


East Asian Mendelian-Randomization Evidence Linking PCOS to Gestational Diabetes Mellitus

Jingshu Du^{1,2}, Liwei Xing³, Yang Chen⁴, Zhe He⁵, Lanping Zhong², Rong Zhao^{1,3} 

¹Graduate School, The First School of Clinical Medicine, Nanjing University of Chinese Medicine, Nanjing, Jiangsu, People's Republic of China;

²Department of Traditional Chinese Medicine, The First People's Hospital of Yunnan Province, Kunming, Yunnan, People's Republic of China;

³Postgraduate Affairs Office, The First School of Clinical Medicine, Yunnan University of Chinese Medicine, Kunming, Yunnan, People's Republic of China; ⁴Department of Gynecology, Yunnan Provincial Hospital of Traditional Chinese Medicine, Kunming, Yunnan, People's Republic of China;

⁵Department of Acupuncture, Second Affiliated Hospital of Yunnan University Chinese Medicine, Kunming, Yunnan, People's Republic of China

Correspondence: Rong Zhao, Graduate School, The First School of Clinical Medicine, Nanjing University of Chinese Medicine, 138 Xianlin Avenue, Qixia District, Nanjing, Jiangsu, People's Republic of China, Email zhaorongkm@163.com

Background: Previous Mendelian randomization (MR) studies explored causal links between polycystic ovary syndrome (PCOS) and pregnancy complications and adverse outcomes mainly in European populations, with limited generalizability. This two-sample MR study addresses this limitation.

Methods: Independent genetic instruments for PCOS were selected at genome-wide significance ($P < 5 \times 10^{-8}$) and statistical pruned using LDTrait. East Asian (EAS)-based genome-wide association study data were analyzed for gestational diabetes mellitus (GDM), preeclampsia, intrahepatic cholestasis of pregnancy, miscarriage, preterm birth, and gestational age at birth. The inverse variance-weighted (IVW) method with Bonferroni correction served as the primary analysis, supplemented by pleiotropy and heterogeneity tests. Multivariable MR (MVMR) assessed direct effects by adjusting for confounders.

Results: Genetically predicted PCOS was causally linked to increased GDM risk (OR=1.203, 95% CI: 1.00–1.433), with consistent evidence across debiased IVW (OR=1.207, 95% CI: 1.008–1.444), Bayesian Weighted MR (OR=1.204, 95% CI: 1.007–1.440), Contamination Mixture (OR=1.231, 95% CI: 1.141–1.526), and distortion test (OR=1.203, 95% CI: 1.120–1.291). The strongest single nucleotide polymorphisms (SNP)-level effect was observed for rs1894116 (YAP1), followed by rs2268361 (FSHR). MVMR confirmed independence from body mass index (BMI), though the effect disappeared after adjusting for glucose and lipid traits. No causal relationships were found between PCOS and other outcomes ($P > 0.05$).

Conclusion: This MR study provides evidence supporting a potential causal association between genetically predicted PCOS and GDM in East Asian populations, in contrast to previous European-based findings, and highlights possible population-specific genetic differences with implications for precision risk assessment in pregnancy.

Keywords: polycystic ovary syndrome, mendelian randomization, East Asian, gestational diabetes mellitus

Introduction

Polycystic ovary syndrome (PCOS) represents a prevalent endocrine condition among women of reproductive age, typically manifesting as irregular menstruation, elevated androgen levels (hyperandrogenism), and ovarian polycystic morphology.¹ Data from the World Health Organization (WHO) indicate that globally, approximately 6% to 13% of women within reproductive age groups are affected by PCOS; nevertheless, the condition remains undiagnosed in up to 70% of these individuals. PCOS significantly contributes to infertility resulting from anovulation.² Furthermore, the health and economic implications of PCOS are profound, with healthcare-associated expenditures in the United States alone approximating 8 billion USD as of 2020.³

Women with PCOS face a significantly elevated risk of pregnancy-related complications, including gestational diabetes mellitus (GDM), gestational hypertension, preeclampsia, miscarriage, and intrahepatic cholestasis of pregnancy (ICP). Specifically, compared to women without PCOS, those with PCOS have an approximately 2.4-fold higher risk of GDM, a 2.2-fold higher risk of GH, a 2.3-fold increased risk of PE, and a 49% increased risk of miscarriage.⁴ Although

numerous observational studies have reported associations between PCOS and adverse pregnancy outcomes, these studies are limited by their inability to fully exclude confounding factors such as obesity, insulin resistance, and lifestyle, thereby hindering the establishment of definitive causal relationships. In recent years, Mendelian randomization (MR), which utilizes genetic variations as instrumental variables (IVs), has emerged as a powerful approach to investigate potential causal relationships. These genetic variants are randomly assigned at conception, analogous to the randomization process used in controlled clinical trials, and are independent of environmental exposures and other genetic confounders, substantially reducing traditional confounding biases.⁵ Consequently, MR serves as an effective methodological strategy to mitigate confounding effects and address potential reverse causality. Recently, numerous MR studies conducted in populations of European (EUR) ancestry have confirmed causal relationships between PCOS and various pregnancy complications and adverse pregnancy outcomes, notably demonstrating a causal relationship between PCOS and an increased risk of gestational hypertension independent of body mass index (BMI).^{6–9}

However, considerable heterogeneity exists among different populations regarding genetic architecture, prevalence, and clinical phenotypes of PCOS.¹⁰ Additionally, significant differences in risk-associated genetic loci across ancestries have been reported; for example, the FSHB gene is notably implicated in EUR populations, whereas variants in THADA and DENND1A are more prominent in East Asian (EAS) populations.^{11,12} Previous MR studies examining causal associations between PCOS and pregnancy outcomes have exclusively focused on EUR populations,^{6–8} thus limiting the generalizability of their findings due to ancestry-specific genetic structures and risk factor distributions. To address this evidence gap, the present study employs ancestry-specific genetic data from EAS populations using a two-sample MR approach. The aim is to validate previous findings derived from EUR populations and explore the applicability and generalizability of these associations in EAS populations.

Methods

Study Design

This study adheres to the guidelines of the STROBE-MR.¹³ Figure 1 provides a detailed overview of the research content and analytical steps. The present MR study utilized GWAS summary statistics specific to EAS ancestry, distinct from EUR populations, to investigate the causal relationship between genetic susceptibility to PCOS and adverse pregnancy outcomes, including miscarriage, preterm birth (PTB), and gestational age at birth, as well as pregnancy complications such as preeclampsia, GDM, and ICP. This approach allows comparison with prior MR studies and assesses the generalizability of these findings. Genetic variants selected as IVs in this study must satisfy the following core MR assumptions: 1) relevance—the genetic variants must be significantly associated with the exposure; 2) independence—no association with confounders; and 3) exclusion restriction—genetic variants affect outcomes exclusively through the exposure (Figure 1A).¹⁴

Data Sources

Detailed information regarding the datasets used in this study is provided in Table 1. Additionally, there was no sample overlap between exposure and outcome datasets, avoiding potential winner's curse bias. This study exclusively utilized publicly available GWAS summary statistics, and all original investigations had already secured the necessary ethical approvals. Consequently, this secondary analysis did not require additional ethical approval, clinical trial registration, or informed consent. According to Article 32, Items 1 and 2, of the Measures for Ethical Review of Life Science and Medical Research Involving Human Subjects (effective February 18, 2023, China), projects may be exempt from ethical review under either of the following circumstances: (1) the study relies on lawfully obtained public data or on observations that do not interfere with public behavior; or (2) the study uses anonymized information or data. As this research satisfies both criteria, formal ethical review was unnecessary.

PCOS and Genetic Instrument Selection

Genetic instruments for PCOS were derived from publicly available GWAS conducted by Shi et al, which focused on Han Chinese women.¹⁵ PCOS was defined according to the Rotterdam criteria established in 2003.¹⁶ This GWAS

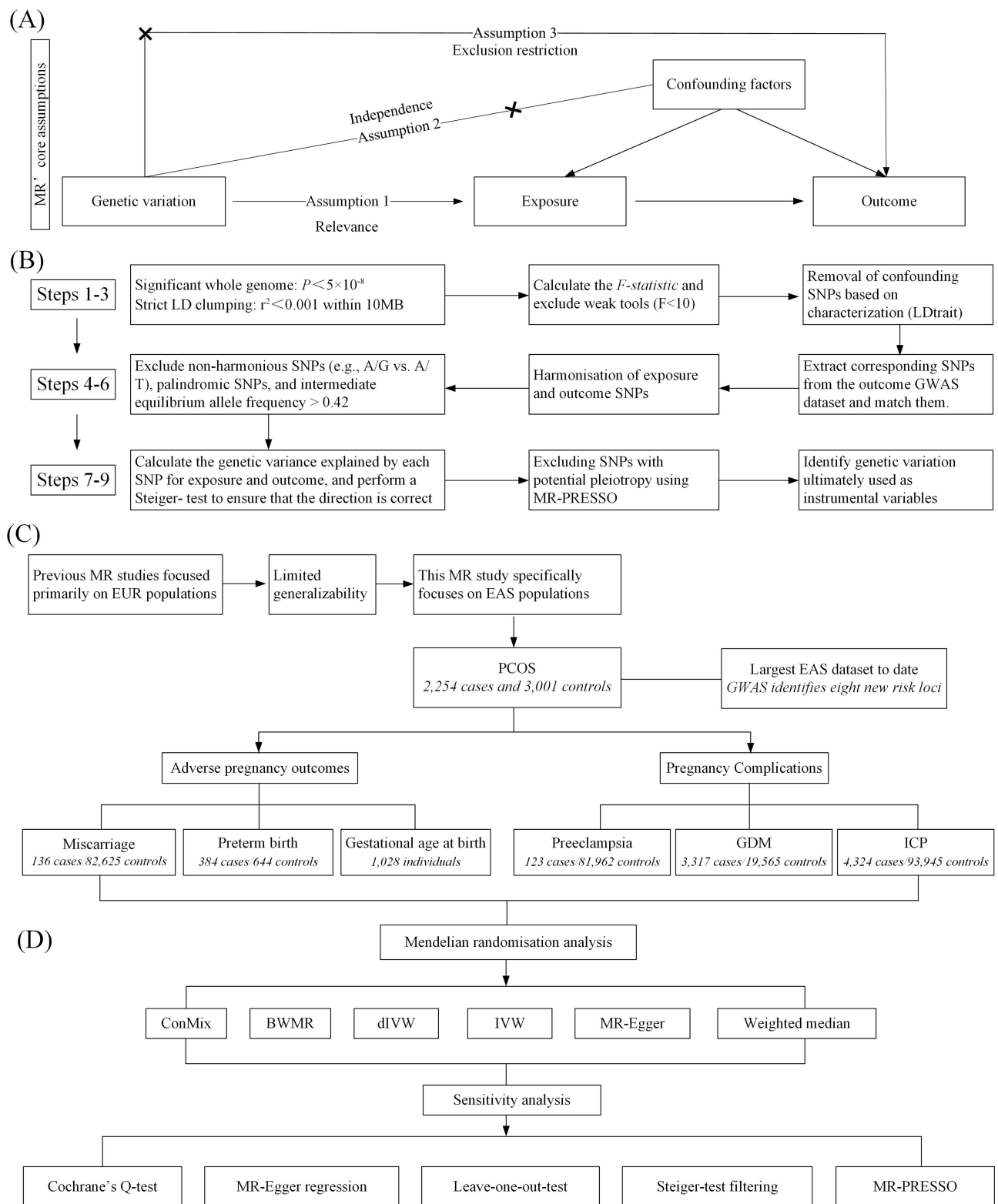


Figure 1 Study Design. **(A)** Genetic variants were used as instrumental variables for exposure phenotypes, which are required to meet the three core assumptions of Mendelian Randomization (MR): relevance, Independence, and exclusion restriction. **(B)** A nine-step process was applied to select valid genetic variants as IVs for causal inference. Detailed procedures are described in the Methods section. **(C)** Data sources and conceptual framework: This study focuses on EAS populations to enhance the generalizability of previous findings. The exposure phenotype, PCOS, was obtained from the largest EAS-specific GWAS to date. Outcomes were categorized into two groups: adverse pregnancy outcomes (including miscarriage, preterm birth, and gestational age at birth) and pregnancy complications (including preeclampsia, GDM, and ICP). **(D)** All MR methods used for causal inference and accompanying sensitivity analyses.

Table 1 Detailed Information of Data Sources

Phenotype	GWAS Catalog ID	Ref	Ancestry	Participants
PCOS	GCST001634	22885925	EAS	2,254 cases and 3,001 controls
Miscarriage	GCST90018566	34594039	EAS	136 cases and 82,625 controls
PTB	GCST90268004	37311799	EAS	384 cases and 644 controls
Gestational age at birth	GCST90268005	37311799	EAS	1,028 individuals
Preeclampsia	GCST90018686	34594039	EAS	123 cases and 81,962 controls
GDM	GCST90297786	38372780	EAS	3,317 cases and 19,565 controls
ICP	GCST90454199	39547589	EAS	4,324 cases and 93,945 controls
HDL-C	GCST90239651	34887591	EAS	146,492 individuals
LDL-C	GCST90239657	34887591	EAS	146,492 individuals
TC	GCST90239675	34887591	EAS	146,492 individuals
TG	GCST90239663	34887591	EAS	146,492 individuals
BMI	GCST90018727	34594039	EAS	163,835 individuals
FI	GCST90002237	34059833	EAS	29,792 individuals
HbA1c	GCST90002243	34059833	EAS	33,307 individuals

involved a two-stage discovery and replication strategy, encompassing 2,254 PCOS cases and 3,001 controls, leading to the identification of eight novel genetic loci significantly associated with PCOS (Table S1). These loci predominantly implicate biological pathways related to insulin signaling, sex hormone regulation, and type 2 diabetes mellitus. To mitigate population stratification, principal component analysis (PCA) was performed, and logistic regression models were adjusted accordingly. SNP genotyping data were integrated and imputation was conducted around newly identified loci. Meta-analysis was performed using fixed-effect models. Original GWAS data underwent inverse-normal transformation, and MR effect estimates were presented per standard deviation (SD) increase in PCOS susceptibility.

Initially, genome-wide significant ($P < 5 \times 10^{-8}$) and independent single nucleotide polymorphisms (SNPs) ($r^2 < 0.001$) were selected for MR analysis. Linkage disequilibrium [LD] clumping was conducted based on the EAS dataset from the 1,000 Genomes Project Phase 3 to account for ancestry-specific genetic variations. The LDtrait tool (<https://ldlink.nih.gov/?tab=ldtrait>) was utilized to evaluate SNP characteristics, removing those associated with confounders to strictly adhere to assumption two. The strength of genetic instruments was assessed by calculating the F-statistic. Matching genetic variants between exposure and outcome datasets involved harmonization of effect alleles, excluding incompatible alleles (eg, C/T vs C/G) and those with intermediate allele frequency (EAF > 0.42). The Steiger test, considering sample sizes and variance explained by SNPs in both exposure and outcome, ensured the correct directionality of causation. Finally, pleiotropic outliers were identified and excluded to finalize the genetic instrument selection (Figure 1B). Notably, proxy SNPs were not used to impute missing data to ensure precision.

Adverse Pregnancy Outcomes

Summary data for miscarriage were obtained from the GWAS conducted by Sakaue et al,¹⁷ integrating BioBank Japan (BBJ), UK Biobank, and FinnGen consortium data across more than 628 phenotypes (Figure 1C). Quality control involved PCA to control for population stratification, employing linear or logistic regression adjusted for age, sex, and the top ten principal components, with inverse-normal transformation for quantitative traits. Phenotypes from BBJ underwent rigorous validation and standardized processing across cohorts. For the current analysis, EAS-specific data (136 cases, 82,625 controls) were selected. It is important to note that miscarriage was defined based on hospital discharge diagnoses of spontaneous abortion (ICD-10: O03, ICD-9: 634, ICD-8: 643). This definition specifically excludes induced or elective abortions.

GWAS summary data for PTB and gestational age at birth were sourced from Hasegawa et al,¹⁸ involving 384 PTB cases and 644 controls from the Japanese population. PTB was defined as delivery occurring before 37 weeks of gestation, evaluated as a maternal effect. Additionally, gestational age at birth (maternal effect) was analyzed as a continuous outcome to determine whether maternal PCOS status was dose-dependently associated with shortened

gestational duration. This provided a more sensitive measure of gestational timing compared to the binary PTB classification.

Pregnancy Complications

Summary statistics for GDM were obtained from the GWAS conducted by Zhen et al,¹⁹ utilizing non-invasive prenatal testing (NIPT) data collected from 30,699 pregnant women during routine prenatal examinations in Shenzhen, China. Pregestational diabetes cases were excluded, and GDM was diagnosed based on abnormal glucose tolerance test (OGTT) results at 24–28 weeks of gestation. This dataset included 3,317 GDM cases and 19,565 controls, with covariate adjustments for gestational age, maternal age, and BMI. The GWAS identified four genetic loci significantly associated with GDM among individuals of EAS ancestry.

Summary statistics for ICP were sourced from a GWAS by Liu et al,²⁰ which analyzed NIPT sequencing data from 98,269 pregnant women from two hospitals in Shenzhen, China, including 4,324 ICP cases and 93,945 controls. Participants with pregnancy-induced hypertension, viral hepatitis, or liver injury were excluded. Covariates such as gestational age, maternal age, and the top ten principal components were incorporated into the GWAS to control for population stratification. Four significant genetic loci associated with ICP risk were identified, including an EAS-specific variant at 14q24.1 (p.Ser267Phe mutation in the SLC10A1 gene), associated with a 16.56-fold increased risk per risk allele, highlighting the unique genetic susceptibility of ICP in EAS populations.

Summary statistics for preeclampsia were extracted from the GWAS conducted by Sakaue et al,¹⁷ as previously described. Preeclampsia was defined by ICD-10-O14, phecode 642, and the FinnGen endpoint O15_PREECLAMPS, comprising 123 cases and 81,962 controls of EAS ancestry.

Statistical Analysis

This study performed two-sample, univariable MR analyses for the primary outcomes, including adverse pregnancy outcomes and pregnancy complications, utilizing the inverse variance weighted (IVW) method²¹ (Figure 1D). Complementary methods, such as MR-Egger, Weighted Median, Contamination Mixture (ConMix),²² Debiased IVW (dIVW),²³ and Bayesian Weighted Mendelian Randomization (BWMR),²⁴ were applied to validate the robustness of IVW results based on differing assumptions and features. Due to multiple testing, statistical significance was adjusted using Bonferroni correction, with $P < 0.008$ (0.05/6) indicating strong causal evidence, and $0.008 < P < 0.05$ considered suggestive causal evidence.

Heterogeneity among SNPs was assessed using Cochran's Q test, employing both IVW and MR-Egger methods.²⁵ A significant Q test ($P < 0.05$) indicated heterogeneity, and the I^2 statistic was calculated to assess the appropriateness of the main effect model. MR-Egger intercept²⁶ and Mendelian Randomization Pleiotropy Residual Sum and Outlier (MR-PRESSO)²⁷ global tests evaluated horizontal pleiotropy and outliers, with $P < 0.05$ considered indicative of pleiotropic bias. Additionally, MR-PRESSO's distortion test validated IVW estimates by fitting residual sums of squares (RSS).

Leave-one-out sensitivity analysis determined if individual SNPs disproportionately influenced observed causal relationships. Statistical power for significant results was calculated using the mRNA online platform (<https://shiny.cnsgenomics.com/mRnd/>). Forest plots visualized individual SNP effects, while scatter plots assessed the directionality of causal effects across methods. All statistical analyses were conducted using R (version 4.2.3), with the following R packages: MVMR, TwoSampleMR, MendelianRandomization, BWMR, and MRPRESSO.

Results

Selection, Strength, and Evaluation of Instrumental Variables

A total of 10 SNPs with genome-wide significance were initially extracted from eight PCOS-associated risk loci to fulfill assumption one (Table S1). LD clumping subsequently excluded rs2349415 (FSHR gene) and rs2272046 (HMG2A gene). Additionally, variant profiling using LDtrait (Table S2) led to the removal of rs6022786 due to its association with diabetes-related traits. MR-PRESSO analysis identified no outliers, effectively minimizing potential horizontal pleiotropy and strictly adhering to assumption two. The Steiger test, which evaluated effect sizes, variances, and sample sizes of

Table 2 Summary of Genetic Variation as the Genetic Instrument for PCOS

SNP	chr	pos	Nearby Gene(s)	A1	A2	eaf	Beta	se	Pval	R ²	F-Statistic
rs2268361	2	48974473	FSHR	T	C	0.496	-0.139	0.020	9.89E-13	0.97%	51
rs4385527	9	94886305	C9orf3	A	G	0.219	-0.174	0.030	5.87E-09	1.04%	55
rs3802457	9	94979054	C9orf3	A	G	0.096	-0.261	0.035	5.28E-14	1.19%	63
rs1894116	11	102199908	YAP1	G	A	0.194	0.239	0.024	1.08E-22	1.79%	96
rs705702	12	55996852	RAB5B, SUOX	G	A	0.245	0.239	0.023	8.64E-26	2.11%	113
rs4784165	16	52313907	TOX3	G	T	0.325	0.140	0.021	3.64E-11	0.86%	45
rs2059807	19	7166098	INSR	G	A	0.301	0.131	0.023	1.09E-08	0.72%	38

Notes: A1 represents the effect allele, and A2 represents the other allele.

SNPs for both exposure and outcome, validated the causal direction stipulated in assumption three (Table S3). Ultimately, seven SNPs were selected as instrumental variables, collectively explaining 8.67% of the genetic variance (R²) in PCOS (Table 2). Each SNP exhibited an F-statistic greater than 10 (range: 38–113), with an average exceeding 66, confirming the strength and validity of these genetic instruments. Detailed SNP characteristics across analyses are summarized in Table S4.

Two-Sample Univariable Mendelian Randomization Causal Inference

Figure 2 illustrates the scatter plots of effect directions for all analytical methods, and detailed results are presented in Table S5. Univariable MR analysis using the IVW method indicated that genetically predicted PCOS increased the risk of GDM by 20.3% per SD increase among individuals of EAS ancestry (odds ratio [OR]=1.203, 95% CI: 1.009–1.433, P=0.039; Figure 3A). The SNP rs1894116 in the YAP1 gene exhibited the largest risk effect, followed by rs2268361 in

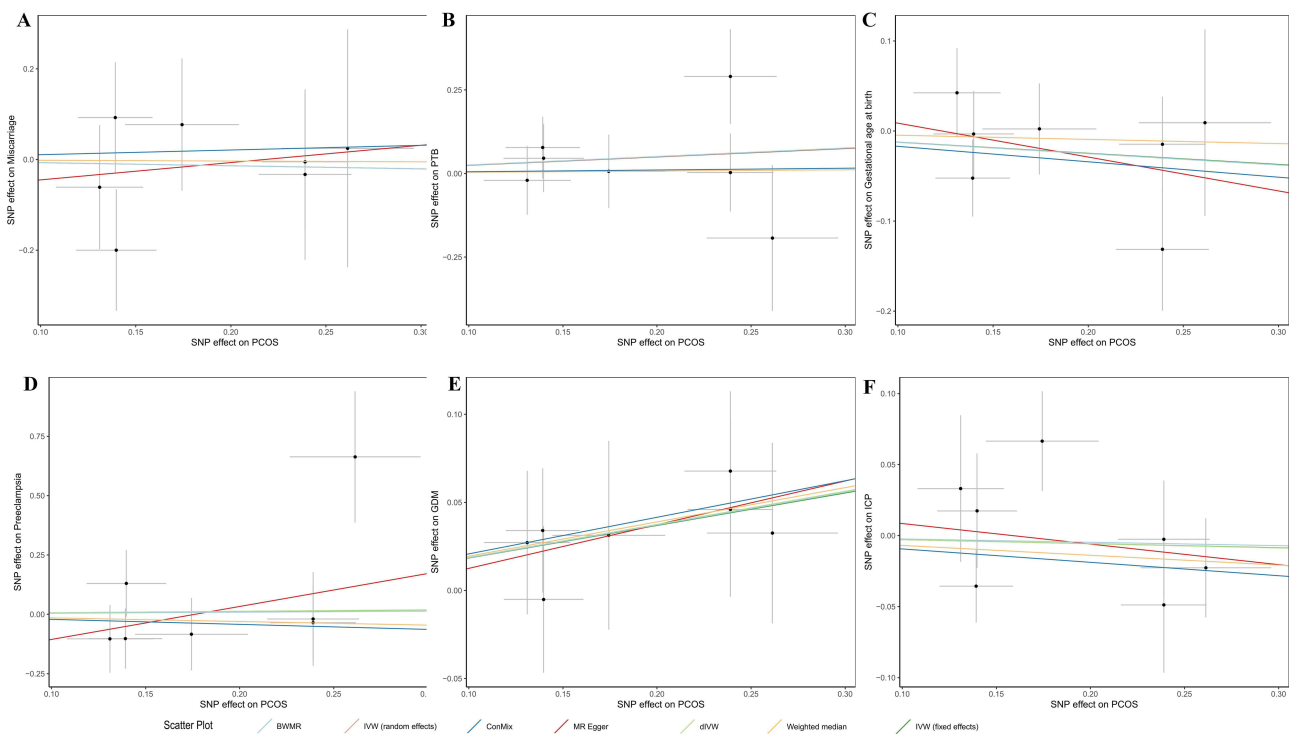


Figure 2 Scatter plot of SNP-specific causal estimates for the effect of PCOS on adverse pregnancy outcomes and pregnancy complications. (A) PCOS on miscarriage (B) PCOS on PTB (C) PCOS on gestational age at birth (D) PCOS on preeclampsia (E) PCOS on GDM (F) PCOS on ICP. Each dot represents the Wald ratio estimate for a single SNP, with horizontal and vertical lines indicating the standard errors of the SNP's effect on PCOS and miscarriage, respectively. The colored lines correspond to different Mendelian randomization (MR) methods used to estimate the causal effect. The overall slope of each line reflects the causal estimate derived from the corresponding method.

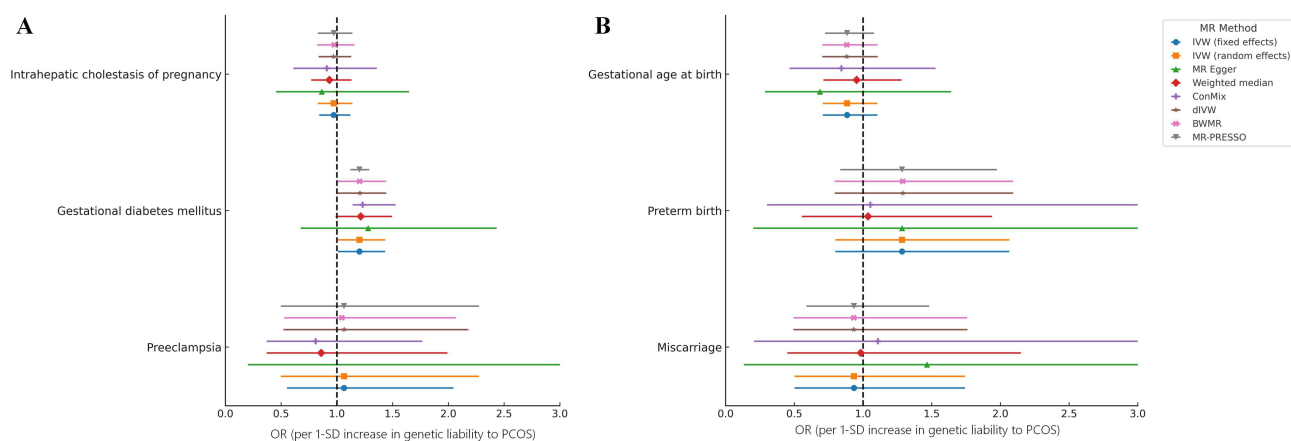


Figure 3 Forest plot of causal associations between PCOS and pregnancy complications and adverse pregnancy outcomes using two-sample univariable Mendelian randomization. ORs and 95% CIs are presented for the associations between genetically predicted PCOS and (A) pregnancy complications and (B) adverse pregnancy outcomes, expressed per 1-SD increase in genetic liability to PCOS. Each color represents a different MR method, including inverse-variance weighted (IVW; fixed- and random-effects), MR-Egger, weighted median, dIVW, BWMR, ConMix, and MR-PRESSO. The vertical dashed line indicates the null hypothesis (OR = 1.0). Due to multiple testing across six outcomes, statistical significance was adjusted using Bonferroni correction, with $P < 0.008$ (0.05/6) indicating strong causal evidence and $0.008 \leq P < 0.05$ considered suggestive causal evidence. Sample sizes for each outcome were as follows: intrahepatic cholestasis of pregnancy (4,324 cases and 93,945 controls), gestational diabetes mellitus (3,317 cases and 19,565 controls), preeclampsia (123 cases and 81,962 controls), miscarriage (136 cases and 82,625 controls), preterm birth (384 cases and 644 controls), and gestational age at birth (1,028 individuals).

the FSHR gene (Figure S1). Validation analyses employing ConMix (OR=1.231, 95% CI: 1.141–1.526, $P=0.040$), dIVW (OR=1.207, 95% CI: 1.008–1.444, $P=0.040$), and BWMR (OR=1.204, 95% CI: 1.007–1.440, $P=0.042$) provided consistent evidence supporting causality. Statistical power exceeded 85% across all methods, affirming the robustness of these findings (Table S6). However, the association did not reach statistical significance after Bonferroni correction and was therefore considered suggestive. Furthermore, no causal associations were detected between PCOS and preeclampsia or ICP ($P > 0.05$).

The univariable MR analyses also indicated no causal relationships between PCOS and adverse pregnancy outcomes, including miscarriage, PTB, and gestational age at birth ($P > 0.05$; Figure 3B).

Sensitivity Analyses

No evidence of heterogeneity was detected by either of the two methods of Cochran's Q test or by the I^2 statistic ($P > 0.05$); therefore, a fixed-effects model was employed (Table S7). MR-Egger regression (Table S8) and MR-PRESSO global tests (Table S9) similarly showed no significant evidence of horizontal pleiotropy ($P > 0.05$). Notably, MR-PRESSO distortion test further supported the PCOS-GDM causal relationship (OR=1.203, 95% CI: 1.120–1.291, $P=0.002$), which remained statistically significant after Bonferroni correction, yielding narrower confidence intervals likely due to methodological differences in standard error estimation (Table S10). Leave-one-out analyses confirmed that the observed causal relationship was not driven by a single SNP, underscoring the robustness of our findings (Table S11 and Figure S2). Figure S3 presents forest plots summarizing the effect contribution of each individual SNP in each analysis, whereas Figure S4 displays the corresponding funnel plots.

Further Analysis: Multivariable and Mediation Mendelian Randomization

To evaluate whether the PCOS-GDM association was indirectly mediated through common metabolic pathways, we conducted a multivariable Mendelian randomization (MVMR) analysis. Given the close association between PCOS and various metabolic abnormalities, we included seven metabolic traits (obesity, glucose metabolism traits, and lipid profiles) in the MVMR model. Specifically, BMI data for individuals of EAS ancestry ($n=163,835$) were sourced from Sakaue et al,¹⁷ glucose metabolism traits ($n_{\text{MAX}}=33,307$) from the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC),²⁸ and lipid traits ($n=146,492$) from the Global Lipids Genetics Consortium (GLGC).²⁹

Table 3 Summary of Results From MVMR to Investigate Direct Causal Relationships

Model	Direct Effect	OR	95% CI	Pval
Adjusted for BMI	0.184	1.202	(1.009–1.433)	0.039
Adjusted for FI	0.194	1.214	(0.994–1.484)	0.058
Adjusted for HbA1c	0.180	1.197	(0.985–1.454)	0.070
Adjusted for HDL-C	0.244	1.277	(0.772–2.111)	0.341
Adjusted for LDL-C	0.217	1.243	(0.999–1.546)	0.051
Adjusted for TG	0.165	1.179	(0.963–1.444)	0.111
Adjusted for TC	0.197	1.218	(0.977–1.520)	0.080

The MVMR analysis revealed that PCOS independently increased GDM risk after adjusting for BMI ($P < 0.05$) (Table 3). However, this causal relationship weakened and became non-significant after further adjustments for glucose traits (glycated hemoglobin [HbA1c], fasting insulin [FI]) and lipid traits (high-density lipoprotein cholesterol [HDL-C], low-density lipoprotein cholesterol [LDL-C], triglycerides [TG], and total cholesterol [TC]; $P > 0.05$).

Building upon this foundation, the present study further employed a two-step mediation MR analysis to explore the potential mediating roles of seven metabolic phenotypes in the causal pathway between PCOS and GDM. This approach estimated the total effect of PCOS on GDM, the direct effect excluding each mediator, and the indirect (mediating) effect through the mediator, while also calculating the proportion attributable to the mediating effect. As shown in Table S12, in all seven models, the 95% confidence intervals for the mediating effects and their proportions spanned zero ($P > 0.05$). Consequently, no single metabolic factor exhibited a statistically significant mediating role between PCOS and GDM.

Discussion

This MR study is the first genetic investigation focusing on individuals of EAS ancestry to explore the causal relationship between PCOS and pregnancy complications and adverse pregnancy outcomes. Our results demonstrated a causal association between genetic susceptibility to PCOS and an increased risk of GDM. Consistent causal evidence was obtained through two primary models and confirmed using up to four validation methods. SNP effect analyses revealed that rs1894116 in the YAP1 gene had the largest contribution to this association, followed by rs2268361 in the FSHR gene. MVMR analyses showed that the causal relationship between PCOS and GDM was independent of BMI; however, after adjusting for lipid and glucose metabolism indicators, the association was no longer statistically significant. Notably, these findings sharply contrast with previous MR studies predominantly conducted in EUR populations, which generally did not observe significant causal effects of PCOS on GDM.^{7–9} Conversely, our findings align closely with meta-analyses of multiple observational studies.⁴ For other phenotypes, our study found no nominally significant causal associations between PCOS and ICP, miscarriage, gestational age at birth, or preeclampsia, consistent with previous MR findings in EUR populations.

Although previous two-sample MR studies in EUR populations have not identified significant causal associations between PCOS and GDM,^{7–9} our study provides novel evidence of such a link in an EAS population, suggesting potential population specificity. This difference may result from genetic heterogeneity between EAS and EUR populations, including allele frequency variations and discrepancies in the distribution of risk loci. The PCOS GWAS data used in our study explained 8.06% of phenotypic variance, potentially offering a more accurate representation of genetic risk distribution in the EAS population. Furthermore, prior studies have reported that EAS women with PCOS often exhibit non-obese phenotypes characterized by higher insulin resistance, specifically higher 2-hour insulin and glucose levels compared with Caucasian women, thus rendering them more susceptible to GDM risk.³⁰ Additionally, MR studies have shown associations between PCOS genetic susceptibility and hypertensive disorders of pregnancy in EUR populations.^{6–8} However, due to the lack of GWAS data on pregnancy-induced hypertension in EAS populations, we could not evaluate this outcome. Future genomic studies are therefore essential to investigate pregnancy-related phenotypes in EAS

populations, especially hypertensive disorders of pregnancy and placental dysfunction, to more comprehensively assess pregnancy risks associated with PCOS across diverse populations.

This study identified SNPs rs1894116 (YAP1) and rs2268361 (FSHR) as having the greatest contributions to the PCOS–GDM association. Specifically, the rs1894116 allele in the YAP1 gene is known to significantly increase PCOS risk (OR ~1.27).³¹ YAP1 protein, a core component of the Hippo signaling pathway, is strongly associated with metabolic disturbances commonly observed in PCOS patients, including insulin resistance and dyslipidemia.³² Activation of YAP1 may exacerbate gestational glucose metabolism disorders by promoting aberrant insulin signaling, lipid deposition, and ovarian dysfunction, thereby elevating GDM incidence. On the other hand, rs2268361 in the FSHR gene is a recognized PCOS susceptibility SNP, whose T allele correlates with reduced follicle-stimulating hormone (FSH) levels.³³ FSHR primarily regulates ovarian follicle development and estrogen production through receptor-mediated activity. Thus, variations in rs2268361 may disrupt ovarian hormone balance by decreasing FSH receptor functionality, impairing insulin sensitivity and glucose metabolism stability, thereby lowering metabolic compensation during pregnancy. Future research should validate the effects of these loci across diverse populations or independent cohorts and investigate early interventions targeting these critical pathways.

The MVMR analysis indicated that the causal association between PCOS and GDM risk is independent of BMI, consistent with previous findings.^{34,35} However, when we additionally adjusted for glucose metabolism and lipid traits, the PCOS–GDM association was attenuated and became non-significant. This pattern is compatible with the hypothesis that insulin resistance and dyslipidemia—hallmark metabolic features of PCOS⁹—may lie on the causal pathway between PCOS and GDM. To further investigate this possibility, we performed a two-step mediation MR analysis using BMI, FI, HbA1c, HDL-C, LDL-C, TG, and TC as candidate mediators. Across all seven models, the estimated indirect effects and mediation proportions were small, with 95% confidence intervals crossing zero, indicating that no single metabolic trait could be identified as a statistically robust mediator in the current setting. These findings suggest that any mediation is likely modest and/or distributed across multiple correlated metabolic pathways, and may also reflect limited statistical power of the available East Asian GWAS. Therefore, our results support a plausible but as yet unconfirmed mediating role of metabolic dysfunction in the PCOS–GDM relationship, which warrants verification in larger, well-phenotyped cohorts and multi-omics mediation frameworks.

From a translational perspective, delineating the genetic architecture linking PCOS to GDM in East Asian women may inform the future development of polygenic risk scores or integrated prediction models that combine genetic liability with clinical and metabolic factors. Such tools could help stratify women with PCOS according to their GDM risk and identify those who might benefit most from closer metabolic surveillance and earlier GDM screening. However, given the current limitations in sample size and ancestry diversity, these applications remain exploratory and require validation in larger, prospective cohorts before clinical implementation.

This study, uniquely conducted within an EAS genetic framework, systematically explored causal relationships between PCOS and multiple adverse pregnancy outcomes, supplementing prior EUR-centric MR investigations and enhancing the generalizability of the findings. In addition, the GWAS summary statistics for pregnancy outcomes were mainly derived from NIPT-recruited pregnancy cohorts, which typically include women who undergo prenatal screening and may not fully represent the general pregnant population in terms of age, underlying risk profile, and healthcare access. This potential selection bias may limit the generalizability of our estimates to all pregnant women with PCOS and should be considered when interpreting the clinical implications of our findings. However, for several pregnancy outcomes such as preeclampsia, miscarriage, and intrahepatic cholestasis of pregnancy, the currently available East Asian GWAS are based on relatively small case numbers. Although we used the largest EAS datasets that could be identified through an updated search of published literature and public resources up to November 2025, the sample sizes remain modest compared with those of European cohorts, which inevitably reduces statistical power and may increase the likelihood of false-negative findings for these outcomes. These results should therefore be interpreted with caution and considered hypothesis-generating until larger EAS and multi-ancestry GWAS become available for replication. Nevertheless, due to the predominance of European and East Asian GWAS summary data, relevant data for other ethnic groups (eg, African [AFR] and Great Middle Eastern [GME] populations) remain unavailable or insufficiently explored. This limitation precludes external validation across broader populations. Future expansions of multi-ancestral genetic

research and accumulation of large-scale non-European data promise to further replicate and extend our findings. Additionally, six analytical methods consistently supported the causal relationship between PCOS and GDM regarding direction and significance. Power analyses demonstrated sufficient statistical robustness without detectable biases such as winner's curse, pleiotropy, or substantial heterogeneity. Importantly, MR-PRESSO distortion tests retained statistical significance after Bonferroni correction ($P = 0.002$), further confirming our findings' reliability and robustness. Future studies should employ methodological triangulation from multiple data sources, within-family MR analyses accounting for relatedness, and non-linear causal inference utilizing individual-level data to further substantiate and elucidate the causal mechanisms identified.

Conclusion

To encapsulate, this study is the first to systematically evaluate the causal relationship between PCOS and pregnancy complications and adverse pregnancy outcomes using MR based on genetic data from EAS populations. In contrast with several previous EUR-based MR studies, our findings provide suggestive evidence that genetic liability to PCOS is associated with an increased risk of GDM in EAS women, independent of BMI. However, this association appears to be at least partly related to co-existing metabolic disturbances, and no single metabolic trait showed a statistically robust mediating effect in two-step mediation analyses. In addition, no causal associations were observed between PCOS and ICP, miscarriage, gestational age at birth, or preeclampsia, broadly consistent with prior research. From a clinical perspective, our findings support closer metabolic assessment and earlier or more intensive GDM screening (eg, advancing or repeating OGTT) in EAS women with PCOS, even in the absence of overt obesity, as well as optimisation of metabolic status and individualized pre-conception counselling to potentially mitigate subsequent pregnancy risk. Future multi-ancestry GWAS and Mendelian randomization studies are needed to validate these findings and to assess their generalizability beyond EAS populations.

Data Sharing Statement

The data used in this study are publicly available from previously published genome-wide association studies (GWAS). All summary statistics can be accessed through the cited references or retrieved from the data sources listed in [Table 1](#) of this manuscript.

Ethics Approval and Consent to Participate

This study is a secondary analysis of publicly available data and does not require ethical approval or clinical registration.

Acknowledgments

We thank all GWAS participants and researchers for making summary statistics publicly available.

Funding

This work was co-supported by the National Natural Science Foundation of China (82460974), Key Project of Basic Research Program of the Science and Technology Department of Yunnan Province (202501AS070140), Research Fund Project of the Education Department of Yunnan Province (2024J0517, 2025Y0619).

Disclosure

The authors report no competing interests in this study.

References

1. Anagnostis P, Tarlatzis BC, Kauffman RP. Polycystic ovarian syndrome (PCOS): long-term metabolic consequences. *Metabolism*. 2018;86:33–43. doi:10.1016/j.metabol.2017.09.016
2. Polycystic ovary syndrome. <https://www.who.int/news-room/fact-sheets/detail/polycystic-ovary-syndrome>. Accessed June 20, 2025.
3. Riestenberg C, Jagasia A, Markovic D, Buyalos RP, Azziz R. Health care-related economic burden of polycystic ovary syndrome in the United States: pregnancy-related and long-term health consequences. *J Clin Endocrinol Metab*. 2022;107(2):575–585. doi:10.1210/clinem/dgab613

4. Bahri Khomami M, Shorakae S, Hashemi S, et al. Systematic review and meta-analysis of pregnancy outcomes in women with polycystic ovary syndrome. *Nat Commun.* 2024;15:5591. doi:10.1038/s41467-024-49749-1
5. Tin A, Köttgen A. Mendelian randomization analysis as a tool to gain insights into causes of diseases: a primer. *J Am Soc Nephrol.* 2021;32(10):2400–2407. doi:10.1681/ASN.2020121760
6. Shao S, Xu Q, Zi Y, et al. The genetic association between polycystic ovary syndrome and the risk of hypertensive disorders of pregnancy: a mendelian randomization study. *Eur J Obstetrics Gynecol Reproduct Biol.* 2025;305:351–355. doi:10.1016/j.ejogrb.2024.12.043
7. Dang C, Wang S, Liu P, et al. Polycystic ovary syndrome and risk of pregnancy-induced hypertension, gestational diabetes mellitus, and preeclampsia: a Bayesian mendelian randomization study. *Diabetol Metabol Syndr.* 2025;17(1):204. doi:10.1186/s13098-025-01739-z
8. Ma Y, Cai J, Liu LW, et al. Causal relationships exist between polycystic ovary syndrome and adverse pregnancy and perinatal outcomes: a mendelian randomization study. *Front Endocrinol.* 2024;15:1327849. doi:10.3389/fendo.2024.1327849
9. Guixue G, Yifu P, Xiaofeng T, et al. Investigating the causal impact of polycystic ovary syndrome on gestational diabetes mellitus: a two-sample mendelian randomization study. *Front Endocrinol.* 2024;15:1337562. doi:10.3389/fendo.2024.1337562
10. Huang Z, Yong EL. Ethnic differences: is there an asian phenotype for polycystic ovarian syndrome? *Best Pract Res Clin Obstet Gynaecol.* 2016;37:46–55. doi:10.1016/j.bpobgyn.2016.04.001
11. Hyeon HS, Hong YS, Jeong K, Chung H, Lee H, Sung YA. Relationship between the characteristic traits of polycystic ovary syndrome and susceptibility genes. *Sci Rep.* 2020;10(1):10479. doi:10.1038/s41598-020-66633-2
12. Louwers YV, Stolk L, Uitterlinden AG, Laven JSE. Cross-ethnic meta-analysis of genetic variants for polycystic ovary syndrome. *J Clin Endocrinol Metab.* 2013;98(12):E2006–2012. doi:10.1210/jc.2013-2495
13. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *JAMA.* 2021;326(16):1614–1621. doi:10.1001/jama.2021.18236
14. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med.* 2008;27(8):1133–1163. doi:10.1002/sim.3034
15. Shi Y, Zhao H, Shi Y, et al. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet.* 2012;44(9):1020–1025. doi:10.1038/ng.2384
16. Group TRE sponsored P consensus workshop. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod.* 2004;19(1):41–47. doi:10.1093/humrep/deh098
17. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet.* 2021;53(10):1415–1424. doi:10.1038/s41588-021-00931-x
18. Hasegawa K, Kumasaka N, Nakabayashi K, et al. Genome-wide association study of preterm birth and gestational age in a japanese population. *Hum Genome Var.* 2023;10(1):19. doi:10.1038/s41439-023-00246-9
19. Zhen J, Gu Y, Wang P, et al. Genome-wide association and mendelian randomisation analysis among 30,699 Chinese pregnant women identifies novel genetic and molecular risk factors for gestational diabetes and glycaemic traits. *Diabetologia.* 2024;67(4):703–713. doi:10.1007/s00125-023-06065-5
20. Liu Y, Wei Y, Chen X, et al. Genetic study of intrahepatic cholestasis of pregnancy in Chinese women unveils East Asian etiology linked to historic HBV epidemic. *J Hepatol.* 2025;82(5):826–835. doi:10.1016/j.jhep.2024.11.008
21. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol.* 2013;37(7):658–665. doi:10.1002/gepi.21758
22. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nat Commun.* 2020;11(1):376. doi:10.1038/s41467-019-14156-4
23. Ye T, Shao J, Kang H. Debiased inverse-variance weighted estimator in two-sample summary-data Mendelian randomization. *Ann Stat.* 2021;49(4):2079–2100. doi:10.1214/20-AOS2027
24. Zhao J, Ming J, Hu X, Chen G, Liu J, Yang C. Bayesian weighted Mendelian randomization for causal inference based on summary statistics. *Bioinformatics.* 2020;36(5):1501–1508. doi:10.1093/bioinformatics/btz749
25. Kulinskaya E, Dollinger MB, Bjørkestøl K. On the moments of Cochran's Q statistic under the null hypothesis, with application to the meta-analysis of risk difference. *Res Synth Methods.* 2020;11(6):920. doi:10.1002/jrsm.1446
26. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol.* 2017;32(5):377–389. doi:10.1007/s10654-017-0255-x
27. 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):693–698. doi:10.1038/s41588-018-0099-7
28. Chen J, Spracklen CN, Marenne G, et al. The trans-ancestral genomic architecture of glycaemic traits. *Nat Genet.* 2021;53(6):840–860. doi:10.1038/s41588-021-00852-9
29. Graham SE, Clarke SL, Wu KHH, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature.* 2021;600(7890):675–679.
30. Chahal N, Quinn M, Jaswa EA, Kao CN, Cedars MI, Huddleston HG. Comparison of metabolic syndrome elements in white and asian women with polycystic ovary syndrome: results of a regional, American cross-sectional study. *F&S Reports.* 2020;1(3):305–313. doi:10.1016/j.xfre.2020.09.008
31. Lidaka L, Bekere L, Lazdane G, Lazovska M, Dzivite-Krisane I, Gailite L. Role of single nucleotide variants in the YAP1 gene in adolescents with polycystic ovary syndrome. *Biomedicines.* 2022;10(7):1688. doi:10.3390/biomedicines10071688
32. Li T, Zhao H, Zhao X, et al. Identification of YAP1 as a novel susceptibility gene for polycystic ovary syndrome. *J Med Genetic.* 2012;49(4):254–257. doi:10.1136/jmedgenet-2011-100727
33. Laven JSE. Follicle stimulating hormone receptor (FSHR) polymorphisms and polycystic ovary syndrome (PCOS). *Front Endocrinol.* 2019;10:23. doi:10.3389/fendo.2019.00023
34. Mills G, Badeghiesh A, Suarathana E, Baghlah F, Dahan MH. Polycystic ovary syndrome as an independent risk factor for gestational diabetes and hypertensive disorders of pregnancy: a population-based study on 9.1 million pregnancies. *Hum Reprod.* 2020;35(7):1666–1674. doi:10.1093/humrep/deaa099
35. Rees DA, Jenkins-Jones S, Morgan CL. Contemporary reproductive outcomes for patients with polycystic ovary syndrome: a retrospective observational study. *J Clin EndocrinolMetabol.* 2016;101(4):1664–1672. doi:10.1210/jc.2015-2682

International Journal of Women's Health

Publish your work in this journal

The International Journal of Women's Health is an international, peer-reviewed open-access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of women's healthcare including gynecology, obstetrics, and breast cancer. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-womens-health-journal>

Dovepress
Taylor & Francis Group