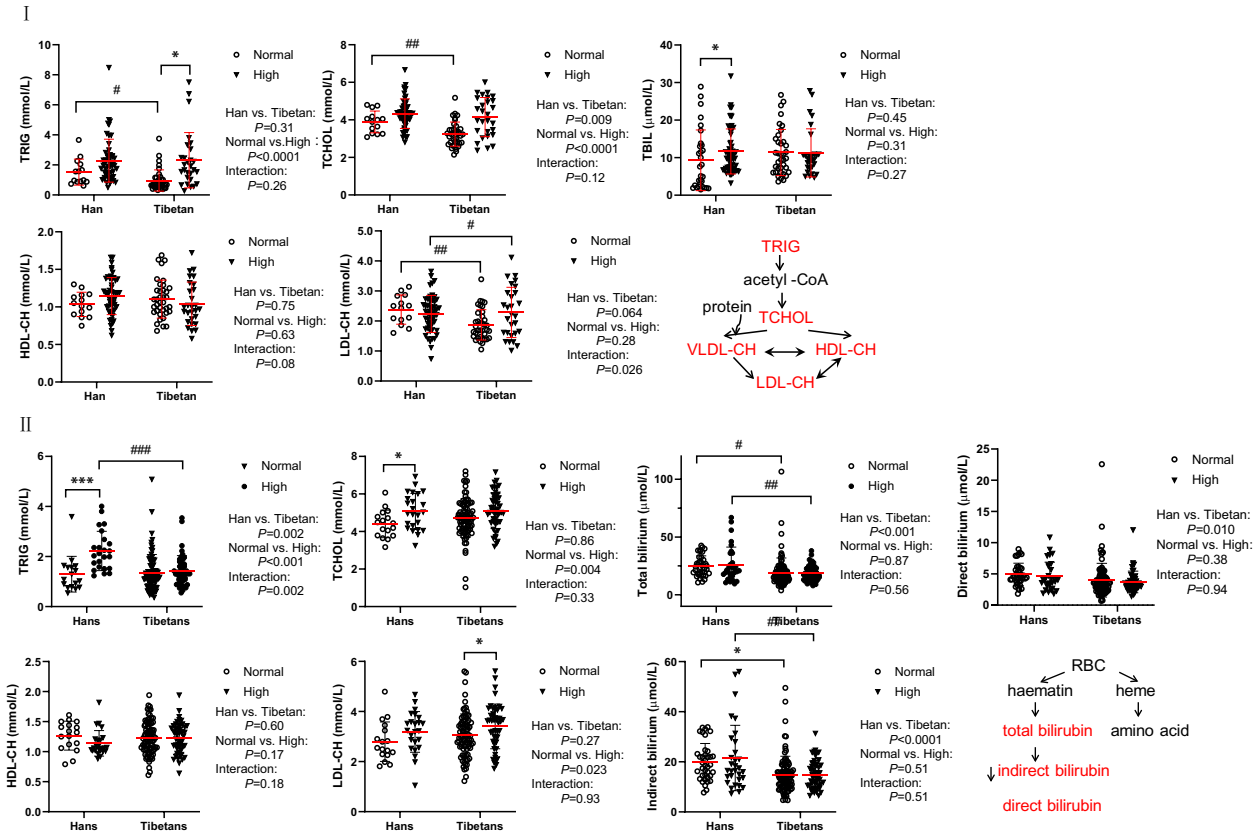



FIGURE 2
Comparison of lipid metabolism and transporters between Hans and Tibetans with normal or high uric acid (UA) levels in Group A (**I**) and Group B (**II**) I: Tibetans showed lower triglyceride (TRIG), total cholesterol (TCHOL), total bilirubin (TBIL), and low-density lipoprotein-cholesterol (LDL-CH) levels than Hans in normal UA populations. Total bilirubin (TBIL) and high-density lipoprotein cholesterol (HDL-CH) were identical between the two populations. II: Tibetans showed lower levels of TRIG, TBIL, and direct and indirect bilirubin than Hans in both normal and high UA settings. TRIG, TCHOL and LDL-CH were increased by hyperuricemia in both populations. TCHOL, HDL-CH, and LDL-CH were similar between the Hans and the Tibetans. *$P$<0.05, ***$P$<0.001, normal UA *vs.* high UA; #$P$<0.05, ##$P$<0.01, ###$P$<0.001, Tibetans *vs.* Hans.

between the Hans and Tibetans in Group A (Figure 5). Although L-thyroxin (T4) levels were similar between the Han and Tibetan populations in normal UA populations, they were lower in Tibetans than in Hans in high UA populations. Similarly, serum free T4 levels were lower in Tibetans than in Hans in high-UA populations. It appears that hyperuricemia exempts Tibetans from T4 synthesis and the release of free T4.

## The Tibetans had lower serum levels of electrolytes than the Hans

In Group A, all serum electrolyte levels were much lower in Tibetans than in Hans in the normal and high UA populations (Figure 6). Interestingly, the electrolyte levels tended to be elevated by high UA levels exclusively in the Hans rather than in the Tibetans. This implies that the proximal convoluted tubule, which is responsible for electrolyte absorption, is more resilient to the dangerous effects of hyperuricemia in Tibetans than it is in Hans.

## Similar circulating hemolytic counts between the Tibetans and the Hans

In Group B, Tibetans had similar counts of blood cells as Hans, except that they had fewer lymphocyte counts but more eosinophil counts than Hans (Figure 7). Tibetans with high UA had increased white blood cell (WBC) counts in comparison with those in the normal UA group. Likewise, in Group C of young adults, Tibetans were not distinguished from Hans in blood cell counts (Table 3). The normality of blood leukocyte parameters in Group C was also tested (Supplementary 1).

## The Tibetans exhibited less erythropoiesis than the Hans under hyperuricemia

In Group B, although both RBC number and [Hb] were beyond the reference ranges (male: RBC 4.3-5.9×$10^{12}$/L, [Hb] 137179 g/L),

10.3389/fendo.2023.1229659

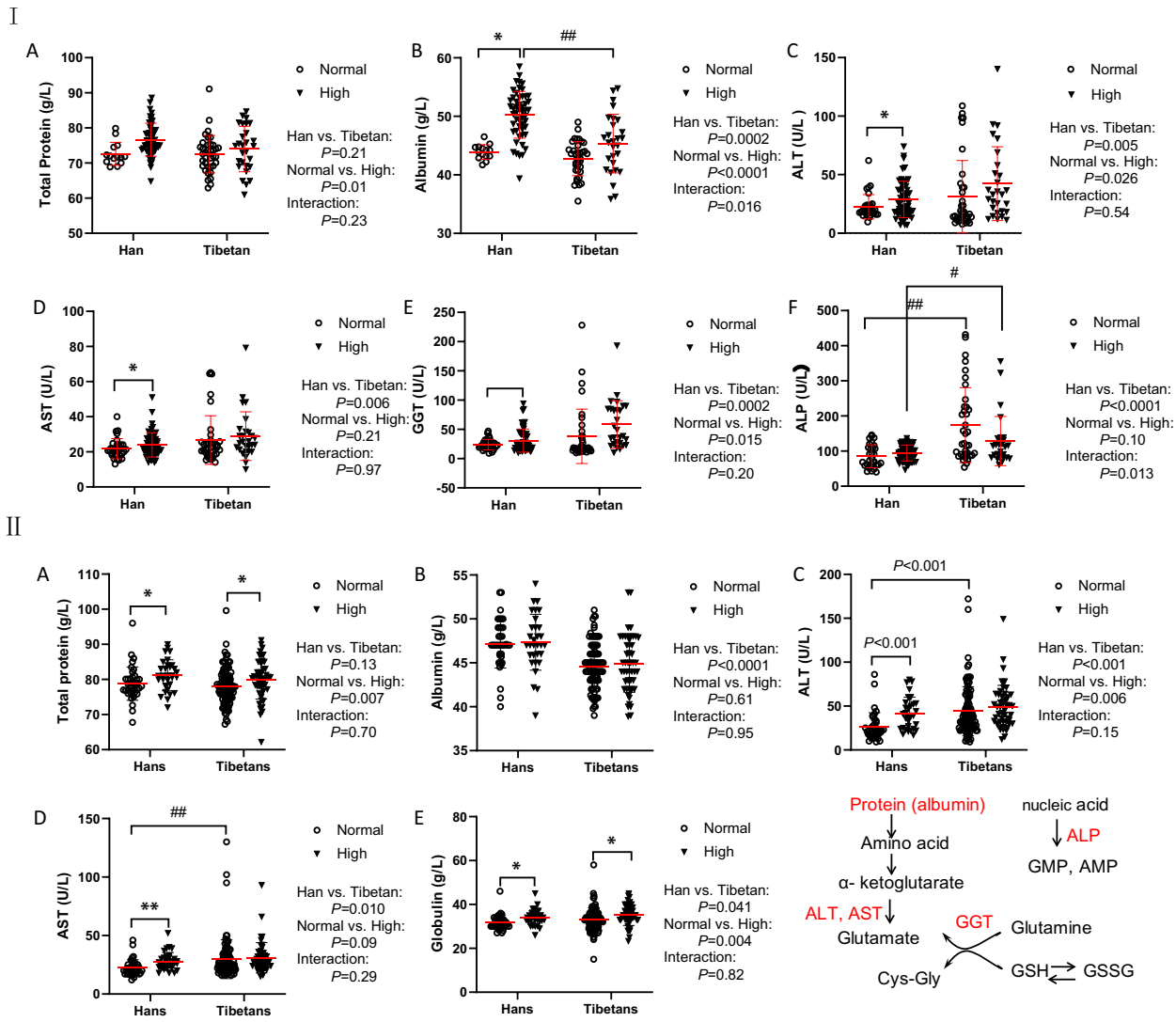



FIGURE 3
Comparisons of the biomarkers for protein metabolism between the Hans and Tibetans with normal or high uric acid (UA) levels in Group A (**I**) and Group B (**II**). The Tibetans exhibited lower levels of albumin but higher levels of aminotransferase than the Hans in both normal and high UA settings in both Group A and Group (B) **P*<0.05, ***P*<0.01, normal UA *vs.* high UA; #*P*<0.05, ##*P*<0.01, Tibetans *vs.* Hans.

they were indistinguishable between the two populations under normal UA conditions. Both were differentially increased between Tibetans and Hans under high UA conditions, resulting in higher levels in Hans than in Tibetans (Figure 8). RBC volume distribution width (RDW) is a conventional biomarker of erythrocyte volume variability and an indicator of homeostasis (33). The RDW standard deviation (SD) and coefficient of variation (CV) were elevated by high UA only in Tibetans (Figure 8).

In Group C, the Hans rather than the Tibetans developed polycythemia with extremely high [Hb] and hematocrit values beyond the normal ranges. In comparison, Tibetans had lower [Hb] values, reduced RDW%, smaller mean platelet volume (MPV) and lower platelet-larger cell ratio (P-LCR) % than Hans (Table 3), indicating greater resilience to severe hypoxia at higher altitudes in Tibet. This implies that severe hypoxia distinguishes the two ethnic populations in the RBC profile, similar to the high UA level in RBC in Figure 8.

Interestingly, in Group B, although the UA levels correlated with RDW-SD and RDW-CV significantly in Tibetans (Table S1), the RBCs profiles were not independently associated with UA levels in either Tibetans or Hans (Table 5). Similar to that in Group A (Table 4), only the glucose level was associated with the UA level in Hans other than Tibetans (Table 5).

## Discussion

Although the prevalence of hyperuricemia is much higher in Tibet than in other places in China, immigrant Hans have accumulated a high incidence of hyperuricemia in Tibet. When

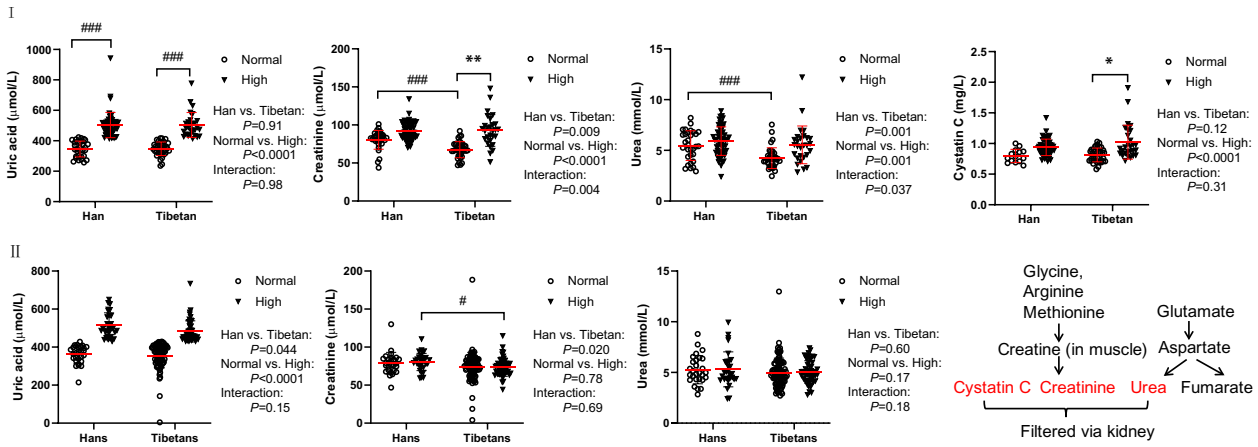



FIGURE 4
Comparisons of the biomarkers for kidney and heart injury between Hans and Tibetans with normal or high uric acid (UA) levels in Group A (**I**) and Group B (**II**). *$P$<0.05, **$P$<0.01, normal UA *vs.* high UA; #$P$<0.05, ###$P$<0.001, Tibetans *vs.* Hans.

we compared the incidence of hyperuricemia between native Tibetans and immigrant Hans at the three altitudes in the Tibetan region, we found a similarly high prevalence of hyperuricemia between the two ethnicities but with distinct biochemical mechanisms. Hyperuricemia is closely associated with obesity and metabolic disturbances such as insulin resistance, dyslipidemia, hypertension, and kidney disease in lowlanders (34) and highlanders in Tibet (1) and Peru (35).

In our study, the high prevalence of hyperuricemia in high-altitude-adapted native Tibetans reached at the similar level as the patients with Acute Coronary Syndrome, e.g., acute Heart Failure (35.8%) (36), which could not be explained simply by traditional metabolic disturbances. First, the factors that were significantly associated with UA levels in the Han population were not applied to Tibetans. Second, Tibetans had higher serum insulin and LDH levels, indicating heightened anaerobic metabolism compared to Hans. Third, Tibetans had increased aminotransferase and ALP activities, suggesting enhanced protein and nucleic acid turnover compared with Hans. Fourth, Tibetans had extremely low serum levels of TRIG, TCHOL, and LDL-CH, implying a lower degree of lipometabolism. Finally, Tibetans had better hypoxic adaptation with a lower degree of polycythemia than Hans. The above biochemical discrepancy between the two populations may be distinctively associated with hyperuricemia between Tibetans and Hans.

## Purine metabolism in hypoxia

UA is primarily produced in the liver as the end product of exogenous and endogenous purine metabolism, covering the catabolism, *de novo* synthesis, and salvage pathways (37). The intake of fructose or a purine-rich diet, ATP depletion induced by

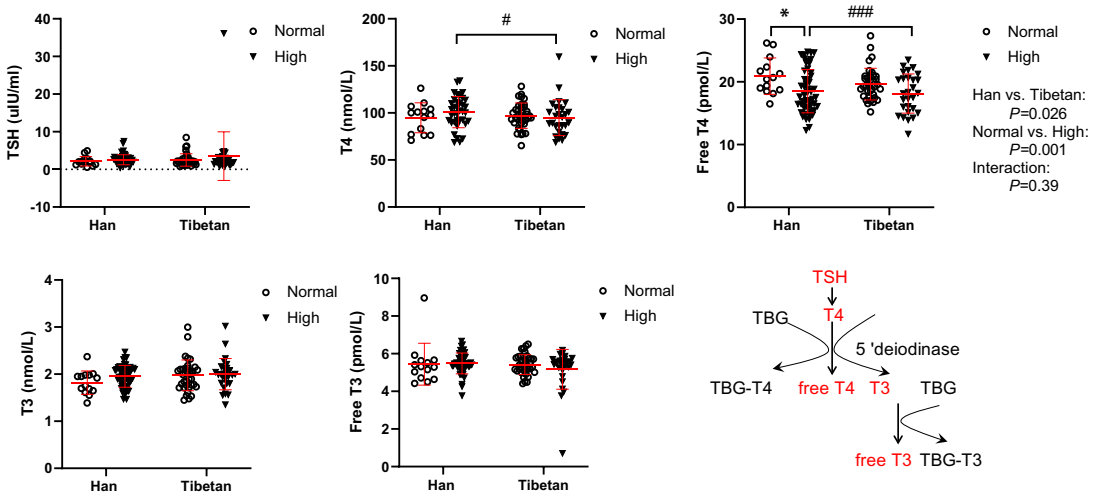



FIGURE 5
Comparisons of serum thyroxin levels between Hans and Tibetans with normal or high uric acid (UA) levels in Group A. *$P$<0.05, normal UA *vs.* high UA; #$P$<0.05, ###$P$<0.001, Tibetans *vs.* Hans. TBG, thyroxin-binding globulin.

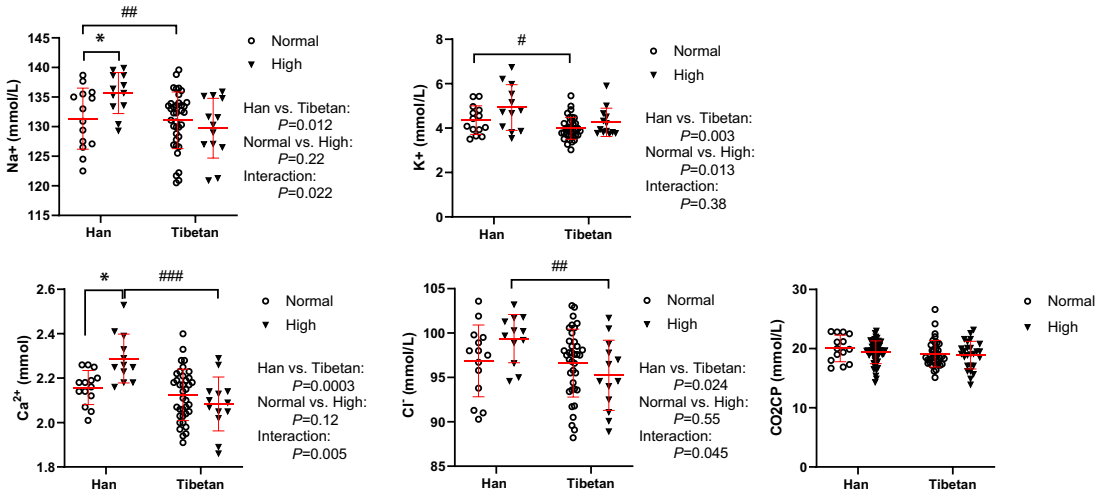



FIGURE 6
Comparisons of serum sodium, potassium, calcium and chlorine concentrations between Tibetans and Hans with normal or high uric acid (UA) levels in Group A. *$P$<0.05, normal UA *vs.* high UA; $^{\#}P$<0.05, $^{\#\#}P$<0.01,$^{\#\#\#}P$<0.001, Tibetans *vs.* Hans. CO2CP: carbon dioxide combining power.

ischemia, and degradation of RNA and DNA can activate the purine metabolism pathway (38).

The high prevalence of hyperuricemia in Tibetans may be attributed to the activation of the purine metabolism pathways. In hypoxia, ATP production is hindered by a lack of oxygen, which accelerates the breakdown of adenosine monophosphate (AMP) to maintain energy levels (26). Under normal conditions, the majority of hypoxanthine is reutilized through the salvage pathway. Under hypoxia, the rate of salvage and degradation decreases because of energy deficiency, which results in hypoxanthine accumulation (39). Therefore, we speculated that metabolic adaptation to hypoxia may contribute to hyperuricemia in Tibetans.

Serum alkaline phosphatase (ALP) is a traditional indirect marker of cholestasis. ATP, adenosine diphosphate (ADP), and AMP can be metabolized to adenosine by two different enzyme systems. Ecto-5-nucleotidase (CD73) converts AMP to adenosine (40) and tissue-nonspecific ALP proteins catabolize nucleotides in a nonspecific manner (41). Soluble adenosine deaminase catabolizes adenosine to inosine. Therefore ALP is involved in the regulation of purinergic signaling by participating in the degradation of

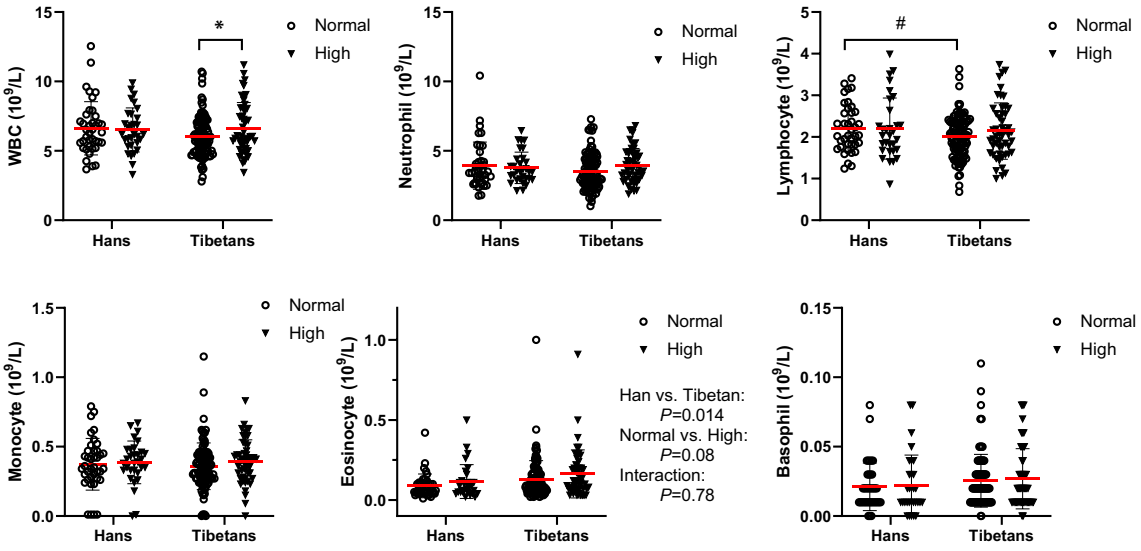



FIGURE 7
Comparisons of the blood cell counts between Tibetans and Hans with normal or high uric acid (UA) levels in Group (B) *$P$<0.05, normal UA *vs.* high UA; $^{\#}P$<0.05.Tibetans *vs.* Hans.

10.3389/fendo.2023.1229659

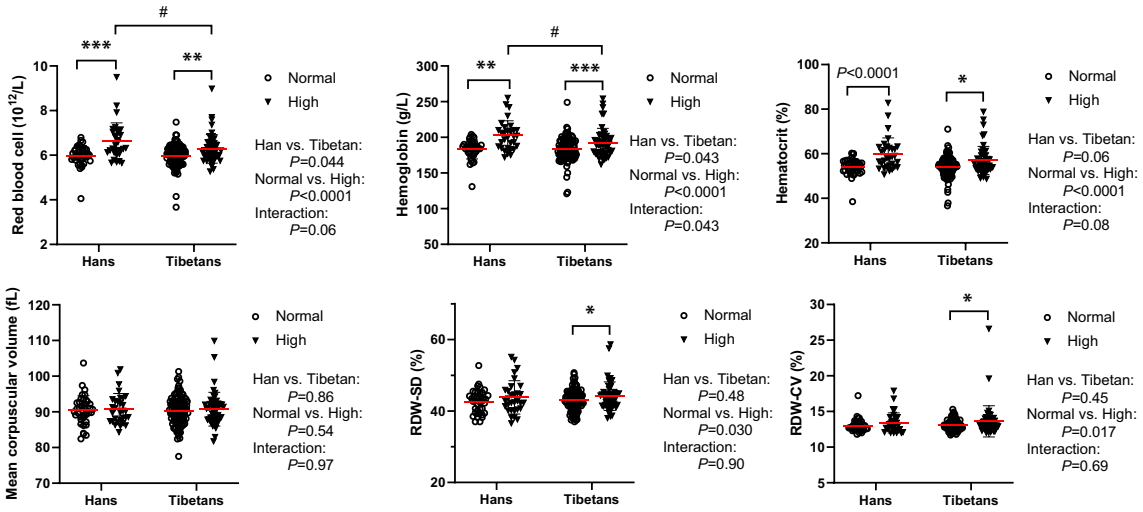



FIGURE 8
Tibetans showed lower levels of erythropoiesis than Hans with normal or high uric acid (UA) levels in Group (B) *$P$<0.05, **$P$<0.01, ***$P$<0.001, normal UA *vs.* high UA; $^{\#}P$<0.05, Tibetans *vs.* Hans.

extracellular nucleotides (42). It has been reported that plasma adenosine concentration and soluble CD73 activity rapidly increase at high altitude (43). Similarly, serum ALP levels increase with serum UA levels in patients with peripheral arterial disease (44). In line with this report, we demonstrated that ALP levels were greatly increased in Tibetans relative to Hans in both the normal and high UA groups. Higher ALP levels may be involved in purine metabolism by mediating nucleotide degradation in Tibetans.

## Metabolic adaptation by promoting glycolytic capacity

At the cellular level, the response to hypoxia results in the promotion of glycolytic capacity, increased glycolytic flux, and lactate efflux in cells (45). LDH is a ubiquitously expressed enzyme that reversibly catalyzes the reduction of pyruvate to L-lactate in the Cori cycle (46). Consistent with elevated muscle LDH activity in highlanders in Sherpas, Nepal (18), native Tibetans had higher LDH levels than Hans in normal UA populations (Figure 2I). This result indicated that Tibetans obtained enhanced glycolysis and gluconeogenesis compared with Hans to maintain adequate ATP levels in a hypoxic environment.

Accordingly, compared with Hans, Tibetans had higher insulin levels in the normal UA population and higher fasting glucose levels in the high UA population in our study (Figure 1). This finding is interesting because the prevalence of diabetes was significantly lower in Tibetans than in Hans in China (47), and people dwelling at high altitudes had a lower diabetes prevalence than

TABLE 5 Multivariate line regression analysis of the blood parameters associated with serum uric acid levels between Hans and Tibetans in Tibet, China (Group B).

| Variable | Hans | | | Tibetan | | |
|---|---|---|---|---|---|---|
| | OR | 95% CI | P value | OR | 95% CI | P value |
| age | 1.303 | -5.461 to 1.183 | 0.20 | 0.728 | -2.062 to 0.9534 | 0.47 |
| RBC | 0.066 | -452.6 to 483.0 | 0.95 | 1.068 | -261.3 to 873.8 | 0.29 |
| HGB | 0.478 | -44.73 to 27.62 | 0.64 | 0.310 | -33.32 to 24.30 | 0.76 |
| HCT | 0.722 | -61.37 to 129.4 | 0.47 | 0.244 | -122.3 to 95.48 | 0.81 |
| MCV | 0.571 | -291.2 to 163.0 | 0.57 | 0.043 | -0.02467 to 0.02577 | 0.97 |
| MCH | 0.629 | -451.1 to 857.9 | 0.53 | 1.033 | -49.98 to 159.1 | 0.30 |
| MCHC | 0.435 | -69.79 to 45.07 | 0.67 | 0.374 | -22.17 to 15.12 | 0.71 |
| RDW-SD | 0.546 | -19.99 to 11.49 | 0.59 | 0.100 | -34.53 to 31.20 | 0.92 |
| RDW-CV | 0.257 | -2.185 to 1.693 | 0.80 | 0.280 | -93.71 to 124.6 | 0.78 |
| Glucose | 2.248 | 3.945 to 74.65 | 0.030 | 1.491 | -3.600 to 0.5060 | 0.14 |

OR, odds ratio; CI, confidence interval; RBC, red blood cell; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red cell distribution width; SD, standard deviation; RDW-CV, RDW-coefficient of variation.

those living at low altitudes (48). However, it was also found that fasting glucose level was significantly higher in the high UA group than in the normal UA group on the Tibetan Plateau (1), and Tibetan highlanders may be vulnerable to glucose intolerance in both China and India (49). Our results suggest that hyperuricemia may modify this protection in subjects at high altitudes, thus increasing the risk of glucose intolerance. Accordingly, the free T4 level was reduced in Tibetans compared with Hans in high UA populations, indicating a relatively lower metabolic rate (glycolytic flux) in this population of Tibetans.

Accumulating evidence suggests that hyperuricemia is associated with impaired glucose metabolism (50) and insulin resistance (51). Hyperinsulinemia can lead to hyperuricemia, but not vice versa (52). Reducing the glycemic quality of carbohydrates over five weeks could reduce UA levels in American subjects (53). Recently, an inverted U-shaped association was observed between major glycemic indices and UA levels in the Chinese population, in which UA levels were elevated with increasing glycemic indices before the inflection points and then decreased with further increases in glycemic indices (54).

In hypoxia-tolerant systems, a shift away from fatty acid oxidation toward a more oxygen-efficient hypometabolic pathway with the downregulation of ATP demand is a common strategy (45). In rats, exposure to hypoxia resulted in the downregulation of fatty acid oxidation and increased pyruvate oxidation (55). Here, we provided corroborative evidence that Tibetans had consistently lower levels of lipometabolism than Hans (Figure 2), indicating better hypoxia tolerance in Tibetans.

## Amino acid metabolism

In the catabolic state of insulin resistance, AST and ALT are responsible for transferring amino acid groups to produce essential intermediate products in the gluconeogenesis pathway. ALT, frequently referred to as glutamic pyruvate transaminase, catalyzes the reversible transamination of alanine and α-ketoglutarate into glutamate and pyruvate (46). Therefore, alanine is a major gluconeogenic precursor. Similar to and often parallel with lactate in the Cori cycle, pyruvate is transformed into alanine by transamination in the muscle, and then alanine is deaminated back to pyruvate in the liver. Therefore, disorders of glucose metabolism are strongly related to liver enzyme abnormalities; for example, AST/ALT levels are inversely correlated with the occurrence of type 2 diabetes (56).

Saliva ALT and glutamic oxaloacetic transaminase levels increase after hypobaric hypoxia in healthy military aircrews (57). Hypoxia-reoxygenation results in the release of LDH, AST, ALT, and XO in the liver of rats (58). Markers of hypoxia correlated significantly with AST and ALT levels in patients with obstructive sleep apnea (59). Specifically, ALT and AST levels were significantly higher in the high UA group than in the normal UA population on the Tibetan Plateau (1). Likewise, regardless of UA levels, we found elevated levels of ALT and AST in Tibetans relative to Hans, in contrast with lower levels of albumin (Figures 3IB, 3IIB) and urea (Figure 4I) in Groups A and B, instead of in Group C. The discrepancy between Groups A, B and C might have contributed to the higher BMI in Tibetans than in Hans in the former two groups and similar BMI between Tibetans and Hans in Group C. These findings indicate that enhanced amino acid utilization and transamination may underlie hyperuricemia in Tibetans.

Glutathione (GSH) is the principal intracellular antioxidant buffer against oxidative stress in the form of reduced GSH and oxidized GSH (GSSG) (19). A favorable reduced/oxidized GSH ratio (GSH/GSSG) is required for cytosolic antioxidant defenses. In short-term exposure to hypoxia, GSH/GSSG was only increased in the muscle of lowlanders but not in highland Sherpa, indicating superior redox homeostasis in highlanders (18). Under hypoxia, the ratio of GSH/GSSG was increased in RBCs (60). Recently, it was found that RBC rely on glutamine to fuel GSH synthesis and pyruvate transamination during hemorrhagic shock (61). GGT is a cell surface enzyme that hydrolyzes the γ-glutamyl bond of extracellular reduced and oxidized GSH into glutamate, cysteine (Cys), and glycine (Gly) (62). In line with the heightened GGT level at high altitudes (35), we found that GGT levels were much higher in Tibetans than in Hans. Moreover, higher GGT levels were reported in the group with impaired fasting glucose than in those with normal fasting glucose in the Chinese population (63), and GGT level was increased to a greater extent by high UA in Hans than in Tibetans (Figure 3I), indicating increased cleavage of GSH in Hans other than Tibetans. Our findings recapitulated the assumption of superior redox homeostasis in the highlander, for example, the Tibetans in our study (18).

## PPP in RBCs

The hematological response to hypoxia is characterized by erythropoiesis, which leads to an increased [Hb] value that increases the oxygen-carrying capacity. It has been reported that native Tibetans have lower [Hb] than Han immigrants (64, 65), which is associated with the positively selected haplotypes of the egl-9 family hypoxia-inducible factor 1 and *PPARA* (66). In Chinese Tibetan immigrants, [Hb] was a positive risk factor for high UA level (5). Accordingly, we found that native Tibetans had lower [Hb] than immigrant Hans under both normal and high UA conditions (Figure 7; Table 3), suggesting a distinctive ethnic difference in the hematological response to altitude.

Recently, it was speculated that the purinergic system may be involved in metabolic adaptations of RBCs to hypoxia. PPP generates ribose sugars for nucleotide synthesis (19). Hypoxia not only promotes glycolysis, but also deregulates PPP and depresses purine catabolism, glutathione homeostasis, and arginine/nitric oxide metabolism in RBCs (60). Likewise, hypoxia can divert

glucose to PPP in the muscle to mitigate the effects of adenosine degradation (19). Consequently, the accumulation of AMP, adenosine, and the PPP product ribose 5-phosphate (ribose-5-P) may activate *de novo* synthesis of purines (67).

In humans, exposure to hypoxia immediately increases RBC glycolysis while shutting down PPP (60). The transient increase in ATP levels during the early response to hypoxia resulted in the accumulation of AMP, adenosine, and the PPP product ribose-5-P in RBCs, proportional to the duration of high-altitude exposure (60). Therefore, severe polycythemia, which is a sign of poor hypoxic adaptation, may be associated with increased UA levels in the Han population (Figure 8; Table 3). In Group B, although the RBC parameters correlated well with UA levels in the two ethnic populations (Table S1), none was independently associated with UA levels in either group (Table 5). However, blood glucose levels were consistently associated with UA levels only in Hans (Tables 4, 5). People with polycythemia seem to be particularly vulnerable to glucose intolerance (68), suggesting poorer adaptation to hypoxia in people with polycythemia than in those without. It should be explored whether severe polycythemia is possibly associated with a high incidence of hyperuricemia in the Hans, as in Group C (Table 3).

## Kidney function

One of the mechanisms underlying hyperuricemia is insulin resistance, which causes a significant decrease in the urinary excretion of UA, sodium, and potassium (69, 70). In healthy individuals, most glucose filtered at the glomerulus is reabsorbed into the epithelial cells from the glomerular filtrate *via* the sodium-glucose cotransporter (SGLT) in the kidney. Glucose then passes into the interstitial fluid and peritubular capillary *via* the glucose transporter maintained by Na+/K+ ATPase in the proximal tubule (71). In our study, in normal UA populations, the higher insulin level in Tibetans relative to Hans indicated possibly worsened insulin resistance in Tibetans. In light of these studies, it is possible that the diabetogenic state in Tibetans would prevent glucose and sodium reabsorption *via* SGLT, resulting in lower serum sodium concentrations. The disparity in serum electrocytes between the two ethnicities was exacerbated by the high UA levels. These results suggest that more severe insulin resistance in Tibetans may be associated with reduced sodium and potassium reabsorption (Figure 6).

Overall, Tibetans had increased glucose metabolism at the expense of lower fatty acid oxidation under anoxia, resulting in enhanced glycolysis and gluconeogenesis, triggering glucose conversion to ribose-5-P, an essential component of nucleotide synthesis *via* PPP (19). Increased glycolysis also promotes the transamination of amino acids, resulting in enhanced gluconeogenesis. Consequently, the higher ALP levels triggered by hypoxia may mediate the degradation of nucleotides, thus increasing the production of purine nucleotides. On the other hand, the immediate shutdown of PPP in RBCs to high altitudes may activate the purinergic system, which may be associated with hyperuricemia in the Hans, who exerted poorer hypoxic adaptation than the Tibetans.

## Study strengths and limitations

These findings are important for the management of metabolic adaptations in hypoxia-related diseases in critical care settings. Our study provides thorough descriptions and comparisons of the biochemical differences between native Tibetans and immigrant Hans in three relatively young populations simultaneously at different altitudes in Tibet, China. Based on these findings, we propose a distinctive etiology underlying the ethnic disparity in hyperuricemia in the Tibetan Plateau. Moreover, we provide corroborative evidence for previous high-altitude adaptation and highlight the complexity of hypoxia-response pathways in humans. Even so, the study on the influence of the environment in Tibetan areas may not offset the influence of genetic background. Future study on the interaction between environment (disease) adaption and ethnicity would provide more information for managing the critical illness.

Although we recruited both sexes in the beginning, for the limitation of the page space, we only present the data from the male population here. The results including the female population would reflect the biochemistry of hyperuricemia at the overall level of population. Because of the religious faith, it is not easy to collect the blood samples from a larger crowd in Tibetan areas, the limitation of insufficient effect due to the small sample size should be considered. In the multiple line regression analysis, other demographic factors such as the economic status, education, and profession may also be the potential confounding factors but were not included in this study. Another limitation was that routine blood tests, including white blood cell count, red RBCs, and [Hb], were not conducted in Group A.

# Conclusions

The risk factors associated with MS for hyperuricemia in immigrant Han individuals did not apply to native Tibetans on the Chinese Tibetan Plateau. However, the higher ALP activity in Tibetans than in Hans may be involved in purine metabolism by mediating the degradation of nucleotides. Moreover, heightened glycolysis, worsened glucose intolerance, increased aminotransferase activity, and reduced UA excretion may underlie hyperuricemia in native Tibetans.

# Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

# Ethics statement

The studies involving humans were approved by Chinese PLA General Hospital (approval identifier S2021-016-01). The studies

were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

X-WR conceived the project, recruited the populations, and collected samples and clinical data. KC, JW, Z-LY, and TJ helped collect and detect the samples. Q-HZ statistically analyzed and interpreted the data and wrote the draft of the manuscript. All the authors have read and approved the final manuscript.

## Funding

The research was supported by Applied Research in Key Fields of Logistics Scientific Research of Chinese PLA (No. B2202).

## Acknowledgments

We acknowledge the medical staff at Tibetan People's Hospital.

## 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.2023.1229659/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The normality test of variables. The biochemistry data in the respective Tibetans and the Hans from the three groups, as well as data in the blood routine test in Group C were analyzed for normality by Shapiro-Wilk test.

SUPPLEMENTARY FIGURE 2
Comparisons of the biomarkers for heart failure between the Hans and the Tibetans with normal or high uric acid (UA) levels in Group A. *$P$<0.05, normal UA *vs.* high UA.

## References

1. Yao S, Zhou Y, Xu L, Zhang Q, Bao S, Feng H, et al. Association between hyperuricemia and metabolic syndrome: A cross-sectional study in Tibetan adults on the Tibetan plateau. *Front Endocrinol (Lausanne)* (2022) 13:964872. doi: 10.3389/fendo.2022.964872

2. Estiverne C, Mandal AK, Mount DB. Molecular pathophysiology of uric acid homeostasis. *Semin Nephrol* (2020) 40:535–49. doi: 10.1016/j.semnephrol.2020.12.006

3. Hua X, Wang YY, Jia P, Xiong Q, Hu Y, Chang Y, et al. Multi-level transcriptome sequencing identifies COL1A1 as a candidate marker in human heart failure progression. *BMC Med* (2020) 18:2. doi: 10.1186/s12916-019-1469-4

4. Zhang X, Meng Q, Feng J, Liao H, Shi R, Shi D, et al. The prevalence of hyperuricemia and its correlates in Ganzi Tibetan Autonomous Prefecture, Sichuan Province, China. *Lipids Health Dis* (2018) 17:235. doi: 10.1186/s12944-018-0882-6

5. Song Z, Zhang A, Luo J, Xiong G, Peng H, Zhou R, et al. Prevalence of high-altitude polycythemia and hyperuricemia and risk factors for hyperuricemia in high-altitude immigrants. *High Alt Med Biol* (2023) 24:132–8. doi: 10.1089/ham.2022.0133

6. Liu R, Han C, Wu D, Xia X, Gu J, Guan H, et al. Prevalence of hyperuricemia and gout in mainland China from 2000 to 2014: A systematic review and meta-analysis. *BioMed Res Int* (2015) 2015:762820. doi: 10.1155/2015/762820

7. 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

8. Dong X, Zhang H, Wang F, Liu X, Yang K, Tu R, et al. Epidemiology and prevalence of hyperuricemia among men and women in Chinese rural population: The Henan Rural Cohort Study. *Mod Rheumatol* (2020) 30:910–20. doi: 10.1080/14397595.2019.1660048

9. Liu F, Du GL, Song N, Ma YT, Li XM, Gao XM, et al. Hyperuricemia and its association with adiposity and dyslipidemia in Northwest China: results from cardiovascular risk survey in Xinjiang (CRS 2008-2012). *Lipids Health Dis* (2020) 19:58. doi: 10.1186/s12944-020-01211-z

10. Maloberti A, Maggioni S, Occhi L, Triglione N, Panzeri F, Nava S, et al. Sex-related relationships between uric acid and target organ damage in hypertension. *J Clin Hypertens (Greenwich)* (2018) 20:193–200. doi: 10.1111/jch.13136

11. Maloberti A, Qualliu E, Occhi L, Sun J, Grasso E, Tognola C, et al. Hyperuricemia prevalence in healthy subjects and its relationship with cardiovascular target organ damage. *Nutr Metab Cardiovasc Dis* (2021) 31:178–85. doi: 10.1016/j.numecd.2020.08.015

12. Zhou D, Liu Y, Zhang X, Gu X, Wang H, Luo X, et al. Functional polymorphisms of the ABCG2 gene are associated with gout disease in the Chinese Han male population. *Int J Mol Sci* (2014) 15:9149–59. doi: 10.3390/ijms15059149

13. Jia Z, Zhao X, Liu X, Zhao L, Jia Q, Shi J, et al. Impacts of the plateau environment on the gut microbiota and blood clinical indexes in han and tibetan individuals. *mSystems* (2020) 5:e00660–19. doi: 10.1128/mSystems.00660-19

14. Gilbert-Kawai ET, Milledge JS, Grocott MP, Martin DS. King of the mountains: Tibetan and Sherpa physiological adaptations for life at high altitude. *Physiol (Bethesda)* (2014) 29:388–402. doi: 10.1152/physiol.00018.2014

15. Lau DS, Connaty AD, Mahalingam S, Wall N, Cheviron ZA, Storz JF, et al. Acclimation to hypoxia increases carbohydrate use during exercise in high-altitude deer mice. *Am J Physiol Regul Integr Comp Physiol* (2017) 312:R400–11. doi: 10.1152/ajpregu.00365.2016

16. Erzurum SC, Ghosh S, Janocha AJ, Xu W, Bauer S, Bryan NS, et al. Higher blood flow and circulating NO products offset high-altitude hypoxia among Tibetans. *Proc Natl Acad Sci U.S.A.* (2007) 104:17593–8.

17. Ge RL, Simonson TS, Gordeuk V, Prchal JT, McClain DA. Metabolic aspects of high-altitude adaptation in Tibetans. *Exp Physiol* (2015) 100:1247–55. doi: 10.1113/EP085292

18. Horscroft JA, Kotwica AO, Laner V, West JA, Hennis PJ, Levett DZH, et al. Metabolic basis to Sherpa altitude adaptation. *Proc Natl Acad Sci U.S.A.* (2017) 114:6382–7.

19. Chicco AJ, Le CH, Gnaiger E, Dreyer HC, Muyskens JB, D'Alessandro A, et al. Adaptive remodeling of skeletal muscle energy metabolism in high-altitude hypoxia: Lessons from AltitudeOmics. *J Biol Chem* (2018) 293:6659–71. doi: 10.1074/jbc.RA117.000470

20. Wu T, Wang X, Wei C, Cheng H, Wang X, Li Y, et al. Hemoglobin levels in Qinghai-Tibet: different effects of gender for Tibetans vs. *Han. J Appl Physiol* (1985) 98:598–604. doi: 10.1152/japplphysiol.01034.2002

21. Ge RL, Simonson TS, Cooksey RC, Tanna U, Qin G, Huff CD, et al. Metabolic insight into mechanisms of high-altitude adaptation in Tibetans. *Mol Genet Metab* (2012) 106:244–7. doi: 10.1016/j.ymgme.2012.03.003

22. Hochachka PW, Clark CM, Monge C, Stanley C, Brown WD, Stone CK, et al. Sherpa brain glucose metabolism and defense adaptations against chronic hypoxia. *J Appl Physiol* (1985) 81:1355–61. doi: 10.1152/jappl.1996.81.3.1355

23. Murray AJ. Energy metabolism and the high-altitude environment. *Exp Physiol* (2016) 101:23–7. doi: 10.1113/EP085317

24. Kayyali US, Donaldson C, Huang H, Abdelnour R, Hassoun PM. Phosphorylation of xanthine dehydrogenase/oxidase in hypoxia. *J Biol Chem* (2001) 276:14359–65. doi: 10.1074/jbc.M010100200

25. Tsushima Y, Nishizawa H, Tochino Y, Nakatsuji H, Sekimoto R, Nagao H, et al. Uric acid secretion from adipose tissue and its increase in obesity. *J Biol Chem* (2013) 288:27138–49. doi: 10.1074/jbc.M113.485094

26. Wheaton WW, Chandel NS. Hypoxia. 2. Hypoxia regulates cellular metabolism. *Am J Physiol Cell Physiol* (2011), 300:C385–93. doi: 10.1152/ajpcell.00485.2010

27. Du Y, Qi M, Wang W, Chen B. Effect of high-altitude hypoxia environment on uric acid excretion, desmin protein level in podocytes, and na+-K+- ATPase activity. *Cell Mol Biol (Noisy-le-grand)* (2022) 68:84–91. doi: 10.14715/cmb/2022.68.6.14

28. Multidisciplinary Expert Task Force on H, Related D. Chinese multidisciplinary expert consensus on the diagnosis and treatment of hyperuricemia and related diseases. *Chin Med J (Engl)* (2017) 130:2473–88. doi: 10.4103/0366-6999.216416

29. Yanai H, Adachi H, Hakoshima M, Katsuyama H. Molecular biological and clinical understanding of the pathophysiology and treatments of hyperuricemia and its association with metabolic syndrome, cardiovascular diseases and chronic kidney disease. *Int J Mol Sci* (2021) 22:9221. doi: 10.3390/ijms22179221

30. Maloberti A, Vanoli J, Finotto A, Bombelli M, Facchetti R, Redon P, et al. Uric acid relationships with lipid profile and adiposity indices: Impact of different hyperuricemic thresholds. *J Clin Hypertens (Greenwich)* (2023) 25:78–85. doi: 10.1111/jch.14613

31. Creeden JF, Gordon DM, Stec DE, Hinds TD Jr. Bilirubin as a metabolic hormone: the physiological relevance of low levels. *Am J Physiol Endocrinol Metab* (2021) 320:E191–207. doi: 10.1152/ajpendo.00405.2020

32. Maries L, Manitiu I. Diagnostic and prognostic values of B-type natriuretic peptides (BNP) and N-terminal fragment brain natriuretic peptides (NT-pro-BNP). *Cardiovasc J Afr* (2013) 24:286–9. doi: 10.5830/CVJA-2013-055

33. Salvagno GL, Sanchis-Gomar F, Picanza A, Lippi G. Red blood cell distribution width: A simple parameter with multiple clinical applications. *Crit Rev Clin Lab Sci* (2015) 52:86–105. doi: 10.3109/10408363.2014.992064

34. Matsuzawa Y, Funahashi T, Nakamura T. The concept of metabolic syndrome: contribution of visceral fat accumulation and its molecular mechanism. *J Atheroscler Thromb* (2011) 18:629–39. doi: 10.5551/jat.7922

35. Gonzales GF, Tapia V. Increased levels of serum gamma-glutamyltransferase and uric acid on metabolic, hepatic and kidney parameters in subjects at high altitudes. *J Basic Clin Physiol Pharmacol* (2015) 26:81–7. doi: 10.1515/jbcpp-2013-0162

36. Rebora P, Centola M, Morici N, Sacco A, Occhino G, Viola G, et al. Uric acid associated with acute heart failure presentation in Acute Coronary Syndrome patients. *Eur J Intern Med* (2022) 99:30–7. doi: 10.1016/j.ejim.2022.01.018

37. Nishino T, Okamoto K. Mechanistic insights into xanthine oxidoreductase from development studies of candidate drugs to treat hyperuricemia and gout. *J Biol Inorg Chem* (2015) 20:195–207. doi: 10.1007/s00775-014-1210-x

38. Furuhashi M. New insights into purine metabolism in metabolic diseases: role of xanthine oxidoreductase activity. *Am J Physiol Endocrinol Metab* (2020) 319:E827–34. doi: 10.1152/ajpendo.00378.2020

39. Chu WY, Allegaert K, Dorlo TPC, Huitema ADR, Group AS. Semi-mechanistic modeling of hypoxanthine, xanthine, and uric acid metabolism in asphyxiated neonates. *Clin Pharmacokinet* (2022) 61:1545–58. doi: 10.1007/s40262-022-01164-9

40. Zimmermann H, Zebisch M, Strater N. Cellular function and molecular structure of ecto-nucleotidases. *Purinergic Signal* (2012) 8:437–502. doi: 10.1007/s11302-012-9309-4

41. Picher M, Burch LH, Hirsh AJ, Spychala J, Boucher RC. Ecto 5'-nucleotidase and nonspecific alkaline phosphatase. Two AMP-hydrolyzing ectoenzymes with distinct roles in human airways. *J Biol Chem* (2003) 278:13468–79. doi: 10.1074/jbc.M300569200

42. Sebastian-Serrano A, de Diego-Garcia L, Martinez-Frailes C, Avila J, Zimmermann H, Millan JL, et al. Tissue-nonspecific alkaline phosphatase regulates purinergic transmission in the central nervous system during development and disease. *Comput Struct Biotechnol J* (2015) 13:95–100. doi: 10.1016/j.csbj.2014.12.004

43. Liu H, Zhang Y, Wu H, D'Alessandro A, Yegutkin GG, Song A, et al. Beneficial role of erythrocyte adenosine A2B receptor-mediated AMP-activated protein kinase activation in high-altitude hypoxia. *Circulation* (2016) 134:405–21. doi: 10.1161/CIRCULATIONAHA.116.021311

44. Cheung BM, Ong KL, Wong LY. Elevated serum alkaline phosphatase and peripheral arterial disease in the United States National Health and Nutrition Examination Survey 1999-2004. *Int J Cardiol* (2009) 135:156–61. doi: 10.1016/j.ijcard.2008.03.039

45. Murray AJ, Montgomery HE, Feelisch M, Grocott MPW, Martin DS. Metabolic adjustment to high-altitude hypoxia: from genetic signals to physiological implications. *Biochem Soc Trans* (2018) 46:599–607. doi: 10.1042/BST20170502

46. Gray LR, Tompkins SC, Taylor EB. Regulation of pyruvate metabolism and human disease. *Cell Mol Life Sci* (2014) 71:2577–604. doi: 10.1007/s00018-013-1539-2

47. Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. *JAMA* (2017) 317:2515–23. doi: 10.1001/jama.2017.7596

48. Santos JL, Perez-Bravo F, Carrasco E, Calvillan M, Albala C. Low prevalence of type 2 diabetes despite a high average body mass index in the Aymara natives from Chile. *Nutrition* (2001) 17:305–9. doi: 10.1016/S0899-9007(00)00551-7

49. Okumiya K, Sakamoto R, Ishimoto Y, Kimura Y, Fukutomi E, Ishikawa M, et al. Glucose intolerance associated with hypoxia in people living at high altitudes in the Tibetan highland. *BMJ Open* (2016) 6:e009728. doi: 10.1136/bmjopen-2015-009728

50. Di Bonito P, Valerio G, Licenziati MR, Campana G, Del Giudice EM, Di Sessa A, et al. Uric acid, impaired fasting glucose and impaired glucose tolerance in youth with overweight and obesity. *Nutr Metab Cardiovasc Dis* (2021) 31:675–80. doi: 10.1016/j.numecd.2020.10.007

51. Spatola L, Ferraro PM, Gambaro G, Badalamenti S, Dauriz M. Metabolic syndrome and uric acid nephrolithiasis: insulin resistance in focus. *Metabolism* (2018) 83:225–33. doi: 10.1016/j.metabol.2018.02.008

52. McCormick N, O'Connor MJ, Yokose C, Merriman TR, Mount DB, Leong A, et al. Assessing the causal relationships between insulin resistance and hyperuricemia and gout using bidirectional mendelian randomization. *Arthritis Rheumatol* (2021) 73:2096–104. doi: 10.1002/art.41779

53. Juraschek SP, McAdams-Demarco M, Gelber AC, Sacks FM, Appel LJ, White KJ, et al. 3rd, effects of lowering glycemic index of dietary carbohydrate on plasma uric acid levels: the omniCarb randomized clinical trial. *Arthritis Rheumatol* (2016) 68:1281–9. doi: 10.1002/art.39527

54. Zhu YY, Zheng RZ, Wang GX, Chen L, Shi LX, Su Q, et al. Inverted U-shaped associations between glycemic indices and serum uric acid levels in the general chinese population: findings from the China cardiometabolic disease and cancer cohort (4C) study. *BioMed Environ Sci* (2021) 34:9–18.

55. Essop MF, Razeghi P, McLeod C, Young ME, Taegtmeyer H, Sack MN. Hypoxia-induced decrease of UCP3 gene expression in rat heart parallels metabolic gene switching but fails to affect mitochondrial respiratory coupling. *Biochem Biophys Res Commun* (2004) 314:561–4. doi: 10.1016/j.bbrc.2003.12.121

56. Niu H, Zhou Y. Nonlinear relationship between AST-to-ALT ratio and the incidence of type 2 diabetes mellitus: A follow-up study. *Int J Gen Med* (2021) 14:8373–82. doi: 10.2147/IJGM.S341790

57. Mominzadeh M, Mirzaii-Dizgah I, Mirzaii-Dizgah MR, Mirzaii-Dizgah MH. Stimulated saliva aminotransaminase alteration after experiencing acute hypoxia training. *Air Med J* (2014) 33:157–60. doi: 10.1016/j.amj.2014.03.004

58. Kooij A, Schiller HJ, Schijns M, Van Noorden CJ, Frederiks WM. Conversion of xanthine dehydrogenase into xanthine oxidase in rat liver and plasma at the onset of reperfusion after ischemia. *Hepatology* (1994) 19:1488–95. doi: 10.1002/hep.1840190626

59. Norman D, Bardwell WA, Arosemena F, Nelesen R, Mills PJ, Loredo JS, et al. Serum aminotransferase levels are associated with markers of hypoxia in patients with obstructive sleep apnea. *Sleep* (2008) 31:121–6. doi: 10.1093/sleep/31.1.121

60. D'Alessandro A, Nemkov T, Sun K, Liu H, Song A, Monte AA, et al. AltitudeOmics: red blood cell metabolic adaptation to high altitude hypoxia. *J Proteome Res* (2016) 15:3883–95. doi: 10.1021/acs.jproteome.6b00733

61. Reisz JA, Slaughter AL, Culp-Hill R, Moore EE, Silliman CC, Fragoso M, et al. Red blood cells in hemorrhagic shock: a critical role for glutaminolysis in fueling alanine transamination in rats. *Blood Adv* (2017) 1:1296–305. doi: 10.1182/bloodadvances.2017007187

62. Hanigan MH. Gamma-glutamyl transpeptidase: redox regulation and drug resistance. *Adv Cancer Res* (2014) 122:103–41. doi: 10.1016/B978-0-12-420117-0.00003-7

63. Wu J, Qiu L, Yan WH, Cheng XQ, Wu W, Guo XZ, et al. Serum gamma-glutamyltransferase and uric acid levels are associated with impaired fasting glucose in adults from Inner Mongolia, China. *BMC Public Health* (2013) 13:294. doi: 10.1186/1471-2458-13-294

64. Beall CM, Goldstein MC. Hemoglobin concentration of pastoral nomads permanently resident at 4,850-5,450 meters in Tibet. *Am J Phys Anthropol* (1987) 73:433–8. doi: 10.1002/ajpa.1330730404

65. Curran LS, Zhuang J, Sun SF, Moore LG. Ventilation and hypoxic ventilatory responsiveness in Chinese-Tibetan residents at 3,658 m. *J Appl Physiol* (1985) 83 (1997):2098–104.

66. Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, et al. Genetic evidence for high-altitude adaptation in Tibet. *Science* (2010) 329:72–5. doi: 10.1126/science.1189406

67. Mandal AK, Mount DB. The molecular physiology of uric acid homeostasis. *Annu Rev Physiol* (2015) 77:323–45. doi: 10.1146/annurev-physiol-021113-170343

68. Okumiya K, Sakamoto R, Kimura Y, Ishimoto Y, Wada T, Ishine M, et al. Strong association between polycythemia and glucose intolerance in elderly high-altitude dwellers in Asia. *J Am Geriatr Soc* (2010) 58:609–11. doi: 10.1111/j.1532-5415.2010.02753.x

69. Muscelli E, Natali A, Bianchi S, Bigazzi R, Galvan AQ, Sironi AM, et al. Effect of insulin on renal sodium and uric acid handling in essential hypertension. *Am J Hypertens* (1996) 9:746–52. doi: 10.1016/0895-7061(96)00098-2

70. Perez-Ruiz F, Aniel-Quiroga MA, Herrero-Beites AM, Chinchilla SP, Erauskin GG, Merriman T. Renal clearance of uric acid is linked to insulin resistance and lower excretion of sodium in gout patients. *Rheumatol Int* (2015) 35:1519–24. doi: 10.1007/s00296-015-3242-0

71. Cowie MR, Fisher M. SGLT2 inhibitors: mechanisms of cardiovascular benefit beyond glycaemic control. *Nat Rev Cardiol* (2020) 17:761–72. doi: 10.1038/s41569-020-0406-8

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 08 December 2023
DOI 10.3389/fendo.2023.1212647



OPEN ACCESS

EDITED BY
Di Liu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Benson Hamooya,
Mulungushi University, Zambia
Kourosh Zarea,
Ahvaz Jundishapur University of Medical Sciences, Iran

*CORRESPONDENCE
Yixuan Niu
niuyx2003@126.com
Weiping Guan
guanweiping@126.com

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

RECEIVED 26 April 2023
ACCEPTED 07 November 2023
PUBLISHED 08 December 2023

CITATION
Li C, Tao T, Tang Y, Lu H, Zhang H, Li H, Liu X, Guan W and Niu Y (2023) The association of psychological stress with metabolic syndrome and its components: cross-sectional and bidirectional two-sample Mendelian randomization analyses. *Front. Endocrinol.* 14:1212647. doi: 10.3389/fendo.2023.1212647



# The association of psychological stress with metabolic syndrome and its components: cross-sectional and bidirectional two-sample Mendelian randomization analyses

Cancan Li[1†], Tianqi Tao[2†], Yanyan Tang[2], Huimin Lu[1], Hongfeng Zhang[2], Huixin Li[2], Xiuhua Liu[2], Weiping Guan[2*] and Yixuan Niu[2*]

[1]Beijing Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing, China, [2]Department of Geriatrics, The Second Medical Center and National Clinical Research Center for Geriatric Diseases, Chinese PLA General Hospital, Beijing, China

**Background:** Metabolic syndrome (MetS) is a group of co-occurring conditions that increase the risk of cardiovascular disease, which include the conditions of hypertension, overweight or obesity, hyperglycemia, and dyslipidemia. Psychological stress is gradually being taken seriously, stemming from the imbalance between environmental demands and individual perceptions. However, the potential causal relationship between psychological stress and MetS remains unclear.

**Method:** We conducted cross-sectional and bidirectional Mendelian randomization (MR) analyses to clarify the potential causal relationship of psychological stress with MetS and its components. Multivariable logistic regression models were used to adjust for potential confounders in the cross-sectional study of the Chinese population, including 4,933 individuals (70.1% men; mean age, 46.13 ± 8.25). Stratified analyses of sexual characteristics were also performed. Bidirectional MR analyses were further carried out to verify causality based on summary-level genome-wide association studies in the European population, using the main analysis of the inverse variance-weighted method.

**Results:** We found that higher psychological stress levels were cross-sectionally associated with an increased risk of hypertension in men (odds ratio (OR), 1.341; 95% confidence interval (CI), 1.023–1.758; p = 0.034); moreover, higher levels of hypertension were cross-sectionally associated with an increased risk of psychological stress in men and the total population (men: OR, 1.545 (95% CI, 1.113–2.145); p = 0.009; total population: OR, 1.327 (95% CI, 1.025–1.718); p = 0.032). Genetically predicted hypertension was causally associated with a higher risk of psychological stress in the inverse-variance weighted MR model (OR, 2.386 (95% CI, 1.209–4.710); *p* = 0.012). However, there was no association between psychological stress and MetS or the other three risk factors

(overweight or obesity, hyperglycemia, and dyslipidemia) in cross-sectional and MR analyses.

**Conclusion:** Although we did not observe an association between psychological stress and MetS, we found associations between psychological stress and hypertension both in cross-sectional and MR studies, which may have implications for targeting hypertension-related factors in interventions to improve mental and metabolic health. Further study is needed to confirm our findings.

KEYWORDS

metabolic syndrome, psychological stress, hypertension, risk factor, cross-sectional study, Mendelian randomization analysis

# Introduction

Metabolic syndrome (MetS), also known as syndrome X or insulin resistance, is a cluster of co-occurring conditions, including hypertension, elevated fasting glucose, elevated triglycerides (TG), lowered high-density lipoprotein cholesterol (HDL-C), and abdominal obesity (1). Individuals with MetS are more susceptible to developing cardiovascular disease (CVD), type 2 diabetes mellitus, and cancers and have a higher risk of death (1, 2). MetS and MetS-related conditions are becoming major public health burdens worldwide. It is reported that over a quarter of the entire world population (about a billion people) has MetS, including one-third of the Chinese population (3, 4). Early recognition and intervention are important to prevent the development of MetS and its progression to chronic diseases, such as CVD (1, 3).

Psychological stress is a major public health challenge that can induce a range of physiological responses involving the neurological, endocrine, and immune systems (5, 6). Because both psychological stress and MetS are risk factors for CVD, their association has become a widespread concern in recent years (1, 6). Epidemiological studies suggested that psychological stress may predict the risk of MetS, hypertension, and obesity (7, 8). This could be attributed to the chronic nature of psychological stress, which can induce long-term alterations in emotional, physiological, and behavioral responses, subsequently influencing susceptibility to diseases such as MetS (9). In the context of existing Chinese studies, two focused on occupational stress (10, 11), while one focused on psychological stress with a relatively small sample size of 345 participants (7). This underscores the necessity of investigating the association between psychological stress and MetS in more extensive and representative Chinese populations. Nonetheless, some data from cross-sectional and cohort studies indicated that psychological factors, such as psychological stress, were outcomes of MetS rather than risk factors (12), while other studies reported no significant associations (13, 14). The aforementioned inconsistent results emphasize the need to investigate the causal relationship between psychological stress and MetS and its components. Such inquiry could provide a scientific foundation for developing targeted prevention policies aimed at mitigating psychological stress, MetS, and associated risk factors.

Mendelian randomization (MR) is a novel approach used to estimate the causal relationship between psychological stress and MetS using genetic variants robustly related to exposure as instrumental variables (IVs), which could overcome the limitations of observational research (15, 16). Due to the random allocation of genotypes from parents to offspring, the relationship between genetic variants and outcomes remains unaffected by common confounding factors, making a causal sequence plausible (15). Accordingly, in this current study, we aim to investigate the association of psychological stress with MetS and its components in general Chinese populations and to assess the causality using a bidirectional two-sample MR technique.

# Method

## Study design and population

This cross-sectional study was used to examine the association of psychological stress with MetS and its components, which included 4933 patients from the Chinese People's Liberation Army General Hospital (Beijing, China) between July 2017 and June 2019. We included individuals aged 18 years and older who provided signed informed consent, had no missing data on standardized questionnaires or clinical characteristics, and were not enrolled in a clinical trial. Participants were excluded from the study if they failed to meet the inclusion criteria or had undergone surgery for cancer or other severe illnesses.

## Sample size estimation

Based on one published cross-sectional study in Asia (17), the psychological stress risk (23%) between the MetS and non-MetS groups was 24% and 22%, respectively. At 80% power (two-sided significance level of 0.05), using the sample size estimation formula

for an independent sample comparison, the total sample size was estimated as:

$$n = 4 \times \left[\frac{(Z_{\frac{\alpha}{2}} + Z_{\beta})\sigma}{\delta}\right]^2 = 4 \times \left[\frac{(1.96 + 0.84) \times 23}{2}\right]^2 \approx 4,147 \quad (1)$$

Consequently, the required sample size would be estimated to be 4,147. The sample size (4,933) of this current cross-sectional study meets the criteria of 4,147.

## Ethical consideration

This study conforms to the principles of the Declaration of Helsinki and relevant ethical guidelines. Approval for this study was granted by the Medical Ethics Committee of the Chinese People's Liberation Army General Hospital (S2019-131-01).

## Data collection of demographic data and blood samples

Participants' demographic data, including age, sex, educational attainment, marital status, smoking, alcohol consumption, physical activity, family history of diabetes, family history of hypertension, family history of CVD, and family history of stroke, was collected through face-to-face interviews with trained nurses conducting the interviews. Physical inactivity was defined as less than 2 h of physical activity per week (18). In addition, participants' height (with a standiometer while wearing socks), body weight (with a digital weighing scale clothed in a light examination gown), waist circumference (with a measuring tape positioned at the midpoint between the lowest rib and iliac crest), and hip circumference (with a measuring tape) were measured by trained nurses. Body mass index (BMI) was calculated as weight (kg) divided by the square of height ($m^2$), and the waist-to-hip ratio was calculated as waist circumference divided by hip circumference. The participants were seated for at least 5 min before two blood pressure measurements were taken by trained nurses using an automated sphygmomanometer, and the average of the two measurements was recorded.

Blood samples were collected from the antecubital vein after overnight fasting. These samples were processed, transported to the Clinical Laboratory Department of the Chinese People's Liberation Army General Hospital, and analyzed within 24 h. Fasting blood glucose (FBG), TG, total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), and HDL-C levels were determined using a Roche C8000 automatic biochemical analyzer (Roche, Mannheim, Germany). C-reactive protein (CRP) was measured using an immunoturbidimetric assay (Siemens Healthcare Diagnostics, Germany).

## Measurement of psychological stress

The Chinese version of the Perceived Stress Scale (CPSS) was used to reflect psychological stress levels. The CPSS comprises seven positive and seven negative items rated on a 5-point Likert scale: 0 = never, 1 = rarely, 2 = sometimes, 3 = often, and 4 = always (19, 20). The total CPSS score ranges from 0 to 56, with higher scores indicating greater psychological stress; a score< 29 was defined as participants with no or low psychological stress, and a score ≥ 29 was defined as participants with moderate or high psychological stress (19, 20). The CPSS was verified in a smoking population and showed good reliability (Cronbach's alpha = 0.85), structural validity, and co-validity (20).

## Measurement of depression and anxiety symptoms

Depressive- and anxiety-related symptoms were measured using the Chinese version of the Zung Self-Rating Depression Scale (SDS) and the Zung Self-Rating Anxiety Scale (SAS) (21, 22). Both the SDS and SAS questionnaires are composed of 20 items (10 positive and 10 negative items) scored on a 4-point scale (1 = never or rarely; 2 = sometimes; 3 = frequently; and 4 = most of the time), with higher scores representing higher depression or anxiety symptoms. The index score (range, 25–100) was equal to the raw score (range, 20–80) × 1.25, and an index score ≥ 50 was defined as participants with depression or anxiety symptoms; otherwise, they were classified as not having depression or anxiety symptoms according to the Chinese norm (21–24). Furthermore, the Chinese version of the SDS and SAS questionnaires were shown to have good reliability (Cronbach's alpha = 0.796; Cronbach's alpha = 0.850) and validity in the Chinese population (23, 24).

## Measurement of sleep quality

Sleep quality was assessed using the Pittsburgh Sleep Quality Index (PSQI). The PSQI consists of 19 items under seven components (subjective sleep quality, sleep latency, sleep duration, habitual sleep, efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction) rated on a 4-point scale (0 = never to 3 = often). The total score on the PSQI scale ranges from 0 to 21, with a score of > 5 indicating poor sleep quality (25). The Chinese version of the PSQI has been verified in a Chinese group and has shown good reliability (Cronbach's alpha = 0.850) and validity (26).

## Definition of MetS and its risk components

In this study, MetS was defined according to the Chinese Diabetes Society (CDS) criteria as having at least three of the following metabolic abnormalities: (1) overweight or obesity: BMI ≥ 25 $kg/m^2$; (2) hypertension: systolic blood pressure (SBP) ≥ 140 mmHg, diastolic blood pressure (DBP) ≥ 90 mmHg, and (or) being treated for hypertension; (3) hyperglycemia: FBG ≥ 6.1 mmol/L, 2-h oral glucose tolerance test ≥ 7.8 mmol/L, and (or) being drug treated for type 2 diabetes; and (4) dyslipidemia: TG ≥ 1.7 mmol/L and (or) HDL-C< 0.9 mmol/L in men,< 1.0 mmol/L in women (27). The CDS has been validated in the Chinese

population, showing good validity (specificity = 0.989) and reliability (28).

## Statistical analyses

The Kolmogorov–Smirnov test was performed for continuous data. Continuous data of normal distribution are represented as mean ± standard deviation (SD) ($\bar{x} \pm s$), and the analysis was performed using the two independent samples *t*-test (Student's *t*-test). Non-normally distributed continuous data were represented as median and interquartile range (IQR), and the analysis was performed using the Mann–Whitney *U* test. The chi-square test ($\chi^2$-test) was performed to analyze categorical variables, which were expressed as frequencies, percentages, or ratios (%). The least absolute shrinkage and selection operator (Lasso) algorithm was used to screen potential confounding factors that were significantly associated with psychological stress, MetS, and its components, thus avoiding overfitting and effectively controlling the model's complexity. Significant potential confounding factors selected with Lasso were then introduced into multivariate logistic regression analyses. SPSS (version 25, IBM) statistical software was used for statistical analysis of the data, and a two-tailed *p*-value below 0.05 was considered statistically significant.

## MR analysis

A bidirectional two-sample MR analysis was performed to evaluate the causality between psychological stress and MetS and its components (i.e., hypertension, BMI, TG, HDL-C, and FBG) to validate the cross-sectional results. MR depends on three premises: (1) genetic variation as an instrumental variable (IV) is significantly associated with exposure, (2) IVs are not related to any confounders of the exposure–outcome association, and (3) IVs can affect the outcome only via exposure (Supplementary Figure S1). To avoid bias due to participant overlap, this MR study relied on the largest available genome-wide association studies (GWASs) on different international consortia for exposure and outcomes. For instance, we obtained summary GWAS data associated with MetS from the most comprehensive GWAS in the UK Biobank, which included 291,107 individuals (59,677 cases and 231,430 controls) (29). Summary-level data on psychological stress were collected from the FinnGen Biobank (ID: finn-b-F5_NEUROTIC), which included 218,792 individuals (20,682 cases and 198,110 controls) (https://gwas.mrcieu.ac.uk/). The sources of GWAS data on hypertension (30), BMI, FBG (31), HDL-C, and TG (32) are shown in Supplementary Table S1.

The inverse variance-weighted (IVW) method, which assumes that all genetic variants are valid IVs (with no heterogeneity or horizontal pleiotropy), was used as the primary approach for evaluating potential causality (33). Thereafter, five alternative analyses (MR-Egger regression method, weighted median estimator (WME), MR pleiotropy residual sum and outlier (MR-PRESSO) weighted mode, and simple mode) were performed to assess the causal effects. Of these, the WME was regarded as a valid estimation when there was heterogeneity in the genetic variants without horizontal pleiotropy (34). MR-Egger regression was used as the main evaluation when there was heterogeneity and pleiotropy, and its intercept was used to test horizontal pleiotropy (35). Meanwhile, the MR-PRESSO global test was conducted to analyze the directional horizontal pleiotropy and identify outliers (36). For the selection of IVs, we chose single nucleotide polymorphisms (SNPs) of psychological stress that reached the genome-wide significance threshold ($p< 1\times10^{-5}$), MetS, and its components at $p< 5\times10^{-8}$. Significant SNPs at linkage disequilibrium (LD) ($r^2$ threshold< 0.001 within a 10-Mb window) were excluded to minimize the effect of strong LD on the results. In addition, we illustrated the magnitude of heterogeneity across all IVs using Cochran's *Q* statistic and a funnel plot (37). Furthermore, the leave-one-out method was used for the sensitive analysis (15). The $R^2$ (Eq. 1: $R^2 = 2 \times \text{eaf} \times (1 - \text{eaf}) \times \text{Beta}^2$] and F statistics (Eq. 2: F statistic $= \frac{R^2 \times (N-2)}{(1-R^2)}$) of each SNP were used to verify the strength of exposure, with an F statistic of > 10 indicating a lower risk of IV bias. We then summed them up to assess the $R^2$ and *F* statistics (38). Power calculations were performed using the mRnd software (https://cnsgenomics.com/shiny/mRnd/) (39). All data analyses were conducted using the "TwoSampleMR" and "MR-PRESSO" packages in R version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at a two-tailed *p*-value< 0.05.

# Results

## Cross-sectional study

### Participants' characteristics

After excluding 28,591 individuals due to incomplete questionnaire results, incomplete blood sample information, or falling under the exclusion criteria, the data of 4,933 participants (70.1% men; mean age, 46.13 ± 8.25) were ultimately used for final analysis (Figure 1). Most participants had completed high school (87.4%) and were nonsmokers (70.7%). Almost all participants were married (93.2%). Health-related information revealed that the percentage of participants who reported a family history of diabetes, a family history of CVD, a family history of hypertension, and a family history of stroke were 25.0%, 22.6%, 48.3%, and 10.4%, respectively. A total of 1,489 participants (30.2%) had MetS, and 543 participants (11.0%) experienced psychological stress. The characteristics of all participants are shown in Table 1.

### Descriptive data and comparison of all variables in participants with and without MetS by sex

According to the CDS criteria, the percentage of participants who reported hypertension, overweight or obesity, hyperglycemia, and dyslipidemia were 39.8%, 50.0%, 41.5%, and 44.2%, respectively. Overall, the prevalence of MetS among participants was 30.2%. Notably, MetS was present in 1,345 (38.9%) and 144 (9.7%) men and women, respectively ($p< 0.001$). There were significant differences in CPSS, SAS, and PSQI scores between participants with and without MetS in the total population ($p< 0.001$). Compared to participants without MetS, those with MetS

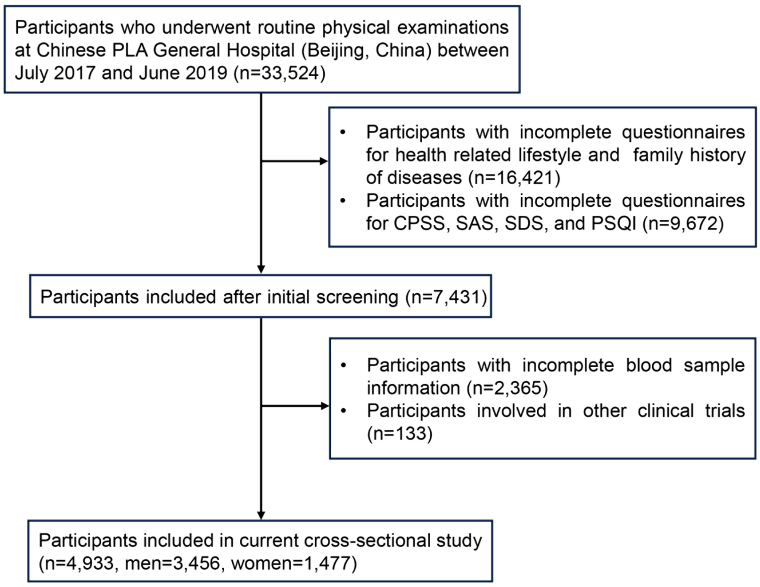



FIGURE 1
Flow chart for the selection of participants in the current cross-sectional study. Chinese PLA General Hospital, Chinese People's Liberation Army General Hospital; CPSS, Chinese Perceived Stress Scale; SAS, Self-Rating Anxiety Scale; SDS, Self-Rating Depression Scale; PSQI, Pittsburgh Sleep Quality Index.

TABLE 1 Characteristics of study population with and without MetS, shown by sex.

| Variables | Total (n = 4933) | | | Men (n = 3456) | | | Women (n = 1477) | | |
|---|---|---|---|---|---|---|---|---|---|
| | MetS (1489, 30.2%) | Non-MetS (3444, 69.8%) | $t/z/\chi^2$ (*P*) | MetS (1345, 38.9%) | Non-MetS (2111, 61.1%) | $t/z/\chi^2$ (*P*) | MetS (144, 9.7%) | Non-MetS (1333, 90.3%) | $t/z/\chi^2$ (*P*) |
| Men, n (%) | 1345 (90.3) | 2111 (61.3) | 417.764 (<0.001) | – | – | – | – | – | – |
| Age (mean ± SD) | 48.33 ± 7.23 | 45.17 ± 8.48 | 13.334 (<0.001) | 48.07 ± 7.15 | 45.66 ± 8.18 | 9.146 (<0.001) | 50.75±7.58 | 44.41 ± 8.90 | 9.365 (<0.001) |
| BMI, n (%) | | | 1425.443 (<0.001) | | | 809.457 (<0.001) | | | 346.645 (<0.001) |
| < 25 | 136 (9.1) | 2331 (67.7) | | 105 (7.8) | 1177 (55.8) | | 31 (21.5) | 1154 (86.6) | |
| ≥ 25 | 1353 (90.9) | 1113 (32.3) | | 1240 (92.2) | 934 (44.2) | | 113 (78.5) | 179 (13.4) | |
| Educational attainment, n (%) | | | 2.622 (0.269) | | | 3.191 (0.203) | | | 8.365 (0.015) |
| Less than high school | 190 (12.8) | 429 (12.5) | | 162 (12.0) | 262 (12.4) | | 28 (19.4) | 167 (12.5) | |
| High school | 957 (64.3) | 2290 (66.5) | | 880 (65.4) | 1427 (67.6) | | 77 (53.5) | 863 (64.7) | |
| College degree or more | 342 (23.0) | 725 (21.1) | | 303 (22.5) | 422 (20.0) | | 39 (27.1) | 303 (22.7) | |
| Marital status, n (%) | | | 41.737 (<0.001) | | | 27.399 (<0.001) | | | 3.537 (0.171) |
| Unmarried | 20 (1.3) | 181 (5.3) | | 16 (1.2) | 92 (4.4) | | 4 (2.8) | 89 (6.7) | |
| Married | 1433 (96.2) | 3166 (91.9) | | 1298 (96.5) | 1976 (93.6) | | 135 (93.8) | 1190 (89.3) | |
| Divorced or widowed | 36 (2.4) | 97 (2.8) | | 31 (2.3) | 43 (2.0) | | 5 (3.5) | 54 (4.1) | |

*(Continued)*

TABLE 1 Continued

| Variables | Total (n = 4933) | | | Men (n = 3456) | | | Women (n = 1477) | | |
|---|---|---|---|---|---|---|---|---|---|
| | MetS (1489, 30.2%) | Non-MetS (3444, 69.8%) | $t/z/\chi^2$ (P) | MetS (1345, 38.9%) | Non-MetS (2111, 61.1%) | $t/z/\chi^2$ (P) | MetS (144, 9.7%) | Non-MetS (1333, 90.3%) | $t/z/\chi^2$ (P) |
| Smoking, n (%) | 573 (38.5) | 870 (25.3) | 87.801 (<0.001) | 563 (41.9) | 838 (39.7) | 1.593 (0.207) | 10 (6.9) | 32 (2.4) | 9.712 (0.002) |
| Alcohol consumption, n (%) | 1037 (69.6) | 1705 (49.5) | 170.757 (<0.001) | 1019 (75.8) | 1445 (68.5) | 21.458 (<0.001) | 18 (12.5%) | 260 (19.5%) | 4.174 (0.041) |
| Physical activity, n (%) | | | | | | | | | |
| ≥ 2 hours/week | 696 (46.7) | 1694 (49.2) | 2.487 (0.115) | 603 (44.8) | 893 (42.3) | 2.143 (0.134) | 93 (64.6) | 801 (60.1) | 1.098 (0.295) |
| < 2 hours/week | 793 (53.3) | 1750 (50.8) | | 742 (55.2) | 1218 (57.7) | | 51 (35.4) | 532 (39.9) | |
| Waist-to-hip ratio (IQR) | 0.96 (0.93, 0.99) | 0.89 (0.81, 0.93) | 32.820 (<0.001) | 0.96 (0.93, 0.99) | 0.93 (0.90, 0.95) | 23.494 (<0.001) | 0.84 (0.82, 0.86) | 0.80 (0.75, 0.83) | 13.405 (<0.001) |
| Family history of diabetes, n (%) | 462 (31.0) | 773 (22.4) | 40.802 (<0.001) | 427 (31.7) | 466 (22.1) | 40.110 (<0.001) | 35 (24.3) | 307 (23.0) | 0.119 (0.731) |
| Family history of hypertension, n (%) | 870 (58.4) | 1514 (44.0) | 87.138 (0.001) | 784 (58.3) | 948 (44.9) | 58.852 (<0.001) | 86 (59.7) | 566 (42.5) | 15.705 (<0.001) |
| Family history of CVD, n (%) | 356 (23.9) | 761(22.1) | 1.949 (0.163) | 308 (22.9) | 439 (20.8) | 2.146 (0.143) | 48 (33.3) | 322 (24.2) | 5.830 (0.016) |
| Family history of stroke, n (%) | 184 (12.4) | 331 (9.6) | 8.386 (0.004) | 173 (12.9) | 212 (10.0) | 6.599 (0.010) | 11 (7.6) | 119 (8.9) | 0.269 (0.604) |
| CRP (IQR) | 0.12 (0.06, 0.22) | 0.08 (0.05, 0.14) | 10.556 (<0.001) | 0.11 (0.06, 0.23) | 0.09 (0.05, 0.15) | 7.673 (<0.001) | 0.15 (0.07, 0.26) | 0.08 (0.05, 0.13) | 7.024 (<0.001) |
| CPSS, n (%) | | | 14.861 (<0.001) | | | 2.776 (0.096) | | | 1.288 (0.256) |
| < 29 | 1364 (91.6) | 3026 (87.9) | | 1238 (92.0) | 1908 (90.4) | | 126 (87.5) | 1118 (83.9) | |
| ≥ 29 | 125 (8.4) | 418 (12.1) | | 107 (8.0) | 203 (9.6) | | 18 (12.5) | 215 (16.1) | |
| CPSS (IQR) | 17.0 (12.0, 22.0) | 18.0 (13.0, 24.0) | 3.561 (<0.001) | 17.0 (12.0, 22.0) | 18.0 (13.0, 23.0) | 1.899 (0.058) | 20.0 (13.0, 25.0) | 19.0 (14.0, 25.0) | 0.239 (0.811) |
| SAS, n (%) | | | 6.201 (0.013) | | | 1.484 (0.223) | | | 4.571 (0.033) |
| < 50 | 1284 (86.2) | 2873 (83.4) | | 1182 (87.9) | 1825 (86.5) | | 102 (70.8) | 1048 (78.6) | |
| ≥ 50 | 205 (13.8) | 571 (16.6) | | 163 (12.1) | 286 (13.5) | | 42(29.2) | 285 (21.4) | |
| SDS, n (%) | | | 0.009 (0.923) | | | 2.521 (0.112) | | | 5.644 (0.018) |
| < 50 | 1006 (67.6) | 2322 (67.4) | | 934 (69.4) | 1519 (72.0) | | 72 (50.0) | 803 (60.2) | |
| ≥ 50 | 483 (32.4) | 1122 (32.6) | | 411 (30.6) | 592 (28.0) | | 72 (50.0) | 530 (39.8) | |
| PSQI, n (%) | | | 4.171 (0.041) | | | 0.438 (0.508) | | | 0.084 (0.772) |
| ≤5 | 553 (37.1) | 1175 (34.1) | | 512 (38.1) | 780 (36.9) | | 41 (28.5) | 395 (29.6) | |
| >5 | 936 (62.9) | 2269 (65.9) | | 833 (61.9) | 1331 (63.1) | | 103 (71.5) | 938 (70.4) | |

MetS, metabolic syndrome; SD, standard deviation; BMI, Body Mass Index; IQR, interquartile range; CVD, cardiovascular diseases; CRP, C-reactive protein; CPSS, Chinese Perceived Stress Scale; SAS, Self-Rating Anxiety Scale; SDS, Self-Rating Depression Scale; PSQI, Pittsburgh Sleep Quality Index.

were older ($p< 0.001$), had higher rates of smoking ($p< 0.001$) and alcohol consumption ($p< 0.001$), higher CRP values ($p< 0.001$), and family histories of diabetes, hypertension, and stroke ($p< 0.001$) in both sexes. In addition, the prevalence of participants with low educational attainment and a family history of CVD was higher in women with MetS than in those without MetS. Further information is provided in Table 1.

### Descriptive data and comparison of all variables in participants with and without psychological stress by sex

As shown in Table 2, the prevalence of psychological stress (11.0%) in women (15.8%) was higher than that in men (9.0%) ($p< 0.001$). For MetS and its components, there were significant differences in MetS, hyperglycemia, overweight or obesity, dyslipidemia, SBP, DBP, FBG, TG, and HDL between individuals with and without psychological stress in the total population, but not in subgroup analysis by sex ($p< 0.05$). For the potential confounding factors, compared to participants without psychological stress, those with psychological stress had significant differences in age, marital status, waist-to-hip ratio, SAS, SDS, and PSQI ($p< 0.05$). Further information is shown in Table 2.

### Associations among psychological stress and MetS and its components

Logistic regression models with MetS and its risk components as dependent variables were used to assess whether psychological stress was associated with MetS, overweight or obesity, hypertension, hyperglycemia, and dyslipidemia, after adjusting for potential confounding factors (age, marital status, smoking, alcohol consumption, physical activity, SAS, SDS, family history of diabetes, and family history of hypertension) selected via Lasso. The results indicated that psychological stress was linked to the risk of hypertension (odds ratio (OR), 1.341 (95% confidence interval (CI), 1.023–1.758); p = 0.034) in men (Table 3; Figure 2). In contrast, psychological stress was not associated with MetS or the three other components. Further information is provided in Table 3; Figure 2. Additionally, logistic regression models with psychological stress as dependent variables were used to assess whether MetS and its individual risk components were independent risk factors for psychological stress. These models were adjusted for age, marital status, smoking, alcohol consumption, physical activity, SAS, and SDS, which were also selected via Lasso. The results indicated that hypertension could be an independent risk factor in total participants and men (total population: OR, 1.327 (95% CI, 1.025–1.718); $p = 0.032$; men: OR, 1.545 (95% CI, 1.113–2.145); $p = 0.009$). Further information is shown in Table 4; Figure 2.

## MR analysis

### The causal effect of psychological stress on MetS and its components

Among the 40 psychological stress-associated variants ($p< 1 \times 10^{-5}$, LD $r^2< 0.001$) (Supplementary Table S2), two SNPs were not available in the summary-level datasets of MetS and hypertension, 21 SNPs were unavailable for the overweight dataset, 17 SNPs were unavailable for the obesity dataset, 20 SNPs were unavailable for the BMI dataset, and 23 SNPs were unavailable for the hyperlipidemia dataset and HDL-C dataset. In addition, owing to incompatible alleles and ambiguous palindromes, we excluded two variants of MetS, hypertension, overweight, obesity, BMI, hyperlipidemia, TG, FBG, and HDL-C. Therefore, we included 36, 36, 17, 21, 18, 38, 15, 15, and 38 variants as IVs for MetS, hypertension, overweight, obesity, BMI, FBG, hyperlipidemia, HDL-C, and TG levels, respectively, in the MR analyses.

The causations were analyzed using IVW, MR-Egger, WME, weighted mode, simple mode, and MR-PRESSO methods. As depicted in Supplementary Table S3; Figure 3; Supplementary Figure S2, the ORs with 95% CIs for each log-odd increment in genetically predicted causal associations between psychological stress and MetS were obtained using the IVW method (OR, 0.989 (95% CI, 0.853–1.146); $p = 0.226$). These findings were consistent with the results from the five other models. The results of the MR-Egger intercept ($p = 0.689$) and MR-PRESSO global tests ($p = 0.151$) showed no indication of potential horizontal pleiotropy. The Cochran's $Q$ value for the IVW model was $p = 0.023$, but the funnel plot considered no significant heterogeneity obtained from individual variants (Supplementary Figure S3). Moreover, leave-one-out analysis showed that no IVs influenced this causal inference after gradually eliminating any single SNP (Supplementary Figure S4). Similarly, no causal relationship was found between psychological stress and the MetS components. The results of the MR-Egger regression analyses, MR-PRESSO global tests, Cochran's $Q$ value of the IVW model, funnel plot, and leave-one-out analyses for MetS components are shown in Supplementary Table S3; Figure 3; Supplementary Figures S2-S5. Most IVs had an F statistic greater than 10, indicating that IV bias was unlikely to exist. The statistical power for MR of psychological stress on MetS and its components was higher than 75% (Supplementary Table S4).

### The causal effect of MetS and its components on psychological stress

In the reverse MR analysis, after excluding the SNPs for palindromic alleles, palindromic alleles with intermediate allele frequencies, and unavailable SNPs in the summary-level dataset of psychological stress, we utilized 68, 66, 14, 13, 37, 11, 69, 31, and 94 variants for MetS, hypertension, overweight, obesity, BMI, hyperlipidemia, HDL-C, TG, and FBG as IVs ($p< 5 \times 10^{-8}$, LD $r^2< 0.001$), respectively (Supplementary Tables S5-S13).

As shown in Figure 4, Supplementary Table S14; Supplementary Figure S6, the MR results showed that hypertension and psychological stress had a positive causal relationship in the IVW model (OR, 2.386 (95% CI, 1.209–4.710); $p = 0.012$), which was in line with the results of the WME, simple mode, weighted mode, and MR-PRESSO models. No potential horizontal pleiotropy was observed in the MR-Egger intercept test ($p = 0.330$) or the MR-PRESSO global test ($p = 0.051$). The Cochran's $Q$ value for the IVW method indicated that heterogeneity may exist ($p = 0.021$); however, the symmetry of the

10.3389/fendo.2023.1212647

TABLE 2 Characteristics of study population according to the presence of psychological stress, shown by sex.

| Variable | Total (4933) | | | Men (3456, 70.1%) | | | Women (1477, 29.9%) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Stressed (543, 11.0%) | Non-stressed (4390, 89.0%) | $t/z/\chi^2$ (*P*) | Stressed (310, 9.0%) | Non-stressed (3146, 91.0%) | $t/z/\chi^2$ (*P*) | Stressed (233, 15.8%) | Non-stressed (1244, 84.2%) | $t/z/\chi^2$ (*P*) |
| Men, n (%) | 310 (57.1 %) | 3146 (71.7%) | 48.921 (<0.001) | – | – | – | – | – | – |
| Age (mean ± SD) | 42.74 ± 9.04 | 46.55 ± 8.05 | 9.372 (<0.001) | 43.17 ± 8.70 | 46.93 ±7.71 | 7.346 (<0.001) | 42.16 ± 9.46 | 45.56 ± 8.79 | 5.091 (<0.001) |
| BMI, n (%) | | | 23.644 (<0.001) | | | 4.924 (0.026) | | | 0.824 (0.364) |
| < 25 | 325 (59.9) | 2142 (48.8) | | 133 (42.9) | 1149 (36.5) | | 192 (82.4) | 993 (79.8) | |
| ≥ 25 | 218 (40.1) | 2248(51.2) | | 177 (57.1) | 1997 (63.5) | | 41 (17.6) | 251 (20.2) | |
| Educational attainment, n (%) | | | 5.626 (0.060) | | | 3.317 (0.190) | | | 1.521 (0.467) |
| Less than high school | 72 (13.3) | 547 (12.5) | | 41 (13.2) | 383 (12.2) | | 31 (13.3) | 164 (13.2) | |
| High school | 334 (61.5) | 2913 (66.4) | | 193 (62.3) | 2114 (67.2) | | 141 (60.5) | 799 (64.2) | |
| College degree or more | 137 (25.2) | 930 (21.2) | | 76 (24.5) | 649 (20.6) | | 61 (26.2) | 281 (22.6) | |
| Marital status, n (%) | | | 44.252 (<0.001)) | | | 27.529 (<0.001) | | | 11.175 (0.004) |
| Unmarried | 51 (9.4) | 150 (3.4) | | 25 (8.1) | 83 (2.6) | | 26 (11.2) | 67(5.4) | |
| Married | 477 (87.8) | 4122 (93.9) | | 278 (89.7) | 2996 (95.2) | | 199 (85.4) | 1126 (90.5) | |
| Divorced or widowed | 15 (2.8) | 118 (2.7) | | 7 (2.3) | 67 (2.1) | | 8 (3.4) | 51 (4.1) | |
| Smoking, n (%) | 167 (30.8) | 1276 (29.1) | 0.666 (0.414) | 155 (50.0) | 1246 (39.6) | 12.648 (<0.001) | 12 (5.2) | 30 (2.4) | 5.328 (0.021) |
| Alcohol consumption, n (%) | 253 (46.6) | 2489 (56.7) | 19.983 (<0.001) | 205 (66.1) | 2259 (71.8) | 4.443 (0.035) | 48 (20.6) | 230 (18.5) | 0.573 (0.449) |
| Physical activity, n (%) | 227 (41.8) | 2163 (49.3) | 10.785 (0.001) | 98 (31.6) | 1398 (44.4) | 18.905 (<0.001) | 129 (55.4) | 765 (61.5) | 3.087 (0.079) |
| Waist-to-hip ratio (IQR) | 0.87 (0.80, 0.93) | 0.92 (0.84, 0.96) | 8.902 (<0.001) | 0.93 (0.90, 0.96) | 0.94 (0.92, 0.97) | 4.605 (<0.001) | 0.80 (0.74, 0.82) | 0.81 (0.77, 0.83) | 3.020 (0.003) |
| Family history of diabetes, n (%) | 139 (25.6) | 1096 (25.0) | 0.103 (0.748) | 79 (25.5) | 814 (25.9) | 0.022 (0.881) | 60 (25.8) | 282 (22.7) | 1.048 (0.306) |
| Family history of hypertension, n (%) | 252 (46.4) | 2132 (48.6) | 0.900 (0.343) | 147 (47.4) | 1585 (50.4) | 0.990 (0.320) | 105 (45.1) | 547 (44.0) | 0.095 (0.758) |
| Family history of CVD, n (%) | 129 (23.8) | 988 (22.5) | 0.432 (0.511) | 74 (23.9) | 673 (21.4) | 1.023 (0.312) | 55 (23.6) | 315 (25.3) | 0.308 (0.579) |
| Family history of stroke, n (%) | 50 (9.2) | 465 (10.6) | 0.990 (0.320) | 30 (9.7) | 355 (11.3) | 0.736 (0.391) | 20 (8.6) | 110 (8.8) | 0.016 (0.898) |
| CRP (IQR) | 0.09 (0.05, 0.15) | 0.09 (0.05, 0.16) | 0.304 (0.761) | 0.10 (0.05, 0.19) | 0.10 (0.05, 0.18) | 0.447 (0.655) | 0.08 (0.05, 0.13) | 0.08 (0.05, 0.15) | 0.371 (0.752) |
| MetS, n (%) | 125 (23.0) | 1364 (31.1) | 14.861 (<0.001) | 107 (34.5) | 1238 (39.4) | 2.776 (0.096) | 18 (7.7) | 126 (10.1) | 1.288 (0.256) |
| Hypertension, n (%) | 205 (37.8) | 1756 (40.0) | 1.019 (0.313) | 154 (49.7) | 1472 (46.8) | 0.945 (0.331) | 51 (21.9) | 284 (22.8) | 0.099 (0.753) |
| Hyperglycemia, n (%) | 193 (35.5) | 1854 (42.2) | 8.906 (0.003) | 127 (41.0) | 1421 (45.2) | 2.014 (0.156) | 66 (28.3) | 433 (34.8) | 0.685 (0.055) |
| Overweight or obesity, n (%) | 218 (40.1) | 2248 (51.2) | 23.644 (<0.001) | 177 (57.1) | 1997 (63.5) | 4.924 (0.026) | 41 (17.6) | 251 (20.2) | 0.824 (0.364) |

*(Continued)*

TABLE 2 Continued

| Variable | Total (4933) | | | Men (3456, 70.1%) | | | Women (1477, 29.9%) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Stressed (543, 11.0%) | Non-stressed (4390, 89.0%) | $t/z/\chi^2$ (*P*) | Stressed (310, 9.0%) | Non-stressed (3146, 91.0%) | $t/z/\chi^2$ (*P*) | Stressed (233, 15.8%) | Non-stressed (1244, 84.2%) | $t/z/\chi^2$ (*P*) |
| Dyslipidemia, n (%) | 204 (37.6) | 1975 (45.0) | 10.787 (0.001) | 162 (52.3) | 1717 (54.6) | 0.612 (0.434) | 42 (18.0) | 258 (20.7) | 0.893 (0.345) |
| SBP (IQR) | 117.0 (105.0, 132.0) | 122.0 (109.0, 136.0) | 4.701 (<0.001) | 123.0 (110.0, 137.0) | 126.0 (113.0, 138.0) | 1.182 (0.237) | 109.0 (98.0, 122.0) | 112.0 (101.0, 128.0) | 3.000 (<0.001) |
| DBP (IQR) | 79.0 (71.0, 89.0) | 81.0 (74.0, 89.0) | 3.123 (<0.001) | 83.0 (75.0, 92.0) | 82.0 (75.0, 90.0) | 0.884 (0.377) | 74.0 (67.0, 82.0) | 78.0 (71.0, 85.0) | 3.880 (<0.001) |
| FBG (mmol/L, IQR) | 5.09 (4.79, 5.55) | 5.24 (4.88, 5.74) | 3.796 (<0.001) | 5.28 (4.94, 5.83) | 5.34 (4.96, 5.92) | 1.465 (0.143) | 4.96 (4.63, 5.30) | 5.00 (4.71, 5.32) | 1.323 (0.186) |
| Triglycerides (mmol/L, IQR) | 1.30 (0.91, 2.05) | 1.53 (1.05, 2.29) | 4.230 (<0.001) | 1.68 (1.20, 2.50) | 1.73 (1.22, 2.53) | 0.482 (0.630) | 1.00 (0.75, 1.37) | 1.08 (0.80, 1.52) | 2.454 (0.014) |
| HDL (mmol/L, IQR) | 1.29 (1.06, 1.53) | 1.20 (1.00, 1.46) | 3.717 (<0.001) | 1.14 (0.95, 1.32) | 1.13 (0.96, 1.33) | 0.439 (0.661) | 1.48 (1.25, 1.76) | 1.48 (1.23, 1.74) | 0.812 (0.417) |
| LDL (mmol/L, IQR) | 3.03 (2.44, 3.66) | 3.11 (2.54, 3.69) | 1.325 (0.185) | 3.20 (2.57, 3.75) | 3.11 (2.54, 3.70) | 0.896 (0.370) | 2.94 (2.36, 3.60) | 3.12 (2.55, 3.70) | 3.097 (0.002) |
| CHO (mmol/L, IQR) | 4.61 (3.99, 5.19) | 4.70 (4.11, 5.31) | 1.910 (0.056) | 4.63 (4.07, 5.30) | 4.69 (4.09, 5.31) | 0.008 (0.994) | 4.64 (3.94, 5.15) | 4.75 (4.15, 5.33) | 3.084 (0.002) |
| SAS, n (%) | | | 745.856 (<0.001) | | | 427.120 (<0.001) | | | 269.124 (<0.001) |
| < 50 | 239 (44.0) | 3918 (89.2) | | 153 (49.4) | 2854 (90.7) | | 86 (36.9) | 1064 (85.5) | |
| ≥ 50 | 304 (56.0) | 472 (10.8) | | 157 (50.6) | 292 (9.3) | | 147 (63.1) | 180 (14.5) | |
| SDS, n (%) | | | 735.605 (<0.001) | | | 457.243 (<0.001) | | | 246.306 (<0.001) |
| < 50 | 87 (16.0) | 3241 (73.8) | | 57 (18.4) | 2396 (76.2) | | 30 (12.9) | 845 (67.9) | |
| ≥ 50 | 456 (84.0) | 1149 (26.2) | | 253 (81.6) | 750 (23.8) | | 203 (87.1) | 399 (32.1) | |
| PSQI, n (%) | | | 161.350 (<0.001) | | | 91.845 (<0.001) | | | 60.693 (<0.001) |
| ≤5 | 57 (10.5) | 1671 (38.1) | | 38 (12.3) | 1254 (39.9) | | 19 (8.2) | 417 (33.5) | |
| >5 | 486 (89.5) | 2719 (61.9) | | 272 (87.7) | 1892 (60.1) | | 214 (91.8) | 827 (66.5) | |

MetS, metabolic syndrome; SD, standard deviation; BMI, Body Mass Index; IQR, interquartile range; CVD, cardiovascular diseases; CHO, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; FBG, fasting blood-glucose; TG, Triglycerides; CRP, C-reactive protein; SBP, systolic blood pressure; DBP, diastolic blood pressure; CPSS, Chinese Perceived Stress Scale; SAS, Self-Rating Anxiety Scale; SDS, Self-Rating Depression Scale; PSQI, Pittsburgh Sleep Quality Index.

funnel plot showed no evidence of heterogeneity (Supplementary Figure S7). Furthermore, leave-one-out analysis suggested that the MR results were stable after the removal of any single SNP. Nonetheless, neither MetS nor its five other factors were causally related to psychological stress. Further information is presented in Supplementary Tables S4, S14; Figure 4; Supplementary Figures S6-S9.

## Discussion

MetS has been recognized as a serious health problem worldwide because of its growing prevalence (3). According to previous studies, the association between psychological stress and MetS remains unclear. In this study, we used a cross-sectional design to investigate the association of psychological stress with MetS and its risk components and used bi-directional MR analyses to explore its causal relationship. We found that psychological stress was associated with hypertension in men after controlling for potential covariates in the present cross-sectional study but not in MR analyses; conversely, hypertension was a risk factor for psychological stress in cross-sectional and MR analyses.

Psychological stress and MetS are associated with alterations in CVD; however, their relationship has not yet been fully elucidated. To reduce the limitations of observational studies, such as the disturbance of confounding effects, we used MR analysis, a scientific method, to explore the relationship between psychological stress and MetS. In the present study, we found no association between psychological stress and MetS, and similar results were obtained

10.3389/fendo.2023.1212647

TABLE 3 Multivariate analysis of psychological stress on MetS and its risk components, shown by sex.

| Variables | Total | | Men | | Women | |
|---|---|---|---|---|---|---|
| | OR (95% CI) | *p*-value | OR (95% CI) | *p*-value | OR (95% CI) | *p*-value |
| MetS | 0.823 (0.644, 1.052) | 0.120 | 0.921 (0.696, 1.219) | 0.565 | 0.630 (0.347, 1.144) | 0.129 |
| Hypertension | 1.140 (0.916, 1.420) | 0.239 | 1.341 (1.023, 1.758) | 0.034 | 0.919 (0.613, 1.378) | 0.683 |
| Overweight or obesity | 0.786 (0.601, 1.029) | 0.082 | 0.786 (0.601, 1.032) | 0.083 | 0.813 (0.532, 1.241) | 0.337 |
| Dyslipidemia | 0.816 (0.655, 1.016) | 0.069 | 0.868 (0.666, 1.130) | 0.293 | 0.762 (0.500, 1.161) | 0.206 |
| TG (≥ 1.7, mmol/L) | 0.828 (0.665, 1.031) | 0.092 | 0.899 (0.690, 1.171) | 0.429 | 0.743 (0.478, 1.156) | 0.188 |
| HDL-C (< 0.9 in men,< 1.0 in women, mmol/L) | 0.798 (0.590, 1.080) | 0.144 | 0.839 (0.594, 1.185) | 0.319 | 0.728 (0.382, 1.386) | 0.333 |
| Hyperglycemia | 1.006 (0.804, 1.260) | 0.957 | 1.131 (0.851, 1.503) | 0.396 | 0.872 (0.598, 1.271) | 0.476 |
| FBG (≥ 6.1, mmol/L) | 0.930 (0.673, 1.285) | 0.659 | 0.950 (0.662, 1.364) | 0.782 | 1.076 (0.477, 2.428) | 0.859 |

All associations were tested using logistic regression, and all results of multivariate analysis were adjusted by age, marital status, smoking, alcohol consumption, physical activity, family history of diabetes, family history of hypertension, SAS, and SDS. MetS, metabolic syndrome; OR, odd ratio; CI, confidence interval; SAS, Self-Rating Anxiety Scale; SDS, Self-Rating Depression Scale.

from the MR analyses. In line with our findings, previous cross-sectional and longitudinal studies have indicated no relationship between psychological stress and MetS, regardless of the instruments used to measure psychological pressure or the definition of MetS (40, 41). Considering salivary cortisol as an objective indicator of psychological stress, prior studies have indicated no significant difference in salivary cortisol levels between populations with and without MetS (13, 42, 43), thereby offering an interpretation of our results. Nevertheless, cross-sectional studies in Japan, Europe, and Pakistan have reported stress scores of 28, 25, and 31, respectively, observing a positive association between psychological stress and MetS (44–46). Indeed, increased psychological stress scores have been associated with an increased risk of metabolic disorders (9). Consistent with our results (mean CPSS score: 18.4), one prior cross-sectional study reporting a low stress score of 22.7 did not support the effect of stress on MetS (7).

Hypertension is a major modifiable risk factor for MetS and CVD. There is growing evidence for an association between hypertension and the progression of psychological stress (47, 48). Our cross-sectional and MR analyses revealed that hypertension may increase the risk of psychological stress. Prior research has found that hypertension causes damage to small blood vessels, contributing to neuronal damage in multiple areas, including the hippocampus, which could promote the development of psychological stress (49). One animal experiment showed that a highly activated region in the spontaneously hypertensive rat (the locus coeruleus) could awaken and regulate autonomic function

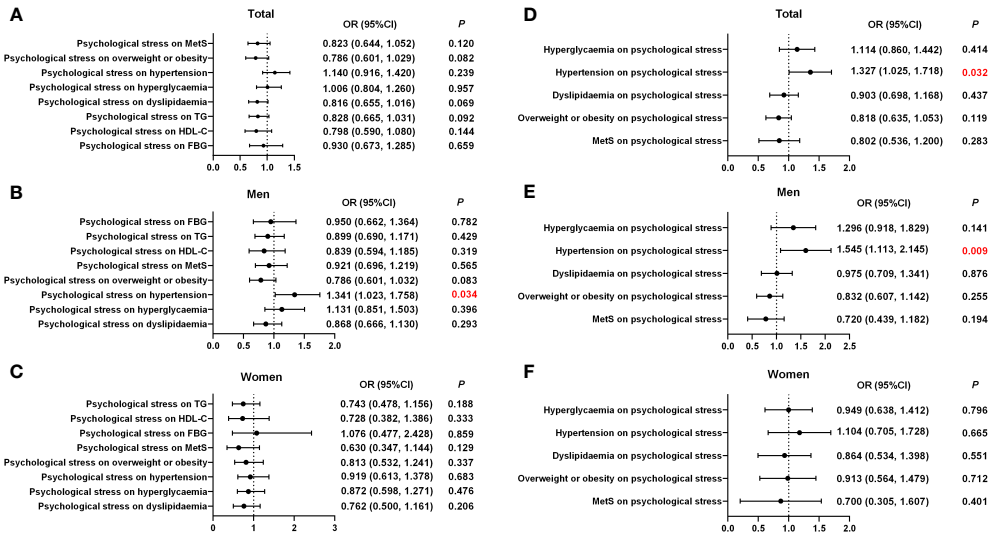



FIGURE 2
Associations of psychological stress with MetS and its components according to sex. **(A)** The effect of psychological stress on MetS and its components in the total population. **(B)** The effect of psychological stress on MetS and its components in men. **(C)** The effect of psychological stress on MetS and its components in women. **(D)** The effect of MetS and its components on psychological stress in the total population. **(E)** The effect of MetS and its components on psychological stress in men. **(F)** The effect of MetS and its components on psychological stress in women. OR, odd ratio; CI, confidence interval; MetS, metabolic syndrome.

10.3389/fendo.2023.1212647

TABLE 4 Multivariate analysis of MetS and its risk components on psychological stress, shown by sex.

| Variables | Total | | Men | | Women | |
|---|---|---|---|---|---|---|
| | OR (95% CI) | *p*-value | OR (95% CI) | *p*-value | OR (95% CI) | *p*-value |
| MetS | 0.802 (0.536, 1.200) | 0.283 | 0.720 (0.439, 1.182) | 0.194 | 0.700 (0.305, 1.607) | 0.401 |
| Hypertension | 1.327 (1.025, 1.718) | 0.032 | 1.545 (1.113, 2.145) | 0.009 | 1.104 (0.705,1.728) | 0.665 |
| Hyperglycemia | 1.114 (0.860, 1.442) | 0.414 | 1.296 (0.918, 1.829) | 0.141 | 0.949 (0.638, 1.412) | 0.796 |
| Overweight or obesity | 0.818 (0.635, 1.053) | 0.119 | 0.832 (0.607, 1.142) | 0.255 | 0.913 (0.564, 1.479) | 0.712 |
| Dyslipidemia | 0.903 (0.698, 1.168) | 0.437 | 0.975 (0.709, 1.341) | 0.876 | 0.864 (0.534, 1.398) | 0.551 |

All results of multivariate analysis were adjusted by age, marital status, smoking, alcohol consumption, physical activity, SAS, and SDS. MetS, metabolic syndrome; OR, odd ratio; CI, confidence interval; SAS, Self-Rating Anxiety Scale; SDS, Self-Rating Depression Scale.

and that enhanced autonomic reactivity is a true indicator of perceived stress levels (50). Therefore, it is particularly important to pay attention to the psychological stress experienced by patients with hypertension to reduce the occurrence of hypertension-related complications. Conversely, based on MR results, psychological stress may not be involved in the development of hypertension. In addition, our cross-sectional study found that psychological stress may be related to hypertension in men but found no association in women or the total population. Indeed, gender plays a role in influencing the aforementioned relationship. In the current cross-sectional survey, a higher prevalence of hypertension was observed in men (47.0%) than in women (22.7%), consistent with results reported in other published studies (51, 52). Research revealed that women tend to manifest emotions such as anxiety or depression more frequently, while men, under chronic stress conditions, are more likely to exhibit an elevated incidence of alcohol consumption and an increased risk of hypertension and MetS (53–55). The mechanisms underlying the relationship between psychosocial stress and hypertension are diverse and complex (56). Furthermore, many cross-sectional and cohort studies have reported that psychological stress is not involved in the progression of hypertension. Therefore, based on current evidence, we cannot conclude that psychological stress is a risk factor for hypertension in the general population (57, 58).

Regarding the relationship between psychological stress and overweight or obesity, hyperglycemia, and dyslipidemia, no significant association was observed in our cross-sectional and MR results, supporting the findings of previous cross-sectional and cohort studies (59, 60). However, several publications that additionally adjusted for the confounding effects of dietary behavior showed a significant relationship between psychological stress and the aforementioned factors (61, 62). Research related to behavioral

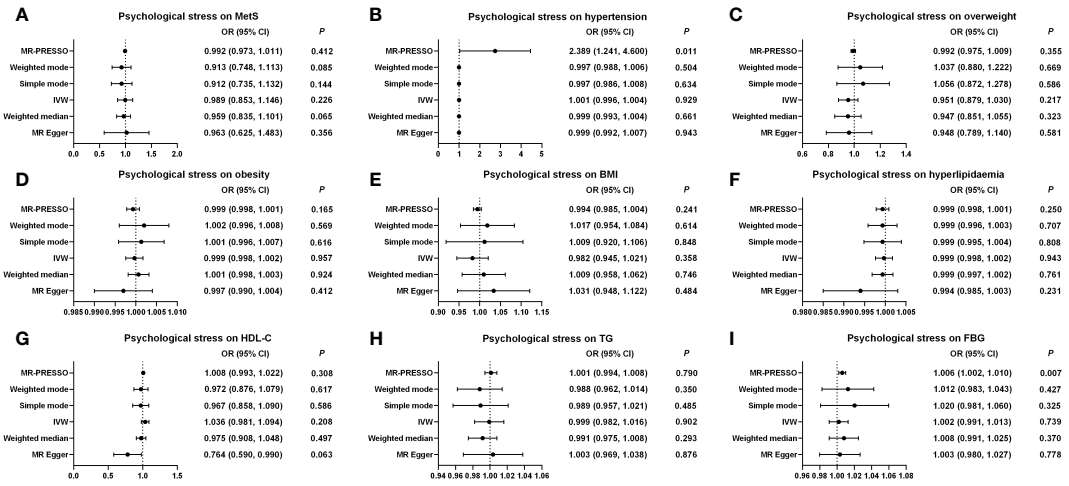



FIGURE 3
Causal estimates of genetically predicted psychological stress on MetS and its components. **(A)** Causal estimates of genetically predicted psychological stress on MetS. **(B)** Causal estimates of genetically predicted psychological stress on hypertension. **(C)** Causal estimates of genetically predicted psychological stress on overweight. **(D)** Causal estimates of genetically predicted psychological stress on obesity. **(E)** Causal estimates of genetically predicted psychological stress on BMI. **(F)** Causal estimates of genetically predicted psychological stress on hyperlipidemia. **(G)** Causal estimates of genetically predicted psychological stress on HDL-C. **(H)** Causal estimates of genetically predicted psychological stress on TG. **(I)** Causal estimates of genetically predicted psychological stress on FBG. MetS, metabolic syndrome; MR, Mendelian randomization; OR, odds ratio; CI, confidence interval; IVW, inverse-variance weighted; MR-PRESSO, MRPleiotropy Residual Sum and Outlier; BMI, body mass index; FBG, fasting blood-glucose; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides.

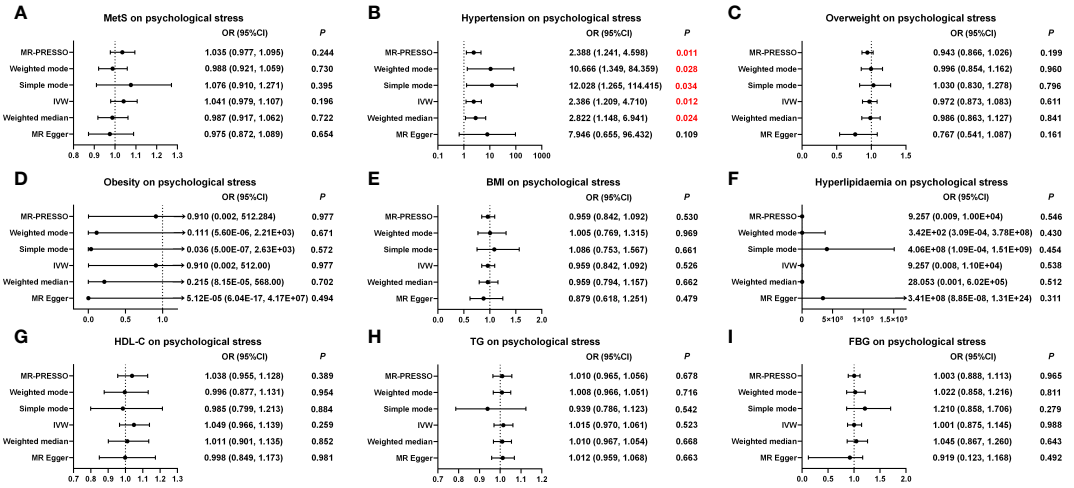



FIGURE 4
Causal estimates of genetically predicted MetS and its components on psychological stress. **(A)** Causal estimates of genetically predicted MetS on psychological stress. **(B)** Causal estimates of genetically predicted hypertension on psychological stress. **(C)** Causal estimates of genetically predicted overweight on psychological stress. **(D)** Causal estimates of genetically predicted obesity on psychological stress. **(E)** Causal estimates of genetically predicted BMI on psychological stress. **(F)** Causal estimates of genetically predicted hyperlipidemia on psychological stress. **(G)** Causal estimates of genetically predicted HDL-C on psychological stress. **(H)** Causal estimates of genetically predicted TG on psychological stress. **(I)** Causal estimates of genetically predicted FBG on psychological stress. MetS, metabolic syndrome; MR, Mendelian randomization; OR, odds ratio; CI, confidence interval; IVW, inverse-variance weighted; MR-PRESSO, Pleiotropy Residual Sum and Outlier; BMI, body mass index; FBG, fasting blood-glucose; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides.

psychology has indicated that high-income populations respond to high levels of psychological stress through physical activity, whereas some low-income populations are more likely to cope with it through compensatory eating (63). Due to limitations in data collection for this project, we did not include dietary habits as covariates in the regression analysis. Additionally, it is worth noting that most of the study population consisted of high-income populations, which could be one possible reason for the non-significant results. Furthermore, the range of CPSS scores in this current study cannot reflect the psychological stress of highly stressed individuals, potentially explaining the lack of a significant correlation (64).

## Strengths and limitations

This study had several limitations that should be considered. Firstly, compared to clinical diagnosis, the self-reported questionnaires (i.e., SDS, SAS, and PSQI) used in this cross-sectional study provided limited evidence. Secondly, due to limitations in data collection for this project, we did not include dietary habits as covariates in the regression analysis. Additionally, it is worth noting that most of the study populations consisted of high-income populations, which could be one possible reason for the nonsignificant results. Furthermore, the range of CPSS scores in this current study cannot reflect the psychological stress of highly stressed individuals, potentially explaining the lack of a significant correlation. Moreover, the cross-sectional study design cannot avoid the influence of traditional confounding factors and inverse causal associations. The reason for the lack of detailed demographic information is that we did not perform subgroup analyses in the MR analyses. Finally, owing to data limitations, the current observational study in the Chinese population and the MR study in the European population both constrain the generalizability of our study results. The strengths of this study are as follows: the confounding effects of depression, anxiety, and sleep quality, which have rarely been accounted for in previous epidemiological studies, were adjusted using regression analysis in the current cross-sectional study (9). In MR analysis, we investigated the causal relationship between psychological stress and MetS and its components from a genetic perspective.

# Conclusion

In conclusion, our findings did not indicate a significant association between psychological stress and MetS. However, we observed associations between psychological stress and hypertension, with evidence that individuals with hypertension may be more susceptible to psychological stress. These findings may have implications for targeting factors related to hypertension and psychological stress in interventions aimed at improving mental and metabolic health. The relationship between psychological stress and MetS and its components requires further study and careful interpretation.

# 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 humans were approved by Medical Ethics Committee of Chinese People's Liberation Army General Hospital (S2019-131-01). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

YN and WG contributed to the design and supervision of this study. CL and TT participated in the design and planning process. YT, HZ, and HXL collected and compiled the data. CL and HML analyzed the data. CL and TT wrote the first draft of the manuscript. YN, XL, and TT revised the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was supported by the National Key R&D Programme of China (2017YFE0118800) and the National Natural Science Foundation of China (No. 82100265, No. 31971049).

## Acknowledgments

We sincerely thank all the researchers for sharing the GWAS-pooled data on psychological stress-related disorders and MetS and its components.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1212647/full#supplementary-material

## References

1. Gonzalez-Chávez A, Chávez-Fernández JA, Elizondo-Argueta S, González-Tapia A, León-Pedroza JI, Ochoa C. Metabolic syndrome and cardiovascular disease: A health challenge. *Arch Med Res* (2018) 49(8):516–21. doi: 10.1016/j.arcmed.2018.10.003

2. Battelli MG, Bortolotti M, Polito L, Bolognesi A. Metabolic syndrome and cancer risk: The role of xanthine oxidoreductase. *Redox Biol* (2019) 21:101070. doi: 10.1016/j.redox.2018.101070

3. Saklayen MG. The global epidemic of the metabolic syndrome. *Curr Hypertens Rep* (2018) 20(2):12. doi: 10.1007/s11906-018-0812-z

4. Lu J, Wang L, Li M, Xu Y, Jiang Y, Wang W, et al. Metabolic syndrome among adults in China: the 2010 China noncommunicable disease surveillance. *J Clin Endocrinol Metab* (2017) 102(2):507–15. doi: 10.1210/jc.2016-2477

5. Haykin H, Rolls A. The neuroimmune response during stress: A physiological perspective. *Immunity* (2021) 54(9):1933–47. doi: 10.1016/j.immuni.2021.08.023

6. Kivimäki M, Steptoe A. Effects of stress on the development and progression of cardiovascular disease. *Nat Rev Cardiol* (2018) 15(4):215–29. doi: 10.1038/nrcardio.2017.189

7. Chen YJ, Lin CL, Li CR, Huang SM, Chan JY, Fang WH, et al. Associations among integrated psychoneuroimmunological factors and metabolic syndrome. *Psychoneuroendocrinology* (2016) 74:342–9. doi: 10.1016/j.psyneuen.2016.09.027

8. Tenk J, Mátrai P, Hegyi P, Rostás I, Garami A, Szabó I, et al. Perceived stress correlates with visceral obesity and lipid parameters of the metabolic syndrome: A systematic review and meta-analysis. *Psychoneuroendocrinology* (2018) 95:63–73. doi: 10.1016/j.psyneuen.2018.05.014

9. Kuo WC, Bratzke LC, Oakley LD, Kuo F, Wang H, Brown RL. The association between psychological stress and metabolic syndrome: A systematic review and meta-analysis. *Obes Rev* (2019) 20(11):1651–64. doi: 10.1111/obr.12915

10. Chen WL, Wang CC, Chiang ST, Wang YC, Sun YS, Wu WT, et al. The impact of occupational psychological hazards and metabolic syndrome on the 8-year risk of cardiovascular diseases-A longitudinal study. *PloS One* (2018) 13(8):e0202977. doi: 10.1371/journal.pone.0202977

11. Loerbroks A, Shang L, Angerer P, Li J. Effort-reward imbalance at work increases the risk of the metabolic syndrome: a prospective study in Chinese university staff. *Int J Cardiol* (2015) 182:390–1. doi: 10.1016/j.ijcard.2015.01.030

12. Räikkönen K, Matthews KA, Kuller LH. The relationship between psychological risk attributes and the metabolic syndrome in healthy women: antecedent or consequence? *Metabolism: Clin Exp* (2002) 51(12):1573–7. doi: 10.1053/meta.2002.36301

13. Abraham SB, Rubino D, Sinaii N, Ramsey S, Nieman LK. Cortisol, obesity, and the metabolic syndrome: a cross-sectional study of obese subjects and review of the literature. *Obes (Silver Spring Md.)* (2013) 21(1):E105–17. doi: 10.1002/oby.20083

14. Garcez A, Weiderpass E, Canuto R, Lecke SB, Spritzer PM, Pattussi MP, et al. Perceived stress, and metabolic syndrome: A matched case-control study in female shift workers. *Hormone Metab Res = Hormon- und Stoffwechselforschung = Hormones metabolisme* (2017) 49(7):510–9. doi: 10.1055/s-0043-101822

15. Yao S, Zhang M, Dong SS, Wang JH, Zhang K, Guo J, et al. Bidirectional two-sample Mendelian randomization analysis identifies causal associations between relative carbohydrate intake and depression. *Nat Hum Behav* (2022) 6(11):1569–76. doi: 10.1038/s41562-022-01412-9

16. Ebrahim S, Davey Smith G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum Genet* (2008) 123(1):15–33. doi: 10.1007/s00439-007-0448-6

17. Cho KI, Kim BH, Je HG, Jang JS, Park YH. Gender-specific associations between socioeconomic status and psychological factors and metabolic syndrome in the Korean population: findings from the 2013 Korean national health and nutrition examination survey. *BioMed Res Int* (2016) 2016:3973197. doi: 10.1155/2016/3973197

18. Guo XH, Zhang PH, Zeng ZC, Wang W, Li C, Shi Y, et al. Combination patterns of cardiovascular risks and sequelae at different stage of hypertension in natural hypertensive population in Beijing. *Clin Exp Hypertens (New York N.Y. 1993)* (2009) 31(2):142–55. doi: 10.1080/10641960802621309

19. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* (1983) 24(4):385–96. doi: 10.2307/2136404

20. Leung DY, Lam TH, Chan SS. Three versions of Perceived Stress Scale: validation in a sample of Chinese cardiac patients who smoke. *BMC Public Health* (2010) 10:513. doi: 10.1186/1471-2458-10-513

21. Zung WW. A rating instrument for anxiety disorders. *Psychosomatics* (1971) 12 (6):371–9. doi: 10.1016/s0033-3182(71)71479-0

22. Zung WW. A SELF-RATING DEPRESSION SCALE. *Arch Gen Psychiatry* (1965) 12:63–70. doi: 10.1001/archpsyc.1965.01720310065008

23. Zhang J, Xu CP, Wu HX, Xue XJ, Xu ZJ, Li Y, et al. Comparative study of the influence of diabetes distress and depression on treatment adherence in Chinese patients with type 2 diabetes: a cross-sectional survey in the People's Republic of China. *Neuropsychiatr Dis Treat* (2013) 9:1289–94. doi: 10.2147/ndt.S49798

24. Zhou J, Yang Y, Qiu X, Yang X, Pan H, Ban B, et al. Relationship between anxiety and burnout among Chinese physicians: A moderated mediation model. *PloS One* (2016) 11(8):e0157013. doi: 10.1371/journal.pone.0157013

25. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* (1989) 28(2):193–213. doi: 10.1016/0165-1781(89)90047-4

26. Tsai PS, Wang SY, Wang MY, Su CT, Yang TT, Huang CJ, et al. Psychometric evaluation of the Chinese version of the Pittsburgh Sleep Quality Index (CPSQI) in primary insomnia and control subjects. *Qual Life Res an Int J Qual Life Aspects Treatment Care Rehabil* (2005) 14(8):1943–52. doi: 10.1007/s11136-005-4346-x

27. Wei D, Chen T, Li J, Gao Y, Ren Y, Zhang X, et al. Association of serum gamma-glutamyl transferase and ferritin with the metabolic syndrome. *J Diabetes Res* (2015) 2015:741731. doi: 10.1155/2015/741731

28. Cheng L, Yan W, Zhu L, Chen Y, Liu J, Xu Y, et al. Comparative analysis of IDF, ATPIII and CDS in the diagnosis of metabolic syndrome among adult inhabitants in Jiangxi Province, China. *PloS One* (2017) 12(12):e0189046. doi: 10.1371/journal.pone.0189046

29. Lind L. Genome-wide association study of the metabolic syndrome in UK biobank. *Metab Syndrome Related Disord* (2019) 17(10):505–11. doi: 10.1089/met.2019.0070

30. Elsworth B, Mitchell R, Raistrick CA, Paternoster L, Hemani G, Gaunt TR. MRC IEU UK Biobank GWAS pipeline version 1. (2017). doi: 10.5523/bris.2fahpksont1zi26xosyamqo8rr

31. 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(6):840–60. doi: 10.1038/s41588-021-00852-9

32. 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

33. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* (2017) 46(6):1985–98. doi: 10.1093/ije/dyx102

34. 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

35. 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

36. 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

37. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *Int J Epidemiol* (2018) 47(4):1264–78. doi: 10.1093/ije/dyy101

38. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol* (2011) 40(3):740–52. doi: 10.1093/ije/dyq151

39. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* (2013) 42(5):1497–501. doi: 10.1093/ije/dyt179

40. Frisman GH, Kristenson M. Psychosocial status and health related quality of life in relation to the metabolic syndrome in a Swedish middle-aged population. *Eur J Cardiovasc Nurs* (2009) 8(3):207–15. doi: 10.1016/j.ejcnurse.2009.01.004

41. Ortega-Montiel J, Posadas-Romero C, Ocampo-Arcos W, Medina-Urrutia A, Cardoso-Saldaña G, Jorge-Galarza E, et al. Self-perceived stress is associated with adiposity and atherosclerosis. The GEA Study. *BMC Public Health* (2015) 15:780. doi: 10.1186/s12889-015-2112-8

42. Hellhammer DH, Wüst S, Kudielka BM. Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* (2009) 34(2):163–71. doi: 10.1016/j.psyneuen.2008.10.026

43. DeSantis AS, DiezRoux AV, Hajat A, Golden SH, Jenny NS, Sanchez BN, et al. Associations of salivary cortisol levels with metabolic syndrome and its components: the multi-ethnic study of atherosclerosis. *J Clin Endocrinol Metab* (2011) 96(11):3483–92. doi: 10.1210/jc.2011-0483

44. Mimura C, Griffiths P. A Japanese version of the Perceived Stress Scale: cross-cultural translation and equivalence assessment. *BMC Psychiatry* (2008) 8:85. doi: 10.1186/1471-244x-8-85

45. Remor E. Psychometric properties of a European Spanish version of the Perceived Stress Scale (PSS). *Spanish J Psychol* (2006) 9(1):86–93. doi: 10.1017/s1138741600006004

46. Shah M, Hasan S, Malik S, Sreeramareddy CT. Perceived stress, sources and severity of stress among medical undergraduates in a Pakistani medical school. *BMC Med Educ* (2010) 10:2. doi: 10.1186/1472-6920-10-2

47. Von Visger TT, Kuntz KK, Phillips GS, Yildiz VO, Sood N. Quality of life and psychological symptoms in patients with pulmonary hypertension. *Heart Lung J Crit Care* (2018) 47(2):115–21. doi: 10.1016/j.hrtlng.2017.12.004

48. Balint EM, Boseva P, Schury K, Guendel H, Rottbauer W, Waller C. High prevalence of posttraumatic stress in patients with primary hypertension. *Gen Hosp Psychiatry* (2016) 38:53–8. doi: 10.1016/j.genhosppsych.2015.10.002

49. Yang Y, Kimura-Ohba S, Thompson J, Rosenberg GA. Rodent models of vascular cognitive impairment. *Trans stroke Res* (2016) 7(5):407–14. doi: 10.1007/s12975-016-0486-2

50. Zubcevic J, Watkins J, Perez PD, Colon-Perez LM, Long MT, Febo M, et al. MEMRI reveals altered activity in brain regions associated with anxiety, locomotion, and cardiovascular reactivity on the elevated plus maze in the WKY vs SHR rats. *Brain Imaging Behav* (2018) 12(5):1318–31. doi: 10.1007/s11682-017-9798-4

51. Divney AA, Echeverria SE, Thorpe LE, Trinh-Shevrin C, Islam NS. Hypertension prevalence jointly influenced by acculturation and gender in US immigrant groups. *Am J Hypertens* (2019) 32(1):104–11. doi: 10.1093/ajh/hpy130

52. Defianna SR, Santosa A, Probandari A, Dewi FST. Gender differences in prevalence and risk factors for hypertension among adult populations: A cross-sectional study in Indonesia. *Int J Environ Res Public Health* (2021) 18(12):6259. doi: 10.3390/ijerph18126259

53. Lundberg U. Stress hormones in health and illness: the roles of work and gender. *Psychoneuroendocrinology* (2005) 30(10):1017–21. doi: 10.1016/j.psyneuen.2005.03.014

54. Kajantie E, Phillips DI. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology* (2006) 31(2):151–78. doi: 10.1016/j.psyneuen.2005.07.002

55. Chaplin TM, Hong K, Bergquist K, Sinha R. Gender differences in response to emotional stress: an assessment across subjective, behavioral, and physiological domains and relations to alcohol craving. *Alcoholism Clin Exp Res* (2008) 32 (7):1242–50. doi: 10.1111/j.1530-0277.2008.00679.x

56. Liu MY, Li N, Li WA, Khan H. Association between psychosocial stress and hypertension: a systematic review and meta-analysis. *Neurological Res* (2017) 39 (6):573–80. doi: 10.1080/01616412.2017.1317904

57. Fauvel JP, M'Pio I, Quelin P, Rigaud JP, Laville M, Ducher M. Neither perceived job stress nor individual cardiovascular reactivity predict high blood pressure. *Hypertens (Dallas Tex 1979)* (2003) 42(6):1112–6. doi: 10.1161/01.Hyp.0000102862.93418.Ee

58. Agyei B, Nicolaou M, Boateng L, Dijkshoorn H, van den Born BJ, Agyemang C. Relationship between psychosocial stress and hypertension among Ghanaians in Amsterdam, the Netherlands–the GHAIA study. *BMC Public Health* (2014) 14:692. doi: 10.1186/1471-2458-14-692

59. Roos E, Lallukka T, Rahkonen O, Lahelma E, Laaksonen M. Working conditions and major weight gain-a prospective cohort study. *Arch Environ Occup Health* (2013) 68(3):166–72. doi: 10.1080/19338244.2012.686931

60. Kroenke CH, Spiegelman D, Manson J, Schernhammer ES, Colditz GA, Kawachi I. Work characteristics and incidence of type 2 diabetes in women. *Am J Epidemiol* (2007) 165(2):175–83. doi: 10.1093/aje/kwj355

61. Heraclides AM, Chandola T, Witte DR, Brunner EJ. Work stress, obesity and the risk of type 2 diabetes: gender-specific bidirectional effect in the Whitehall II study. *Obes (Silver Spring Md.)* (2012) 20(2):428–33. doi: 10.1038/oby.2011.95

62. Brunner EJ, Chandola T, Marmot MG. Prospective effect of job strain on general and central obesity in the Whitehall II Study. *Am J Epidemiol* (2007) 165(7):828–37. doi: 10.1093/aje/kwk058

63. Richardson AS, Arsenault JE, Cates SC, Muth MK. Perceived stress, unhealthy eating behaviors, and severe obesity in low-income women. *Nutr J* (2015) 14:122. doi: 10.1186/s12937-015-0110-4

64. Barrington WE, Ceballos RM, Bishop SK, McGregor BA, Beresford SA. Perceived stress, behavior, and body mass index among adults participating in a worksite obesity prevention program, Seattle, 2005-2007. *Preventing chronic Dis* (2012) 9:E152. doi: 10.5888/pcd9.120001

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 19 December 2023
DOI 10.3389/fendo.2023.1232266



OPEN ACCESS

EDITED BY
Xiao Wang,
Lund University, Sweden

REVIEWED BY
Abdurezak Ahmed Abdela,
Addis Ababa University, Ethiopia
Aydin Ece,
Dicle University, Türkiye

*CORRESPONDENCE
Alexander Teumer
ateumer@uni-greifswald.de

RECEIVED 31 May 2023
ACCEPTED 24 November 2023
PUBLISHED 19 December 2023

CITATION
Nasr MK, Schurmann C, Böttinger EP and Teumer A (2023) Mendelian randomization indicates causal effects of estradiol levels on kidney function in males.
*Front. Endocrinol.* 14:1232266.
doi: 10.3389/fendo.2023.1232266



# Mendelian randomization indicates causal effects of estradiol levels on kidney function in males

M. Kamal Nasr[1,2,3,4], Claudia Schurmann[2], Erwin P. Böttinger[2,5] and Alexander Teumer[1,3,4]*

[1]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany, [2]Digital Health Center, Hasso Plattner Institute, University of Potsdam, Potsdam, Germany, [3]DZHK (German Centre for Cardiovascular Research), Partner Site Greifswald, Greifswald, Germany, [4]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany, [5]Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, NY, United States

**Context:** Chronic kidney disease (CKD) is a public health burden worldwide. Epidemiological studies observed an association between sex hormones, including estradiol, and kidney function.

**Objective:** We conducted a Mendelian randomization (MR) study to assess a possible causal effect of estradiol levels on kidney function in males and females.

**Design:** We performed a bidirectional two-sample MR using published genetic associations of serum levels of estradiol in men (*n* = 206,927) and women (*n* = 229,966), and of kidney traits represented by estimated glomerular filtration rate (eGFR, *n* = 567,460), urine albumin-to-creatinine ratio (UACR, *n* = 547,361), and CKD (*n* = 41,395 cases and *n* = 439,303 controls) using data obtained from the CKDGen Consortium. Additionally, we conducted a genome-wide association study using UK Biobank cohort study data (*n* = 11,798 men and n = 6,835 women) to identify novel genetic associations with levels of estradiol, and then used these variants as instruments in a one-sample MR.

**Results:** The two-sample MR indicated that genetically predicted estradiol levels are significantly associated with eGFR in men (beta = 0.077; *p* = 5.2E-05). We identified a single locus at chromosome 14 associated with estradiol levels in men being significant in the one-sample MR on eGFR (beta = 0.199; *p* = 0.017). We revealed significant results with eGFR in postmenopausal women and with UACR in premenopausal women, which did not reach statistical significance in the sensitivity MR analyses. No causal effect of eGFR or UACR on estradiol levels was found.

**Conclusions:** We conclude that serum estradiol levels may have a causal effect on kidney function. Our MR results provide starting points for studies to develop therapeutic strategies to reduce kidney disease.

KEYWORDS

**glomerular filtration rate, steroids, albuminuria, genome-wide association study, causality**

# Introduction

Chronic kidney disease (CKD) due to impaired kidney function is a major contributor to death and suffering in the 21st century (1), affecting an estimated 843 million individuals worldwide in 2017. Between 1990 and 2017, the global all-age mortality rate attributed to CKD increased by 41.5%. Studies and research continue to be conducted to identify and evaluate the risk factors associated with the development of CKD. These risk factors include high blood pressure and diabetes mellitus type 1 and 2 (1).

Furthermore, sex-associated differences in the epidemiology of kidney disorders have been observed (2, 3). Studies and trials have shown that most people who reach end-stage kidney disease (ESKD) are men, and with a faster disease progression than women (4, 5). However, randomized controlled trials assessing a causal effect of estradiol levels on kidney disease are lacking.

Several theories exist to explain the sex-associated differences in exposure and prognosis of kidney disease. These include unhealthy lifestyle and habits, which are found to be more prevalent in men than in women (2, 6). However, one important physiological difference is the steroidal sex hormones, including testosterone and estradiol, which play an essential role in the development of sexual characteristics (7).

Similar to the testosterone levels in men, estradiol levels in women can vary depending on age and menstrual status. In premenopausal women, estradiol levels vary throughout the menstrual cycle starting from 20 pg/mL to 80 pg/mL during the early phase of menstruation (8), followed by a gradual increase until the level reaches its maximum at the middle of the cycle, before decreasing again at the end of it. The estradiol levels could reach 300 pg/mL by the end of the second week of the menstruation cycle (8, 9), with an upper limit even reaching higher than 600 pg/mL. Estradiol levels are significantly lower in postmenopausal women. Some studies report average levels between 50 pg/mL and 120 pg/mL in older women who are no longer menstruating (8, 9).

Sex hormones have secondary functions such as organ development and prevention of disorders, such as osteoporosis (10). However, their impact on kidneys is not fully understood. Studies have shown an association between lower testosterone levels in men and increased all-cause mortality risk at advanced stages of CKD (11). A significant association between dialysis and decreased estradiol levels was also found in women, resulting in the lack of ovulation and abnormal menstruation cycles (12). The complexity of the regulation of the estradiol hormone in women has made it difficult to study its role and association with kidney functions. However, studies using animal models, which were designed to uncover the reasons underlying this association, suggested that estradiol and other estrogens could have a nephroprotective effect by antagonizing apoptosis of the podocytes, especially in females (13, 14). Estradiol has also shown protective effects on other kidney-damaging pathways like nitrogen oxide production and collagen synthesis. On the other hand, testosterone has shown destructive effects on kidney function, by either inducing apoptosis of podocytes or other mechanisms such as fibrosis of kidney cells (2, 15). Most of these studies were conducted in animal models, thus creating a need to investigate the effect of sex hormones on kidney functions in humans (16).

Randomized controlled trials are a well-established method to assess causality. However, the high cost of conducting these studies and their challenging feasibility are pertinent drawbacks of this approach (17). Mendelian randomization (MR) is an alternative method using genetic associations as instrumental variables to overcome possible bias due to confounding when drawing causal inference from observational studies (18). MR methods have been utilized to investigate causal effects of several phenotypes, including kidney function (19–22).

Previous studies conducting MR analyses of sex hormones on kidney function focused on testosterone and sex hormone-binding globulin (SHBG) by using data from the UK Biobank cohort study (21, 22). These data revealed that genetically predicted SHBG levels are associated with a protective effect on kidney function and a reduced risk of CKD in the male population (21). In addition, genetically predicted testosterone levels increased the risk of CKD in men (22).

However, there is a lack of studies investigating the causal effect of estradiol hormone levels on kidney function in both male and female populations (23). Estradiol levels are subject to wide intra-individual variation in the premenopausal female population, whereas they are generally lower in postmenopausal women, and often below the detection limit in men. These variations complicate the identification of genetic variants that are significantly associated with the hormone levels (23, 24).

Here, we conducted a bidirectional MR to assess causality between the levels of the estradiol sex hormone and kidney traits in both males and females, using known and novel estradiol-associated genetic variants as instruments. The significant findings of the two-sample MR were aimed for validation by additional pleiotropy-robust MR methods and a one-sample MR using the UK Biobank cohort study data.

# Materials and methods

## Study design

We applied MR to assess causal associations of the estradiol hormone on the urine albumin-to-creatinine ratio (UACR), the estimated glomerular filtration rate (eGFR) based on serum creatinine, and CKD using two-sample MR analyses. We included the single-nucleotide polymorphism (SNP) summary statistics for males ($n$ = 206,729) and females ($n$ = 229,966), from genome-wide association studies (GWAS) of two different publications conducted in the UK Biobank, for instrument selection of estradiol in the two-sample MR. We used the summary statistics on kidney-related traits obtained from the CKDGen Consortium (25, 26). The datasets were limited to individuals of European ancestry aligning them with the estradiol sample population. The GWAS included 480,698 individuals for CKD (41,395 cases), 567,460 individuals for eGFR, and 547,361 individuals for UACR. In these studies, both eGFR and UACR were log-transformed. Additionally, the UACR was inverse-normal transformed before conducting the GWAS.

To validate and test the robustness of the significant two-sample MR findings, we conducted a GWAS on the continuous estradiol levels in the UK Biobank dataset as a means to discover instruments for a subsequent one-sample MR. Finally, we tested for a potential causal effect of the kidney traits on estradiol levels. An overview of the analyses performed, including its main results, is provided in Figure 1.

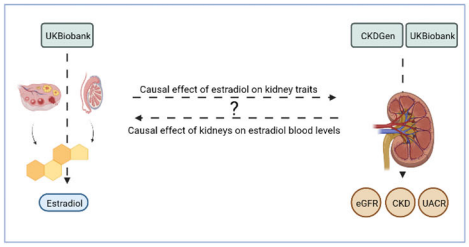
To ensure that the instruments were independent from each other, we used LocusZoom (https://my.locuszoom.org) with an $R^2 <$ 0.01 cut-off value to select the variant with the smallest $p$-value per locus (27).

## Dataset selection for the two-sample Mendelian randomization

We applied a two-sample MR, which can use SNP–outcome and SNP–exposure associations obtained from the GWAS datasets, to assess causality. Three core assumptions on genetic variants have to be fulfilled to act as suitable instruments in a MR analyses (1): association with the exposure (2); independence of the outcome given the exposure and all the confounders of the exposure–outcome association; and (3) independence of the factors that confound the exposure–outcome relationship (28). We used the datasets and applied the methods described in detail below to ensure the validity of the instruments as far as possible. For the exposure data, we used two different published GWASs that have investigated genetic associations with estradiol. We chose instruments with genome-wide significant associations ($p < 5*10^{-8}$) with the exposure, and removed variants with pleiotropic effects on the potential confounders, as described below. Finally, we applied Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) analysis (29) to identify outliers among the instruments, which were then removed prior to the subsequent MR analyses.

The results from Haas et al. (30) included women of European ancestry stratified by their menstruation status. The second source for the genetic instruments was obtained from the study by Ruth et al. (23), in which a GWAS in men with European ancestry was conducted. Both studies used data from the UK Biobank cohort study and identified estradiol as a dichotomous variable being above the detection limit (23, 30). The selection of potential instrumental variables was performed by following the guidelines for MR analyses (31).

To assess the causal effects of estradiol levels on kidney function, we used genetic predictors of log-transformed UACR and eGFR




| Population Group | | Exposure Group GWAS | MR Performed | MR Results |
|---|---|---|---|---|
| Women | Overall | Haas *et al.* summary statistics (n = 229,966)<br>7 SNPs used as instruments<br>associated with estradiol above/below the limit<br>(dichotomous variable) | Two-Sample MR,<br>CKDGen summary<br>statistics as<br>outcome | No Significant association |
| | Pre-menopausal | Haas *et al.* summary statistics (n = 51,081)<br>3 SNPs used as instruments<br>associated with estradiol above/below the limit<br>(dichotomous variable) | Two-Sample MR,<br>CKDGen summary<br>statistics as<br>outcome | Significant association with UACR<br>(beta = -0.045, p-value = 2.1E-11) |
| | Post-menopausal | Haas *et al.* summary statistics (n = 84,194)<br>4 SNPs used as instruments<br>associated with estradiol above/below the limit<br>(dichotomous variable) | Two-Sample MR,<br>CKDGen summary<br>statistics as<br>outcome | Significant association with eGFR<br>(beta = 0.010, p-value = 3.7E-4) |
| Men | | Ruth *et al.* summary statistics (n = 206,729)<br>8 SNPs were used as instruments associated<br>with estradiol above/below the limit<br>(dichotomous variable) | Two-Sample MR,<br>CKDGen summary<br>statistics as<br>outcome | Significant association with eGFR<br>(beta = 0.077, p-value = 5.2E-05) |
| | | UK Biobank Individual Data (n = 11,798)<br>1 SNP used as an Instrument<br>associated with estradiol blood levels<br>(continous variable) | One-Sample MR<br>in the UK Biobank | Significant association with eGFR<br>(beta = 0.199, p-value = 0.017) |

FIGURE 1
Overview of the Mendelian randomization analyses and its results.The upper scheme illustrates the main goal of our Mendelian randomization (MR) study investigating a possible causal effect between estradiol levels and kidney function traits as represented by estimated glomerular filtration rate (eGFR), urinary albumin-to-creatinine ratio (UACR), and risk of chronic kidney disease (CKD). The lower table summarizes the strata, datasets, methods, and the results of the MR analyses. The green boxes represent validated significant results, where the results in the yellow boxes could not be confirmed by sensitivity analyses. The figure was created with BioRender.com.

from the CKDGen Consortium GWAS results as exposure (25, 26). As the log-transformed UACR was included as sex- and age-adjusted inverse-rank-normalized residuals in the GWAS, the unit change corresponds to one standard deviation (SD) change of the log-transformed UACR. We used the generated summary statistics of the GWAS in the UK Biobank for the genetic associations with the log-transformed estradiol levels as an outcome in males and females.

We looked for proxy SNPs with an $R^2$ value > 0.8 if the potential instrument was not available in the outcome GWAS results. However, no proxies could be found. All instruments were primarily associated with the exposure according to the Steiger test (32). The results were subsequently verified for no association with body mass index (BMI), body fat mass, and type 2 diabetes as potential confounders using the PhenoScanner webtool (33). Details on the instrument selection are provided below.

## Details on variant selection in the published datasets used as instruments for estradiol

The first data source was obtained from Haas et al. (30). Of the 229,966 women with European ancestry available in the UK Biobank dataset, 51,081 were premenopausal and 84,194 were identified as being postmenopausal. The GWAS was conducted on estradiol as a continuous phenotype (inverse-rank normalized), and as a dichotomized outcome using PLINK 2.0 (34). The GWAS on the continuous trait revealed only one significant association (rs727428), which also represented a known association with SHBG and testosterone levels (35), and was thus not treated as a valid instrument for estradiol. Therefore, we only used the results of the dichotomized trait for the subsequent MR analyses. We extracted all the independent SNPs with genome-wide significance ($p < 5*10–^{8}$) presented in Haas et al. The GWAS results for women overall provided 10 SNPs with a significant genetic association with estradiol, only seven of them were suitable candidates for the following two-sample MR. Two SNPs (rs774021038 and rs71181755) were excluded, as there were no available corresponding results in the outcome summary statistics, while one SNPs (rs34929649) was excluded due to its association with the fat mass of body parts. Of the selected seven SNPs, four SNPs were eligible as candidates in the postmenopausal women group and three in the premenopausal women.

The second source for the genetic instruments for the two-sample MR was obtained from the study by Ruth et al. (23), in which a GWAS in the UK Biobank cohort was conducted. In that GWAS, genetic variants associated with estradiol levels in men above vs. below the assay detection limit were analyzed using a linear model. The authors identified 22 variants with a statistically significant association with estradiol levels, of which 10 (rs188982745, rs570754094, rs781858752, rs34019140, rs201687269, rs5933688, rs12850857, rs776715248, 4:69958680_GA_G, and 5:35983283_CA_C) were not available in the kidney trait GWAS results, and one SNP (rs117826558) was available only in the summary statistics of the GWAS on CKD.

One SNP, rs1260326 on chromosome 2, was excluded as an instrument because it was significantly associated with eGFR in the CKDGen Consortium summary statistics ($p = 2*10^{-36}$) used as outcome, two SNPs were excluded from the analysis due to their association with possible confounders violating the MR assumptions (31): rs45446698 was associated with BMI and fat percentage, and rs727428 due to its association with levels of testosterone and SHBG. MR-PRESSO identified rs657152 as an outlier. This variant is located near *ABO* on chromosome 9, a gene that shows a considerable association with angiotensin-converting enzymes (i.e., ACE1/ACE2) (36) and is supposed to have a direct effect on kidney functionality (37).

All the instruments had a minor allele frequency (MAF) > 1% and a high imputation quality (info score ≥ 0.8) in both the exposure and outcome data. The final list of instruments for estradiol included in the two-sample MR analyses are provided in Supplementary Tables 1, 2.

## Details on variant selection for kidney trait instruments

Of the 256 genome-wide significant associations associated with serum creatinine-based eGFR in the European ancestry sample in the publication of Wuttke et al. (25), 122 variants that were marked with likely support for kidney function and replicated in the MVP study (if available) were included as potential instruments. In total, 14 SNPs were excluded due to their association with BMI or body fat mass (rs10430743, rs10774625, rs10838702, rs112545201, rs11564722, rs1268176, rs2411192, rs35004449, rs3134605, rs3905668, rs55759218, rs632887, rs9375694, and rs9828976), leaving 108 SNPs that passed the selection criteria for instruments (Supplementary Table 3).

For UACR, the 63 conditionally independent genome-wide associations of the European ancestry meta-analysis, conducted in the CKDGen Consortium (26), were selected as potential instruments. Five of the SNPs (rs17453832, rs557338857, rs141493439, rs45551835, and rs562661763) were not available in the GWAS of the outcome, leaving 58 instruments for the two-sample MR of UACR (Supplementary Table 4).

## Data selection for the GWAS and one-sample Mendelian randomization

The UK Biobank is a prospective cohort study with deep genetic and phenotypic data of more than five hundred thousand individuals recruited from England, Scotland, and Wales (38). Given the large sample size of the UK Biobank study with both estradiol levels and kidney function markers available, we conducted a one-sample MR in this dataset. This additional dataset provided us with the opportunity to assess a causal effect on kidney function, thus extending the published GWASs by using the continuous scale of estradiol levels and restricting the sex hormones to postmenopausal women, which in turn reduced the

heterogeneity in these measurements. The details of the sample selection are provided in Supplementary Figure 1.

The eGFR was calculated with the CKD-EPI study equation using the R (The R Foundation for Statistical Computing, Vienna, Austria) package "nephro" with serum creatinine levels, age, and sex as inputs, and it was log-transformed for subsequent analyses to match with the two-sample MR.

To identify the SNPs with a significant association as being candidates for instruments, we conducted a GWAS of log-transformed estradiol levels on the imputed genotypes using a linear mixed model implemented in BOLT-LMM (39). The GWAS was conducted for male, female, and sex-combined groups.

For the GWAS and the subsequent one-sample MR, we included only European ancestry individuals (identified by field ID: 22006) with available active consent, genotype data, blood estradiol levels, and creatinine levels measured in urine and blood. We excluded the individuals with a recorded estrogen-based treatment to avoid an exogenous confounding effect (24, 31). For the female population, we included only self-reported postmenopausal women to avoid confounding caused by uncontrolled changes in estradiol levels during the menstruation cycle (8, 9).

Out of 502,505 individuals available in the dataset, 92,810 were excluded because of their non-European ancestry. We excluded 407 individuals due to having estradiol-based treatment (field ID 20003) and 12 individuals due to consent withdrawal. Of the remaining 409,276 individuals, 84,567 premenopausal women and 303,087 individuals with missing estradiol or genotype data were excluded, which resulted in 21,622 individuals (6,835 women and 14,797 men). Of these, 14,797 were male and 11,798 individuals had kidney biomarkers available and were thus included in the subsequent one-sample MR analyses.

We used BMI and age, and also sex in the combined analysis as covariates. In each GWAS, SNPs were filtered using a minor allele frequency (MAF) > 0.001, a Hardy–Weinberg equilibrium *p-value* > $10^{-12}$, and an imputation info score ≥ 0.8. We used a *p-value* < $5^{*}10^{-8}$ as a threshold for genome-wide significance. As no instruments in women were found, the one-sample MR was restricted to men.

## Statistical analyses

In the two-sample MR analyses, we used the inverse variance-weighted method (IVW), with multiplicative random effects to assess the causal effect of the exposure on the respective outcome. To test the robustness of the significant MR result, we applied the pleiotropy-robust but less powerful weighted median (40) and MR Egger (41) methods. Cochran's Q was used to test for the heterogeneity of the causal effect of the individual instruments in the IVW MR. The MR Egger intercept was tested for directional pleiotropy. The analyses were conducted using the R package "TwoSampleMR".

For the one-sample MR, we applied a two-stage least squares regression implemented in the R packages "tsls" and "ivreg" for UACR and eGFR, and the control function estimator for CKD (28). The analyses were adjusted for age and BMI. Supplementary Figure 2 provides a schematic overview of the analytical steps performed in the one-sample MR. The power calculation was performed with the "Online sample size and power calculator for Mendelian randomization with a continuous outcome" (https://sb452.shinyapps.io/power/).

For the two-sample MR, a $p$-value < 0.05/4 = 0.0125 was considered statistically significant, correcting for the two different sex strata and the kidney traits included as outcomes, that is, eGFR and CKD for kidney function, and UACR as a marker for kidney damage. For the confirmatory one-sample MR, a $p$-value < 0.05 was considered as significant.

## Ethics statement

In this project only published GWAS summary statistics and the data obtained from the UK Biobank cohort study with ethics approval, as provided on the study website and in the corresponding publication (38), were used.

# Results

## Two-sample Mendelian randomization

The MR using selected known genetic variants as instruments that are associated with estradiol as a dichotomous (above vs. below the assay detection limit) variable in women using the GWAS results of Haas et al. (30) revealed a significant association of genetically predicted higher estradiol levels with a higher eGFR in postmenopausal women (beta = 0.010; $p = 3.7^{*}10^{-4}$; Table 1). This association was not significant in the premenopausal and overall women groups (Tables 2, 3). The significant association in the eGFR had the effects of similar size and with the same direction in the MR

TABLE 1 Associations of the inverse variance-weighted two-sample Mendelian randomization of urinary albumin-to-creatinine ratio (UACR), chronic kidney disease (CKD), and estimated glomerular filtration rate (eGFR) in postmenopausal women using the Haas et al. summary statistics for genetically predicted estradiol levels.

| Outcome | #SNPs | Estimate/[OR] | 95% CI | *p*-value | Q pval |
|---|---|---|---|---|---|
| UACR | 4 | −0.035 | −0.076 to 0.007 | 0.105 | 0.618 |
| CKD | 4 | [0.851] | 0.651 to 1.052 | 0.115 | 0.607 |
| eGFR | **4** | **0.010** | **0.004 to 0.015** | **3.7E-4** | 0.814 |

The Q pval represents the heterogeneity test result p-value. The OR represents the odds ratio of CKD. SNP, number of single nucleotide polymorphism; CI, confidence intervals. The association results in bold for highlighting statistical significance.

TABLE 2 Associations of the inverse variance-weighted two-sample Mendelian randomization of urinary albumin-to-creatinine ratio (UACR), chronic kidney disease (CKD), and estimated glomerular filtration rate (eGFR) in premenopausal women using the Haas et al. summary statistics for genetically predicted estradiol levels.

| Outcome | #SNPs | Estimate/[OR] | 95% CI | *p*-value | Q pval |
|---|---|---|---|---|---|
| UACR | **3** | **−0.045** | **−0.058 to −0.032** | **2.1E-11** | 0.942 |
| CKD | 3 | [0.855] | 0.637 to 1.074 | 0.160 | 0.471 |
| eGFR | 3 | 0.008 | 0.002 to 0.014 | 0.013 | 0.667 |

The association results in bold for highlighting statistical significance.

sensitivity analyses, but without reaching the significance level (Supplementary Table 5).

In the premenopausal women group, there were higher genetically predicted estradiol levels significantly associated with a lower UACR (beta = −0.045; $p = 2.1*10^{-11}$; Table 2). These effect sizes were similar in the sensitivity analyses, but were not statistically significant (Supplementary Table 6).

The MR using the instruments for a dichotomous estimation of the estradiol levels in the male population from Ruth et al. revealed that higher genetically predicted estradiol levels are associated with a higher eGFR (beta = 0.077; $p = 5.2*10–^{5}$; Table 4). Similar results were obtained using the weighted median MR, thus confirming the significant associations (Supplementary Table 7).

No indication of directional pleiotropy or heterogeneity was found for the results (Tables 1–4 and Supplementary Tables 5-7). The MR scatter plots of the significant associations are given in Supplementary Figure 3.

No significant MR results were found for CKD as the outcome (Tables 1–4).

## Genome-wide association study and one-sample Mendelian randomization

The cohort characteristics of the UK Biobank participants included in this analysis are provided in Supplementary Table 8. Our three GWASs on continuous estradiol levels in the male ($n$ = 14,797), female ($n$ = 6,835), and sex-combined ($n$ = 21,632) datasets revealed only one genome-wide significant locus ($p < 5*10–^{8}$) at chromosome 14 in males (Figure 2). The SNP rs7151019 [T/G, MAF = 0.42, beta(T) = -0.026, imputation info = 0.90] represents the variant with the lowest *p*-value at this locus ($p = 6*10^{-22}$) explaining 0.73% of the variation of the log-transformed estradiol levels (SD = 0.15). The variant is located close to the immunoglobulin heavy locus (IGH), a protein-coding gene with no known direct link to estradiol metabolism. This locus did not reach statistical significance in the female or in the sex-combined GWAS (Supplementary Figures 4 and 5). Of note, this locus was not included in the two-sample MR analyses. The quantile–quantile plots of the GWAS results do not indicate inflation of the *p*-values (Supplementary Figure 6). The PhenoScanner (33) did not show an association with BMI, body fat mass, type 2 diabetes, or eGFR. No association in the female or sex-combined samples passed the level of genome-wide significance, thus no MR analyses could be performed in these datasets.

The one-sample MR using the 11,798 males with available kidney biomarkers and the SNP rs7151019 as instrument allowed a detection of at least a 0.3 SD unit change in the kidney trait per SD change in log-estradiol levels at a 80% power. We identified a significant association between estradiol levels and eGFR (beta = 0.199; $p = 0.017$) confirming the two-sample MR results. In concordance with the two-sample MR in males, no significant results were found for CKD and UACR (Table 5).

## Effects of kidney function on estradiol blood levels

The MR for testing causal effects of kidney function traits on estradiol levels revealed no significant association of genetically predicted eGFR or UACR with estradiol levels in males or females (Supplementary Table 9).

# Discussion

Our two-sample MR analysis based on the instruments assessing estradiol levels below vs. above the detection limit revealed a significant causal effect in males with a positive effect direction, implying that higher levels of estradiol could lead to

TABLE 3 Associations of the inverse variance-weighted two-sample Mendelian randomization of urinary albumin-to-creatinine ratio (UACR), chronic kidney disease (CKD), and estimated glomerular filtration rate (eGFR) in the overall women population using the Haas et al. summary statistics for genetically predicted estradiol levels.

| Outcome | #SNPs | Estimate/[OR] | 95% CI | *p*-value | Q pval |
|---|---|---|---|---|---|
| UACR | 7 | −0.032 | −0.088 to 0.024 | 0.262 | 0.698 |
| CKD | 7 | [0.915] | 0.586 to 1.243 | 0.595 | 0.427 |
| eGFR | 7 | 0.007 | −0.005 to 0.018 | 0.270 | 0.500 |

The Q pval represents the heterogeneity test result p-value. The OR represents the odds ratio of CKD.

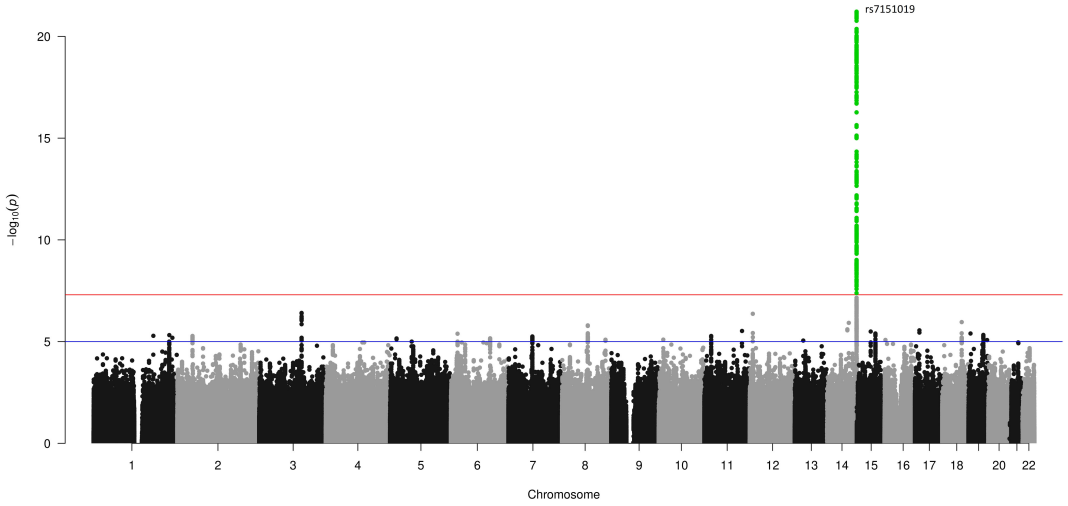



FIGURE 2
GWAS results for estradiol levels in men in the UK Biobank. A Manhattan plot showing the SNP positions on the *x*-axis, and their association $-\log_{10}(p)$ on the *y*-axis. The red line represents the threshold for genome-wide significance of $5*10^{-8}$, the green-colored dots represent variants with a *p*-value equal or smaller than the suggestive threshold of $10^{-6}$.

higher eGFRs. This association was validated in the one-sample MR using continuous levels of estradiol above the detection limit. In addition, such a causal effect on eGFR was suggested by the two-sample MR in postmenopausal women, and an inverse effect in premenopausal women on UACR. However, these results did not reach statistical significance in the sensitivity analyses, and could not be validated using a one-sample MR on continuous outcomes. Overall, the MR analyses suggest a causal effect between higher estradiol levels and better kidney function traits.

To our knowledge, this study is the first MR analysis to investigate the possible causal effect of estradiol levels on kidney function traits. Previous research studies investigated the relationship between the decline in kidney performance in men and women and the change in sex hormone levels in general (13, 42). Based on the results of these studies, researchers sought to identify possible mechanisms of the influence of these hormones on kidney function. For estradiol, most of the studies showed a possible protective effect, either by inhibiting the pathological processes of increasing oxidative stress in the diseased kidney (42), or by inhibiting renal fibrosis aggravation and glomerular sclerosis (13). Animal studies have shown that estradiol plays a protective role by reducing albuminuria and enhancing creatinine clearance (43), which is in line with the effect directions of our MR results. However, these findings were yet not confirmed by sex-stratified analyses. Other studies showed contradicting results. In a nationally representative sample of a United States adult male population, increased levels of estradiol were associated with a decrease in eGFR (44); however, other studies failed to identify an association between endogenous estradiol levels and changes in eGFR or albuminuria (45). Even though most of these studies attempted to find a correlation between sex and the risk of kidney disease, the results of these studies failed to establish a hypothesis for this association.

The aim of our study was to assess a possible causal effect between estradiol levels and kidney function using the MR framework by including summary statistics from published GWASs and data of the large UK Biobank cohort study. To reduce sex-specific heterogeneity in estradiol measurements, we focused on sex-stratified analyses. Our results showed that genetically predicted estradiol levels were significantly associated with an increased eGFR in men. Although we reduced heterogeneity of the estradiol measurements for our two-sample MR analyses in women by using GWAS results that were stratified by pre- and postmenopausal status, our findings were indicative in this sample given the consistent effect direction but not robustly significant. A reason could be the reduced power in the postmenopausal women dataset given the small sample size of individuals above the estradiol detection limit, and the trait variation due to the menstrual cycle in the premenopausal women obtained from the GWAS of Haas et al. (30), where they relied on self-reported menopause and age below 60 years at sampling time. Thereby, the women-combined dataset induces large variation of estradiol levels, thus reducing the statistical power in both the GWAS and MR analyses.

The GWAS results of the log-transformed estradiol levels, which we conducted in the UK Biobank cohort study, differed from the analysis of former studies, and by this also the inclusion of genetic instruments in the one-sample MR analyses. The possible reasons for this difference are due to several aspects. Due to the limited information of the menstruation cycle at time of estradiol measurement of the female study participants, we limited the corresponding GWAS to 6,835 self-reported postmenopausal women. The minimum detectable level in the UK Biobank cohort study was 175 pmol/L, thus the detection of estradiol in postmenopausal women was less sensitive compared with other studies like Pott et al. (24). This detection limit affects to a lesser extent the analyses in men, who, on average, have higher estradiol levels than postmenopausal women. In the study of Pott et al., only one locus harboring the signal peptide peptidase-like 2A gene (rs12913657 on chromosome 15) reached genome-wide

TABLE 4 Associations of the inverse variance-weighted two-sample Mendelian randomization of urinary albumin-to-creatinine ratio (UACR), chronic kidney disease (CKD), and estimated glomerular filtration rate (eGFR) in males after using the Ruth et al. summary statistics for genetically predicted estradiol levels.

| Outcome | #SNPs | Estimate/[OR] | 95% CI | *p*-value | Q pval |
|---|---|---|---|---|---|
| UACR | 7 | −0.024 | −0.226 to 0.178 | 0.819 | 0.243 |
| CKD | 8 | [0.522] | 0.308 to 1.36 | 0.512 | 0.492 |
| eGFR | 7 | **0.077** | **0.040 to 0.114** | **5.2E-05** | 0.216 |

The Q pval represents the heterogeneity test result *p*-value.
The OR represents the odds ratio of CKD. The significant results are marked in bold.

TABLE 5 Associations of the one-sample Mendelian randomization of urinary albumin-to-creatinine ratio (UACR), chronic kidney disease (CKD), and estimated glomerular filtration rate (eGFR) in males of the UK Biobank cohort study.

| Outcome | #SNPs | Estimate/[OR] | 95% CI | *p*-value |
|---|---|---|---|---|
| UACR | 1 | −0.33 | −1.170 to 0.389 | 0.397 |
| CKD | 1 | [0.758] | 0.508 to 1.18 | 0.206 |
| eGFR | 1 | **0.199** | **0.036 to 0.362** | **0.017** |

The OR represents the odds ratio of CKD. The significant results are marked in bold.

significance in 4,191 men, but without a replication sample included (24). This locus was not associated with estradiol level in the larger GWAS of the 14,797 men from the UK Biobank dataset. However, our GWAS revealed a highly significant association on chromosome 14 (rs7151019), which was used as instrument in the one-sample MR.

Ruth et al. also conducted a GWAS in the UK Biobank males, but they dichotomized the estradiol level at the detection limit (23). Their GWAS revealed more loci associated with estradiol levels by including a larger sample size. However, the two-sample MR result using this dataset confirmed our one-sample MR. In addition, these eGFR MR results were also directionally consistent with the CKD MR, although not reaching significance after multiple testing. Of note, our significant GWAS locus at chromosome 14 was also revealed by Ruth et al., where one of their top SNPs (rs34019140) was in moderate linkage disequilibrium with our SNP rs7151019 ($R^2$ = 0.45). However, rs34019140 (or a suitable proxy) was not available in the summary statistics of the kidney traits.

The main strength of our project is the different analyses performed using multiple sources of genetic association data, with two different exposure populations for the use in the large two-sample MR, as well as individual-level data from the UK Biobank cohort study for use in the one-sample analysis. One challenge of the MR approach is avoiding weak instrument bias (31). Thus, we selected only independent variants in the one-sample MR having a strong statistical association with estradiol levels ($p < 5^{*}10^{-8}$). Furthermore, it is important that the instruments in a MR are not associated with (unadjusted) confounders of the estradiol–kidney function relationship. Therefore, we conducted the one-sample MR by adjusting for possible confounders, ensuring the robustness of the association results (46, 47).

On the other hand, our study had several limitations. The first is the lack of available estradiol instruments from the meta-analysis results. This shortage is mainly due to the changes in estradiol levels in premenopausal women during the menstruation cycle. Such daily changes make it hard to develop a uniform study design for measuring estradiol levels. An alternative approach is to include exclusively postmenopausal women, which results in a small sample size and is where measurement of estradiol levels is technically difficult. This heterogeneity of estradiol levels in women is the second limitation for our study. The available meta-analysis results for continuous estradiol include populations with relatively small sample sizes, which results in low power for identifying genetic instruments. There was no sample overlap between the estradiol level assessed in the UK Biobank cohort study and eGFR, but the UACR GWAS included the UK Biobank cohort study dataset, which could in turn bias the effect estimates of the respective two-sample MR (48). Finally, no sex-stratified kidney trait GWASs were available, which could be a reason for the non-significant two-sample MR results. Although the MR analyses identified statistically significant associations, the reported effect sizes are small and hard to interpret. However, the effect sizes itself obtained from an MR are generally less informative (31).

Although we found a robust and significant MR result in men, we cannot exclude that there is also a causal effect of estradiol level on kidney function in women. Taking this into account, our results do not allow a conclusion on whether or not the observed differences in kidney disease prevalence between men and women can be attributed to sex-specific differences in estradiol levels.

Our results also highlight the need to identify additional genetic variants associated with estradiol levels in men and women providing instruments for MR analyses, and also in non-European populations. Finding such associations could be challenging, especially in the female population. Furthermore, studies like randomized controlled trials are required to estimate the magnitude of this potential causal relation. Nevertheless, our MR results provide starting points for subsequent studies focusing on the effects of estradiol levels on kidney function which may finally lead to therapeutic strategies as part of preventing kidney diseases.

## Data availability statement

Publicly available datasets were analyzed in this study. The data underlying this article are available from the UK Biobank (GWAS on estradiol), at https://www.ukbiobank.ac.uk/. The remaining datasets were derived from publications as referenced in the manuscript.

## Ethics statement

Ethical approval was not required for the studies involving humans because only published GWAS summary statistics and data obtained from UK Biobank with ethical approval as provided on the study website and in the corresponding publication was used. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from a previous study for which ethical approval was obtained. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

## Author contributions

MKN: writing and analysis; AT: supervising, reviewing, writing, and editing; CS: supervising, reviewing, and editing; EB: supervising. All authors contributed to the article and approved the submitted version.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. MKN was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—451892213; 455978266. We acknowledge support for the Article Processing Charge by the German Research Foundation and the Open Access Publication Fund of the University of Greifswald.

## Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number 40502. We thank Christoph Lippert for his support with the UK Biobank data, and Matthias Kirchler for assistance with the GWAS analyses. We are grateful to the authors of the publications of Ruth et al. and Haas et al., as well as to the CKDGen Consortium for providing the GWAS summary statistics.

## Conflict of interest

CS is currently an employee of Bayer AG, Pharmaceuticals.

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



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1232266/full#supplementary-material

## References

1. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. *Kidney Int Suppl* (2022) 12:7–11. doi: 10.1016/j.kisu.2021.11.003

2. Carrero JJ, Hecking M, Chesnaye NC, Jager KJ. Sex and gender disparities in the epidemiology and outcomes of chronic kidney disease. *Nat Rev Nephrol* (2018) 14:151–64. doi: 10.1038/nrneph.2017.181

3. Antlanger M, Noordzij M, van de Luijtgaarden M, Carrero JJ, Palsson R, Finne P, et al. Sex differences in kidney replacement therapy initiation and maintenance. *Clin J Am Soc Nephrol* (2019) 14(11):1616–25. doi: 10.2215/CJN.04400419

4. Coggins C. Differences between women and men with chronic renal disease. *Nephrol Dial Transplant* (1998) 13:1430–7. doi: 10.1093/ndt/13.6.1430

5. Farahmand M, Ramezani Tehrani F, Khalili D, Cheraghi L, Azizi F. Endogenous estrogen exposure and chronic kidney disease; a 15-year prospective cohort study. *BMC Endocr Disord* (2021) 21:155. doi: 10.1186/s12902-021-00817-3

6. Michishita R, Matsuda T, Kawakami S, Kiyonaga A, Tanaka H, Morito N, et al. The association between unhealthy lifestyle behaviors and the prevalence of chronic kidney disease (CKD) in middle-aged and older men. *J Epidemiol* (2016) 26:378–85. doi: 10.2188/jea.JE20150202

7. Sanderson JT. The steroid hormone biosynthesis pathway as a target for endocrine-disrupting chemicals. *Toxicol Sci* (2006) 94(1):3–21. doi: 10.1093/toxsci/kfl051

8. Becker JB, Berkley KJ, Geary N, Hampson E, Herman JP, Young E (eds.). *Sex Differences in the Brain:From Genes to Behavior: From Genes to Behavior*. Oxford University Press USA. (2007).

9. Strauss JF, Barbieri RL, Yen SSC. eds. *Physiology, pathophysiology, and clinical management. 6th ed*. Elsevier, Philadelphia PA. (2019). Available at: https://www.sciencedirect.com/science/book/9780323479127. (Retrieved December 6 2023).

10. Tuck SP, Francis RM. Testosterone, bone and osteoporosis. In: *Advances in the management of testosterone deficiency*. Basel: KARGER (2008). p. 123–32. doi: 10.1159/000176049

11. Khurana KK, Navaneethan SD, Arrigain S, Schold JD, Nally JV, Shoskes DA. Serum testosterone levels and mortality in men with CKD stages 3-4. *Am J Kidney Dis* (2014) 64(3):367–74. doi: 10.1053/j.ajkd.2014.03.010

12. Ahmed SB, Ramesh S. Sex hormones in women with kidney disease. *Nephrol Dial Transplant* (2016) 31(11):1787–95. doi: 10.1093/ndt/gfw084

13. Valdivielso JM, Jacobs-Cachá C, Soler MJ. Sex hormones and their influence on chronic kidney disease. *Curr Opin Nephrol Hypertens* (2019) 28:1–9. doi: 10.1097/MNH.0000000000000463

14. Dhindsa S, Reddy A, Karam JS, Bilkis S, Chaurasia A, Mehta A, et al. Prevalence of subnormal testosterone concentrations in men with type 2 diabetes and chronic kidney disease. *Eur J Endocrinol* (2015) 173(3):359–66. doi: 10.1530/EJE-15-0359

15. Duckles SP, Miller VM. Hormonal modulation of endothelial NO production. *Pflugers Arch Eur J Physiol* (2010) 459(6):841–51. doi: 10.1007/s00424-010-0797-1

16. Lapi F, Azoulay L, Niazi MT, Yin H, Benayoun S, Suissa S. Androgen deprivation therapy and risk of acute kidney injury in patients with prostate cancer. *JAMA - J Am Med Assoc* (2013) 310(3):289–96. doi: 10.1001/jama.2013.8638

17. Admon AJ, Donnelly JP, Casey JD, Janz DR, Russell DW, Joffe AM, et al. Emulating a novel clinical trial using existing observational data predicting results of the PREVENT study. *Ann Am Thorac Soc* (2019) 16(8):998–1007. doi: 10.1513/AnnalsATS.201903-241OC

18. Ebrahim S, Davey Smith G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum Genet* (2008) 123:15–33. doi: 10.1007/s00439-007-0448-6

19. Ellervik C, Mora S, Ridker PM, Chasman DI. Hypothyroidism and kidney function: A mendelian randomization study. *Thyroid* (2020) 30:365–79. doi: 10.1089/thy.2019.0167

20. Teumer A, Gambaro G, Corre T, Bochud M, Vollenweider P, Guessous I, et al. Negative effect of vitamin D on kidney function: a Mendelian randomization study. *Nephrol Dial Transplant* (2018) 33(12):2139–45. doi: 10.1093/ndt/gfy074

21. Zhao JV, Schooling CM. Sex-specific associations of sex hormone binding globulin with CKD and kidney function: A univariable and multivariable mendelian randomization study in the UK biobank. *J Am Soc Nephrol* (2021) 32:686–94. doi: 10.1681/ASN.2020050659

22. Zhao JV, Schooling CM. The role of testosterone in chronic kidney disease and kidney function in men and women: a bi-directional Mendelian randomization study in the UK Biobank. *BMC Med* (2020) 18:122. doi: 10.1186/s12916-020-01594-x

23. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med* (2020) 26:252–8. doi: 10.1038/s41591-020-0751-5

24. Pott J, Bae YJ, Horn K, Teren A, Kühnapfel A, Kirsten H, et al. Genetic association study of eight steroid hormones and implications for sexual dimorphism of coronary artery disease. *J Clin Endocrinol Metab* (2019) 104:5008–23. doi: 10.1210/jc.2019-00757

25. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet* (2019) 51:957–72. doi: 10.1038/s41588-019-0407-x

26. Teumer A, Li Y, Ghasemi S, Prins BP, Wuttke M, Hermle T, et al. Genome-wide association meta-analyses and fine-mapping elucidate pathways influencing albuminuria. *Nat Commun* (2019) 10:4130. doi: 10.1038/s41467-019-11576-0

27. Boughton AP, Welch RP, Flickinger M, VandeHaar P, Taliun D, Abecasis GR, et al. LocusZoom.js: Interactive and embeddable visualization of genetic association study results. *Bioinformatics* (2021) 37(18):3017–8. doi: 10.1093/bioinformatics/btab186

28. Teumer A. Common methods for performing mendelian randomization. *Front Cardiovasc Med* (2018) 5:51. doi: 10.3389/fcvm.2018.00051

29. Verbanck M, Chen C-Y, 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

30. Haas CB, Hsu L, Lampe JW, Wernli KJ, Lindström S. Cross-ancestry genome-wide association studies of sex hormone concentrations in pre- and postmenopausal women. *Endocrinol (United States)* (2022) 163:1–12. doi: 10.1210/endocr/bqac020

31. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* (2020) 4:186. doi: 10.12688/wellcomeopenres.15555.2

32. 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

33. 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) 35:4851–3. doi: 10.1093/bioinformatics/btz469

34. 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

35. Harrison S, Davies NM, Howe LD, Hughes A. Testosterone and socioeconomic position: Mendelian randomization in 306,248 men and women in UK Biobank. *Sci Adv* (2021) 7(31):eabf8257. doi: 10.1126/sciadv.abf8257

36. Gemmati D, Tisato V. Genetic hypothesis and pharmacogenetics side of renin-angiotensin-system in COVID-19. *Genes (Basel)* (2020) 11:1–17. doi: 10.3390/genes11091044

37. Mizuiri S. ACE and ACE2 in kidney disease. *World J Nephrol* (2015) 4:74. doi: 10.5527/wjn.v4.i1.74

38. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nat 2018 5627726* (2018) 562:203–9. doi: 10.1038/s41586-018-0579-z

39. Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* (2015) 47:284–90. doi: 10.1038/ng.3190

40. 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

41. Bowden J, Smith GD, 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

42. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, Brinton RD, Carrero J-J, DeMeo DL, et al. Sex and gender: modifiers of health, disease, and medicine. *Lancet* (2020) 396:565–82. doi: 10.1016/S0140-6736(20)31561-0

43. Mankhey RW, Bhatti F, Maric C. 17β-Estradiol replacement improves renal function and pathology associated with diabetic nephropathy. *Am J Physiol Physiol* (2005) 288:F399–405. doi: 10.1152/ajprenal.00195.2004

44. Yi S, Selvin E, Rohrmann S, Basaria S, Menke A, Rifai N, et al. Endogenous sex steroid hormones and measures of chronic kidney disease (CKD) in a nationally representative sample of men. *Clin Endocrinol (Oxf)* (2009) 71:246–52. doi: 10.1111/j.1365-2265.2008.03455.x

45. Kim C, Ricardo AC, Boyko EJ, Christophi CA, Temprosa M, Watson KE, et al. Sex hormones and measures of kidney function in the diabetes prevention program outcomes study. *J Clin Endocrinol Metab* (2019) 104:1171–80. doi: 10.1210/jc.2018-01495

46. Damtie S, Biadgo B, Baynes HW, Ambachew S, Melak T, Asmelash D, et al. Chronic Kidney Disease and Associated Risk Factors Assessment among Diabetes Mellitus Patients at A Tertiary Hospital, Northwest Ethiopia. *Ethiop J Health Sci* (2018) 28(6):691–700. doi: 10.4314/ejhs.v28i6.3

47. Shi H, Clegg DJ. Sex differences in the regulation of body weight. *Physiol Behav* (2009) 97:199–204. doi: 10.1016/j.physbeh.2009.02.017

48. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* (2016) 40:597–608. doi: 10.1002/gepi.21998

# Frontiers in Endocrinology

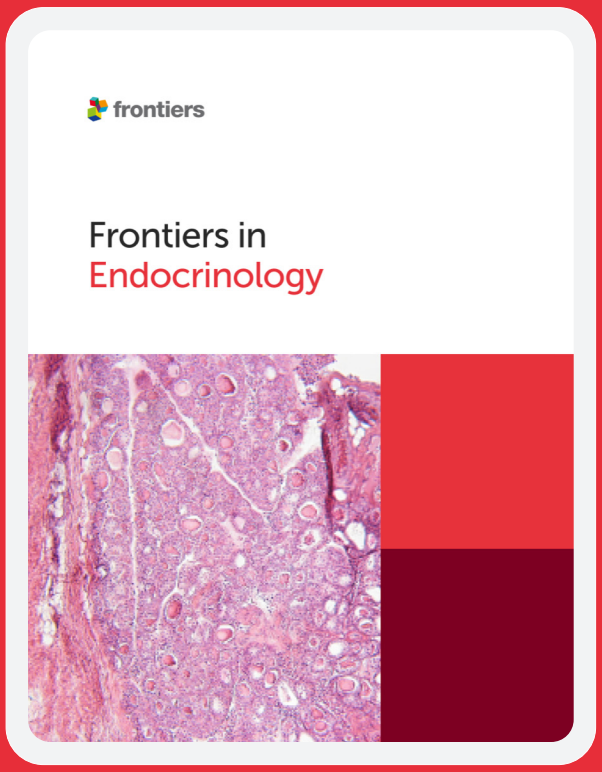
**Explores the endocrine system to find new therapies for key health issues**

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