


RESEARCH

Open Access



Causal association between low vitamin D and polycystic ovary syndrome: a bidirectional mendelian randomization study

Bingrui Gao¹, Chenxi Zhang¹, Deping Wang^{1,2}, Bojuan Li¹, Zhongyan Shan¹, Weiping Teng¹ and Jing Li^{1*} 

Abstract

Background Recent studies have revealed the correlation between serum vitamin D (VD) level and polycystic ovary syndrome (PCOS), but the causality and specific mechanisms remain uncertain.

Objective We aimed to investigate the cause-effect relationship between serum VD and PCOS, and the role of testosterone in the related pathological mechanisms.

Methods We assessed the causality between serum VD and PCOS by using genome-wide association studies (GWAS) data in a bidirectional two-sample Mendelian randomization (TS-MR) analysis. Subsequently, a MR mediation analysis was conducted to examine the mediating action of testosterone in the causality between serum VD and PCOS. Ultimately, we integrated GWAS data with cis-expression quantitative loci (cis-eQTLs) data for gene annotation, and used the potentially related genes for functional enrichment analysis to assess the involvement of testosterone and the potential mechanisms.

Results TS-MR analysis showed that individuals with lower level of serum VD were more likely to develop PCOS (OR=0.750, 95% CI: 0.587–0.959, $P=0.022$). MR mediation analysis uncovered indirect causal effect of serum VD level on the risk of PCOS via testosterone (OR=0.983, 95% CI: 0.968–0.998, $P=0.025$). Functional enrichment analysis showed that several pathways may be involved in the VD-testosterone-PCOS axis, such as steroid hormone biosynthesis and autophagy process.

Conclusion Our findings suggest that genetically predicted lower serum VD level may cause a higher risk of developing PCOS, which may be mediated by increased testosterone production.

Keywords Vitamin D, Polycystic ovary syndrome, Testosterone, Mendelian randomization

Introduction

Vitamin D (VD) is an essential fat-soluble steroid hormone that is necessary for calcium-phosphate metabolism, bone homeostasis, cell differentiation, and immune system function. The prevalence of VD deficiency (VDD) in the population has gradually increased over the past few decades. VDD is associated with various diseases, including cardiovascular disease, inflammation, dyslipidemia, weight gain, and infectious diseases [1, 2]. Furthermore, mounting studies have indicated the potential link between the serum VD status and women's

*Correspondence:

Jing Li

jli23@cmu.edu.cn; lijingendocrine@126.com

¹ Department of Endocrinology and Metabolism, The Institute of Endocrinology, NHC Key Laboratory of Diagnosis and Treatment of Thyroid Diseases, The First Affiliated Hospital of China Medical University, Shenyang, Liaoning 110000, P.R. China

² Department of Endocrinology and Metabolism, Hongqi Hospital Affiliated to Mudanjiang Medical College, Mudanjiang, Heilongjiang 157011, P.R. China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

reproductive health. Firstly, the biological function of VD is mediated via intracellular VD receptors (VDRs), which are distributed among various tissues, encompassing hypothalamic, pituitary tissue, endometrium, and ovary [3, 4]. Secondly, VD participates in regulating genes associated with ovarian and placental functions [5, 6]. All evidences suggest that the serum VD plays a potentially significant role in female reproductive health.

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder that affects women of reproductive age, with a global incidence ranging 20–25% [7, 8]. PCOS will affect woman's endometrial function and oocyte competence [9, 10], which leads to reproductive dysfunction in PCOS patients, including infertility, miscarriage, and pregnancy complications [11–13]. However, the exact pathogenesis of PCOS remains unclear. Prior observational studies have elucidated the correlation between the serum VD and the risk of PCOS. A recent study revealed that serum VD concentration were lower in women diagnosed with PCOS compared to body mass index (BMI)-matched control, suggesting that regardless of BMI, PCOS is correlated with reduced VD level [14]. However, these studies can only prove that there is a correlation between them, they cannot clarify the causality between them. In addition, hyperandrogenemia stands as one of the diagnostic criteria for PCOS and impacts 60–80% of patients [15]. Female are actually more sensitive to testosterone even though it is known as a male hormone [16]. Growing evidences showed that testosterone may play an important role between the serum VD level and the risk of PCOS. Hahn et al. illustrated an association between the serum VD level and the severity of hirsutism in individuals with PCOS [17]. The research conducted by Latic et al. indicates a negative correlation between serum VD level and testosterone production in patients with PCOS [18]. However, a study by Mesinovic et al. suggested no discernible correlation between the serum VD level and androgen production in individuals with PCOS [19]. Moreover, a large observational study by Gallea et al. also showcased the association between serum VD levels, insulin, and body weight among PCOS patients but not specifically with hyperandrogenemia [20]. The reason for these different results may be due to the fact that observational studies are susceptible to confounding factors as well as various biases [21]. Therefore, it is not clear whether testosterone production mediate the relationship between serum VD level and the risk of PCOS, due to the limitations of the study methodology.

In recent years, mendelian randomization (MR) analysis is widely used as an epidemiological method in medical research. Firstly, MR analysis can minimize the impact of confounding factors and various biases on the results by simulating randomized controlled trials (RCTs) at the

genetic level, and secondly, MR analysis can also determine causality and reduce the impact of reverse causality on the results of the study [22].

Thus, in this study, we use the bidirectional two-sample MR (TS-MR) analysis to investigate the cause-effect relationship between the serum VD level and the risk of PCOS. Secondly, we perform the mediation MR analysis to test the mediating role of testosterone production between serum VD level and the risk of PCOS. Finally, we used the bioinformatics analysis to assess the possible biological functions and molecular mechanisms between them.

Materials and methods

Study design of mendelian randomization study

Our study explored the cause-effect of serum VD level as an exposure on the risk of developing PCOS as an outcome trait and the effect of testosterone as a mediator between VD and PCOS through bidirectional TS-MR analysis, multivariable MR (MVMR) and mediator MR analysis (Fig. 1). In order to ensure the study's validity, the study needed to meet the three following crucial assumptions [23] (Fig. 1C): 1) the correlation assumption: instrumental variables (IVs) must be robustly correlated with the exposure factors; 2) the exclusion restriction assumption: IVs are not associated with potential confounders of the exposure or the outcome; and 3) the independence assumption: IVs do not influence the outcome variables through other pathways besides the exposure factors. This study followed guidelines of STROBE-MR [24] checklist (Table S1).

Data source and IVs selection of mendelian randomization study

We obtained data associated with VD from a large genome-wide association study (GWAS) that identified 143 loci among 417,580 participants which was conducted by Revez et al. in 2020 [25]. We accessed the summary data related to PCOS from a meta-analysis in the FinnGen and Estonian Biobank (EstBB), which included 3609 cases and 229,788 controls [7]. Summary data related to bioavailable testosterone (BT) were obtained from the UK Biobank (UKB). Data on serum fasting glucose (FBG) levels were obtained from a UKB GWAS we conducted in 340,002 British participants [26]. Summary data on circulating concentrations of fasting insulin (FI) were obtained from the MAGIC GWAS included 151,013 participants [27]. Pooled data related to BMI were acquired from a GWAS meta-analysis within the (GIANT) consortium, encompassing 681,275 participants [28]. Details of the GWAS database are summarized in Table S2.

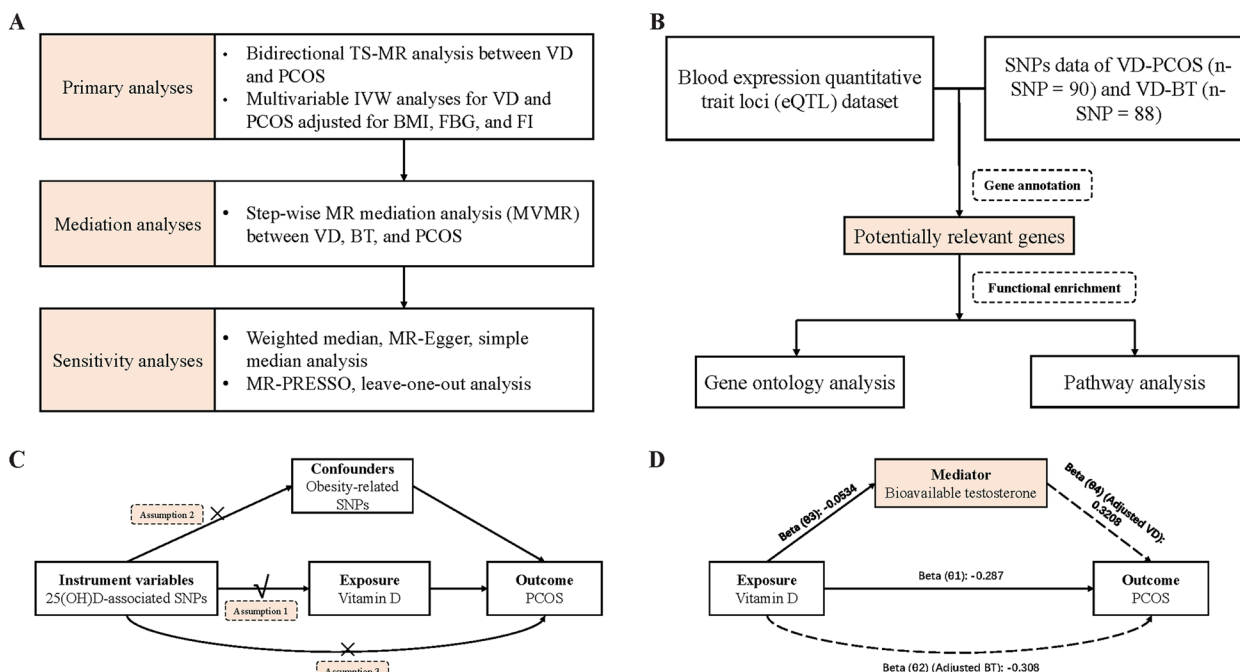


Fig. 1 Flowchart of the study. **A** Flowchart of the MR study; **B** Flowchart of the Bioinformatics study; **C** Diagram of the MR assumptions of the association between VD and PCOS; **D** Illustrative diagram for the mediation MR analysis framework Abbreviations: MR, mendelian randomization; TS-MR, two-sample MR; VD, vitamin D; PCOS, polycystic ovary syndrome; IVW, inverse variance weighted; BMI, body mass index; FBG, fasting glucose; FI, fasting insulin; MVMR, multivariable MR; BT, bioavailable testosterone; SNPs, single-nucleotide polymorphisms

In the bidirectional TS-MR analysis, Single-nucleotide polymorphisms (SNPs) with genome-wide significance ($P < 5 \times 10^{-8}$) were first selected. These SNPs were matched against the SNP-outcome GWAS database to exclude SNPs that could not be matched. To minimize the effects of linkage disequilibrium, we conducted a clumping process with an r^2 threshold of 0.001 and a clumping window of 10,000 kb and excluded these SNPs if present. Subsequently, we performed MR-PRESSO analysis immediately to demonstrate whether there was significant horizontal pleiotropy to exclude outlier SNPs [29]. To ensure that the IVs were not affected by confounding variables, we searched the PhenoScanner V2 [30] and deleted obesity-related SNPs associated with BMI and waist circumference (WC). Finally, 88 SNPs (VD on PCOS) and 2 SNPs (PCOS on VD) were used as IVs in the primary bidirectional TS-MR study, respectively. All SNPs exhibited an F statistic greater than 10. The variance explained for each SNP (R^2) was calculated using the widely-accepted formula [31, 32]. We used the same method as above to screen the SNPs required in the MR mediation analysis. All the IVs SNPs are summarized in Table S3-7.

Statistic analysis of mendelian randomization study

Initially, the primary analysis aimed to explore the causal relationship between VD and PCOS. We used bidirectional TS-MR analysis to assess the causal relationship between VD and PCOS. In this, we used Cochran’s Q test to assess the heterogeneity [33]; if there was no heterogeneity, we would use the fixed-effects inverse variance weighted (IVW) method, otherwise, we would use the random-effects IVW method [34]. Furthermore, considering that obesity, abnormal insulin levels, and abnormal glucose values are common in patients with PCOS, we adjusted genetically predicted BMI, FBG, and FI by MVMR to explore the direct causal effect between VD and PCOS. To make the results more robust.

Secondly, a stepwise MR analysis approach was used to examine whether there exist mediation effects of BT between VD and PCOS. To assess the direct causal effect between VD, BT, and PCOS, we performed an MVMR analysis using the MVMR R package [35]. Conditional F statistics were calculated for assessing the strength of the genetic instruments in MVMR analysis [36]. The product of the coefficients method [37] and the multivariate delta

method [38] were used to calculate the indirect effects of VD on PCOS via mediator.

Sensitivity analysis of mendelian randomization study

The following tests were used as sensitivity analyses to assess the robustness of MR effect estimates to invalid genetic variants. Firstly, we conducted MR-Egger regression [39, 40], weighted median [41], and weighted mode [42] methods. MR-Egger regression can detect and explain horizontal pleiotropy mainly through intercept tests [39, 40]. Weighted median can yield impartial estimations even when over half of the information arise from flawed IVs [43]. We used weighted mode to divide SNPs into multiple subsets based on similar causal effects, and the estimates of causal effects were computed for the subset with the highest number of SNPs [42]. Secondly, the leave-one-out (LOO) analysis can test whether the results are affected by a single SNP [44]. Thirdly, as described above we performed MR-PRESSO analysis [29] to identify the presence of potential horizontal pleiotropic outliers in IVs that could lead to biased results, as well as searching for and removing obesity-related SNPs associated with BMI and WC from the PhenoScanner database [45].

All analyses were conducted using R version 4.2.0 (R Foundation for Statistical Computing, Vienna, Austria). *P* values were considered significant at 0.05.

Bioinformatical analysis

We used the largest whole blood expression quantitative trait loci (eQTL) dataset from the eQTLGen consortium, which includes data on cis-eQTLs for 19,250 whole blood expressed genes from 31,684 individuals [46]. We combined SNPs data of VD-PCOS (*n*-SNP=90) and VD-BT (*n*-SNP=88) with cis-eQTLs data for gene annotation, respectively. Genes with $P < 5 \times 10^{-8}$ and *FDR* < 0.05 were screened as potentially relevant genes for VD-PCOS and VD-BT.

Subsequently, we used these potentially relevant genes for bioinformatics analyses, including Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. GO analyses [47], including biological

process (BP), molecular function (MF), and cellular composition (CC), are commonly used for large-scale functional enrichment studies. KEGG is a database that stores information about genomes, biological pathways, diseases, and drugs. We used the clusterProfiler package, org.Hs.eg.db package, and enrichplot package in the software R to perform GO and KEGG enrichment analyses of the potentially relevant genes. *P* < 0.05 for GO entries and KEGG pathways were considered significant.

Results

Causal effect between serum vitamin D and polycystic ovary syndrome

In our bidirectional TS-MR analysis, the number of IVs of VD on PCOS and PCOS on VD were 90 and 2, respectively. The F-statistic values for each SNP were greater than 10 (Table S3), indicating that the results were almost unaffected by weak instrumental bias. The result of fixed-effects IVW method (Cochran’s Q statistic = 81.42, *P* = 0.704) indicated that genetically predicted higher level of VD led to a lower risk of developing PCOS after excluding obesity-associated SNPs (*n* = 90 SNPs, OR = 0.750, 95% CI: 0.587–0.959, *P* = 0.022) (Table 1). MR-Egger, weighted median, and weighted mode methods all obtained similar magnitude and direction to IVW method (Table 1). The scatter plot demonstrates the inhibitory effect of individual SNP on PCOS (Fig. S1). Since the MR-Egger *P*-intercept was greater than 0.05 (Table S8) and the funnel plot (Fig. S2) was roughly symmetrical, there was no indication of horizontal pleiotropy detected in the study. The results of the LOO analyses indicated that there were no potentially affecting SNPs in the main MR analyses (Fig. S3). The result of the result of the MR-PRESSO test did not show any outlier SNPs. Nevertheless, the results of reverse TS-MR showed that genetically predicted risk of developing PCOS did not affect the VD level (fixed-IVW: *n* = 2 SNPs, OR = 1.004, 95% CI: 0.987–1.022, *P* = 0.640) (Table 1).

We subsequently explored the direct effect of the serum VD level on PCOS by MVMR methods, and the results of both Model 1 (adjusted BMI) and Model 2 (adjusted BMI, FBG, and FI) showed that the negative correlation

Table 1 Bidirectional two-sample Mendelian randomization results of the serum VD level and the risk of PCOS

Exposure	Outcome	SNPs, n	Method	OR (95%CI)	β	<i>P</i>
VD	PCOS	90	IVW (fixed)	0.750 (0.587-0.959)	-0.287	0.022
		90	Weighted median	0.833 (0.564-1.231)	-0.182	0.359
		90	MR Egger	0.846 (0.570-1.255)	-0.168	0.407
		90	Simple median	0.748 (0.499-1.121)	-0.290	0.160
PCOS	VD	2	IVW (fixed)	1.004 (0.987-1.022)	0.004	0.640

Abbreviations: VD Vitamin D, PCOS Polycystic ovary syndrome, SNPs Single-nucleotide polymorphisms, IVW Inverse variance weighted, NA Not available, OR Odds ratio

Table 2 Multivariate mendelian randomization results of the serum VD level on the development of PCOS risk

Method	OR (95%CI)	β	P
IVW (no adjusted)	0.750 (0.587-0.959)	-0.287	0.022
Model 1 (adjusted BMI)	0.704 (0.521-0.952)	-0.351	0.023
Model 2 (adjusted BMI, FBG, and FI)	0.694 (0.510-0.945)	-0.365	0.020

Abbreviations: VD Vitamin D, PCOS Polycystic ovary syndrome, BMI Body mass index, FBG Fasting glucose, FI Fasting insulin, OR Odds ratio, IVW Inverse variance weighted

Table 3 Step-wise MR mediation analysis (MVMR) results between the serum VD level, BT, and the risk of PCOS

	β	SE	OR (95%CI)	P
θ_1	-0.2871	0.1253	0.750 (0.587-0.959)	0.022
θ_2	-0.3082	0.1461	0.735 (0.552-0.978)	0.035
θ_3	-0.0534	0.0162	0.948 (0.905-0.994)	0.026
θ_4	0.3208	0.1044	1.378 (1.123-1.691)	0.002
Indirect effect ($\theta_3 \times \theta_4$)	-0.0171	0.0076	0.983 (0.968-0.998)	0.025

Indirect effect ($\theta_3 \times \theta_4$): indirect causal effect of the serum VD level on the risk of PCOS via BT

Abbreviations: VD Vitamin D, BT Bioavailable testosterone, PCOS Polycystic ovary syndrome, MR Mendelian randomization, MVMR Multivariable MR, OR Odds ratio, SE Standard error

θ_1 : total effect of the serum VD level on the risk of PCOS

θ_2 : direct effect of the serum VD level on the risk of PCOS

θ_3 : direct effect of the serum VD level on BT

θ_4 : direct effect of BT on the risk of PCOS

between serum VD level and the risk of PCOS remained similar (Table 2). This confirms the robustness of the TS-MR results.

Mendelian randomization mediation analysis

After excluding the outlier SNPs and obesity-related SNPs, MVMR analysis (adjusted BT) revealed direct causal effects of serum VD level (OR: 0.735, 95% CI: 0.552–0.978; $P=0.035$) on the risk of developing PCOS (Table 3, Fig. 1D). In the following steps of the MR mediation analysis, we found strong evidence for a causal effect of serum VD level ($\beta: -0.053, P=0.026$) on BT (Table 3). In addition to this, we also found a causal relationship between BT and PCOS (OR: 1.378, 95% CI: 1.123–1.691; $P=0.002$) (Table 3).

Taken together, we found the potential mediation pathways between VD and PCOS: an indirect causal effect of VD on PCOS risk via BT ($\theta_3 \times \theta_4$) (OR: 0.983, 95% CI: 0.968–0.998; $P=0.025$) (Table 3). The pathway mediated 5.96% of the total causal effect of VD on PCOS risk. Detailed estimates of direct and indirect causal effects can be found in Table 3.

Bioinformatics study

The results of the MR study suggested that reduced VD level may lead to the development of PCOS, and BT is a mediator between VD and PCOS, meaning that VD can ultimately influence the development of PCOS by affecting the production of testosterone. On the basis of the above studies, we collected IVs of VD-PCOS (n -SNPs=90) and VD-BT (n -SNPs=88) respectively, and combined these IVs with cis-eQTLs data for gene annotation respectively. Ultimately, 147 (VD-PCOS) and 164 (VD-BT) potentially relevant genes were annotated (Table S9-10), respectively. We then used these genes to perform GO and KEGG analyses.

Firstly, the potentially relevant genes of VD-PCOS were analyzed for enrichment. The results of GO analysis suggested that these genes were mainly related to androgen metabolic process, superoxide metabolic process, cell body membrane, and steroid dehydrogenase activity (Fig. 2A). The KEGG analysis was mainly enriched in the process of autophagy, steroid biosynthesis, cytochrome P450 metabolic process, and vitamin digestion and absorption process (Fig. 2C). Subsequently, potentially relevant genes associated with VD-BT were analyzed for enrichment. The results of GO analysis suggested that these genes were mainly associated with steroid metabolism, superoxide metabolism, autophagosome membrane, nuclear androgen receptor binding, and vitamin transmembrane transporter activity (Fig. 2B), and the KEGG analysis was mainly enriched for autophagy, steroid biosynthesis, vitamin digestion and absorption, and cholesterol metabolism process (Fig. 2C). All information of the enrichment analysis is shown in the additional file (Table S11-S12).

Discussion

In our bidirectional TS-MR analysis, we found that higher serum VD level was causally associated with a lower risk of developing PCOS (OR=0.750, 95% CI: 0.587–0.959, $P=0.022$), whereas there was little evidence for a causal effect of the risk of PCOS on the effect of serum VD level. Furthermore, our MR mediation analysis confirmed that testosterone can act as one of the mediating factors between the causality of VD and PCOS (OR=0.983, 95% CI: 0.968–0.998, $P=0.025$). The mediating effect of testosterone was 5.96%. Ultimately, we utilized potentially relevant genes for GO and KEGG enrichment analysis to assess the involvement of testosterone and the potential biological and molecular mechanisms between them.

VD, a lipid-soluble vitamin, plays a pivotal role in numerous biological processes. Primarily synthesized endogenously through exposure to sunlight, it is also acquired, albeit to a lesser extent, from dietary sources

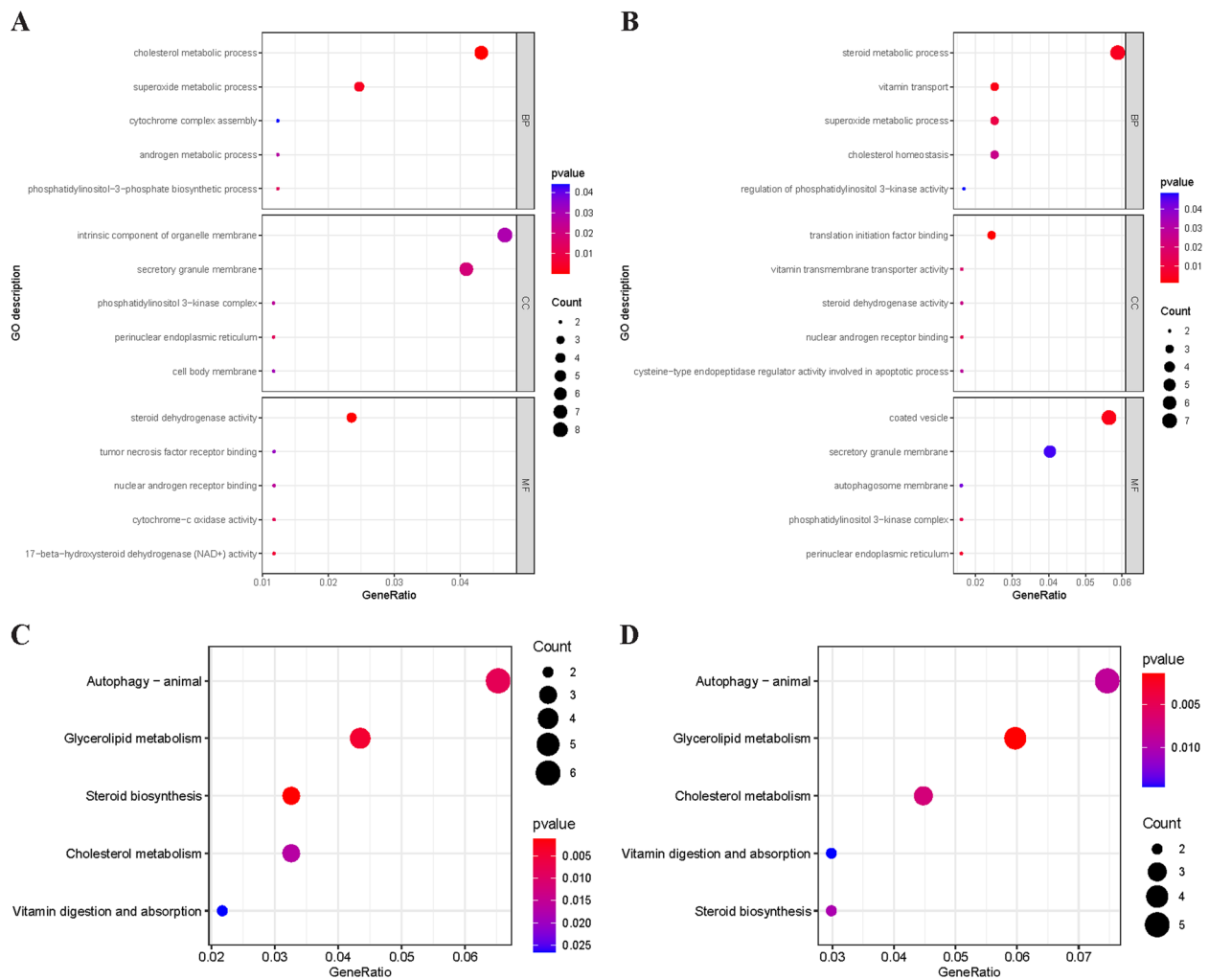


Fig. 2 Gene Ontology and Kyoto Encyclopedia of the Genome pathway enrichment analysis of potentially relevant genes. **A** The GO enrichment analysis for potentially relevant genes related to VD and PCOS; **B** The GO enrichment analysis for potentially relevant genes related to VD and BT; **C**. The KEGG pathway analysis for potentially relevant genes related to VD and PCOS; **D**. The KEGG pathway analysis for potentially relevant genes related to VD and BT. Abbreviations: VD, vitamin D; PCOS, polycystic ovary syndrome; BT, bioavailable testosterone; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of the Genome

[48]. VDD is considered a globally prevalent nutritional deficiency, with various studies reporting prevalence rates of 58–91% among infertile women [49]. A cross-sectional study encompassing 625 women diagnosed with PCOS and 217 control subjects revealed that Chinese women diagnosed with PCOS exhibited notably lower level of VD compared to their healthy [50]. The result from a large observational study conducted by Krul-Poel et al. similarly demonstrated significantly diminished level of VD among women within the PCOS group [51]. Recent research has demonstrated that women with PCOS exhibit lower serum concentrations of VD compared to BMI-matched controls. This implies that the level of VD is linked to PCOS irrespective of BMI [14]. Aligned with

the outcomes of these observational studies, our research indicated that higher serum VD level serves as a protective factor for the risk of PCOS. To eliminate the influence of obesity as a potential confounder on the results, we excluded obesity-related SNPs in our TS-MR analysis. Subsequently, in our MVMR analyses, we adjusted for genetically predicted BMI, FBG, and FI to explore the direct causal relationship between VD and PCOS. These stringent measures significantly enhance the credibility and robustness of our findings.

The precise mechanism through which serum VD operates on PCOS remains elusive. Hyperandrogenemia stands as a pivotal diagnostic criterion for PCOS. Numerous past studies have concentrated on exploring

the correlation between serum VD and hyperandrogenemia in PCOS, yet the conclusions drawn from these studies have not reached a consensus. A study conducted by Latic N et al. revealed a negative correlation between serum VD level and testosterone in PCOS patients. Additionally, Menichini et al. demonstrated a positive impact of VD supplementation (4000 IU) on total testosterone [52]. However, a study by Mesinovic et al. suggested no discernible correlation between serum VD and androgens in individuals with PCOS [19]. Moreover, a large observational study by Gallea et al. also showcased associations between serum VD level, insulin, and body weight among PCOS patients but not specifically with hyperandrogenemia [20]. The inconsistencies observed in these findings might stem from variations in race, sample sizes, seasonal disparities, and the lifestyles of the included subjects. Our study, employing Mendelian randomization, effectively mitigated the impact of sample size, seasonal fluctuations, and diverse lifestyles on the outcomes. Furthermore, our research focused solely on individuals of European ethnicity, and we excluded BMI-related SNPs when incorporating instrumental variables, thereby significantly reducing BMI's potential confounding effect on the results. These measures ensured the robustness and reliability of our findings. Our results suggest that testosterone acts as a mediator between serum VD and PCOS, implying that serum VD may potentially contribute to the development of PCOS by influencing testosterone production.

The mechanism by which serum VD ultimately contributes to the development of PCOS by affecting testosterone remains unclear, but possible explanation has been proposed. Serum VD heightens the activity of aromatase within the ovary, thereby fostering the conversion of androgens to estrogens, ultimately culminating in diminished androgens production [53]. Kinuta et al. demonstrated a marked reduction in aromatase activity within the ovaries of VDR knockout mice in contrast to the control group [54]. In addition, we performed bioinformatics analysis to explore more possible biological mechanisms. Firstly, the results of GO and KEGG analyses of potentially related genes of VD-PCOS showed that steroid biosynthetic process, androgen metabolic process, and nuclear androgen receptor binding process were the possible biological mechanisms between the causality of the serum VD level and PCOS. These results are consistent with the results of our bidirectional TS-MR analysis, demonstrating again that the serum VD can ultimately influence the development of PCOS by modulating testosterone production. Subsequently, we subjected potentially relevant genes associated with VD-BT to bioinformatics analysis. The results suggested

that autophagy process and superoxide metabolism process might be the biological mechanism between serum VD and testosterone.

There are very few studies linking autophagy to PCOS, and the results of these studies suggest that the development of PCOS is closely related to the process of autophagy [55]. Texada et al. showed that autophagy can regulate steroid production by modulating cholesterol transport in endocrine cells [56]. In addition to this, the role of VD-mediated autophagy in disease has been extensively studied, and basic study by Hu et al. showed that VD can mediate the regulation of autophagy function through gastric epithelial cell VD receptors, which ultimately affects the pathogenic effects of *H. pylori* [57]. However, whether VD can mediate autophagy ultimately leading to PCOS remains unknown. The results of the bioinformatics study in this study suggest that autophagy is most likely one of the important mechanisms underlying the relationship between VD and PCOS.

Our study has proved that lower serum VD level causes higher prevalence of PCOS. The latter could have oocyte competence and endometrial function impaired [9, 10], but also cause a few adverse outcomes related to reproduction, such as infertility, miscarriage, and premature delivery [12, 13]. It has been found that VDD could decrease the rates of ovulation and success pregnancy in the PCOS patients, leading to less live birth [58]. In addition, It has been reported that serum VD level was independent predicting factor for live birth in the PCOS patients received ovulation induction [59]. Yasmine et al. have reported that endometrial thickness of PCOS patients maybe improved after VD administration [60]. A recent meta-analysis has shown that VD supplementation to PCOS women could decrease the occurrence rates of early miscarriage and premature delivery [53]. The nuclear receptor of VD (VDR) and 1,25(OH)₂D₃ membrane binding protein are expressed in both ovarian granulosa and theca cells [61, 62]. It has been found that VD can regulate the expression of enzymes in the VDR and ovary, ultimately regulating ovarian function [63]. One study showed that VDR mRNA was significantly less expressed in granulosa cells of the women with PCOS [64]. It may cause PCOS patients to be more sensitive to VDD. Based on the above studies and ours, serum VD level need be monitored in the female population, especially in the women of reproductive age, and timely VD administration in PCOS patients would help to improve their reproductive function and pregnancy outcomes.

Our research has several advantages. Primarily, this study confirms the direct causal relationship of the serum VD level on the risk of PCOS through the utilization of

the TS-MR analysis method. This method avoids the limitation commonly found in most observational studies, thereby fortifying the reliability and validity of our finding. Secondly, we ascertain the mediating function of testosterone in the relationship between serum VD and PCOS via MR mediation analysis, thus laying the groundwork for subsequent mechanistic studies. Finally, this is the first study to combine MR studies and bioinformatics analyses together to explore causal relationship and potential functional mechanisms between serum VD level, testosterone, and the risk of PCOS, which is quite different from other studies. Nonetheless, this study also has limitations. Firstly, our study failed to capture dietary and sun exposure information that may affect serum VD level. Secondly, the use of exclusively European data in a MR analysis may not be generalizable to other ethnic populations, albeit reducing the impact of ethnicity bias on the study outcomes. Finally, the absence of relevant data prevented us from independently exploring the relationship of serum VD₂/D₃ with the risk of PCOS, warranting further investigation.

Conclusions

In conclusion, our studies confirm the causality between lower serum VD level and higher risk of PCOS. Furthermore, testosterone may act as a mediator between serum VD and PCOS. These findings emphasize the clinical importance of testing serum VD level and timely VD supplementation as possible primary prevention and treatment of PCOS.

Abbreviations

VD	Vitamin D
PCOS	Polycystic ovary syndrome
GWAS	Genome-wide association studies
TS-MR	Two-sample Mendelian randomization
Cis-eQTLs	Cis-expression quantitative loci
VDD	VD deficiency
VDRs	VD receptors
BMI	Body mass index
MR	Mendelian randomization
MVMR	Multivariable MR
IVs	Instrumental variables
BT	Bioavailable testosterone
UKB	UK Biobank
FBG	Fasting glucose
FI	Fasting insulin
SNPs	Single-nucleotide polymorphisms
WC	Waist circumference
IWV	Inverse variance weighted
LOO	Leave one out
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
BP	Biological process
MF	Molecular function
CC	Cellular composition

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13048-024-01420-5>.

Additional file 1: Table S1. STROBE-MR Checklist; **Table S2.** Key characteristics of participating studies; **Table S3.** GWAS significant SNPs used as genetic instruments for VD level on PCOS; **Table S4.** GWAS significant SNPs used as genetic instruments for PCOS on VD level; **Table S5.** GWAS significant SNPs used as genetic instruments for VD level on BT; **Table S6.** GWAS significant SNPs used as genetic instruments for BT on PCOS; **Table S7.** GWAS significant SNPs used as genetic instruments for BT and VD level on PCOS; **Table S8.** Heterogeneity and directional pleiotropy test using MR-Egger intercepts; **Table S9.** Potentially relevant genes corresponding to IVs associated with VD and PCOS; **Table S10.** Potentially relevant genes corresponding to IVs associated with VD and PCOS; **Table S11.** GO and KEGG enrichment analysis for potentially relevant genes related to VD and PCOS; **Table S12.** GO and KEGG enrichment analysis for potentially relevant genes related to VD and BT; **Figure S1.** Scatter plot of the MR estimates for the association of VD level with PCOS; **Figure S2.** Funnel plot reveals overall heterogeneity of the impact of VD on PCOS; **Figure S3.** Leave-one-out analysis of the impact of the VD on PCOS.

Acknowledgements

We would like to express our sincere gratitude to the compilers of the GWAS summary dataset for their management of the data collection and data resources.

Authors' contributions

Designed the study: Jing Li, Bingrui Gao; Collected data: Bingrui Gao, Chenxi Zhang; Performed statistical analyses: Bingrui Gao, Deping Wang, Bojuan Li; Drafted the manuscript: Bingrui Gao; Supervised the study and reviewed the manuscript: Jing Li, Zhongyan Shan, Weiping Teng.

Funding

This work was supported by the General Program of National Natural Science Foundation of China (grant number No.81771741), Distinguished Professor at Educational Department of Liaoning Province (grant number No. [2014]187) to JL.

Availability of data and materials

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Our analysis used publicly available genome-wide association study (GWAS) summary statistics. No new data were collected, and no new ethical approval was required.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 23 February 2024 Accepted: 20 April 2024

Published online: 07 May 2024

References

- Holick MF. The vitamin D deficiency pandemic: Approaches for diagnosis, treatment and prevention. *Rev Endocr Metab Disord.* 2017;18:153–65.
- Autier P, Boniol M, Pizot C, Mullie P. Vitamin D status and ill health: a systematic review. *Lancet Diabetes Endocrinol.* 2014;2:76–89.

3. Lerchbaum E, Obermayer-Pietsch B. Vitamin D and fertility: a systematic review. *Eur J Endocrinol*. 2012;166:765–78.
4. Irani M, Merhi Z. Role of vitamin D in ovarian physiology and its implication in reproduction: a systematic review. *Fertil Steril*. 2014;102:460–468. e3.
5. Parikh G, Varadinova M, Suwandhi P, Araki T, Rosenwaks Z, Poretsky L, et al. Vitamin D regulates steroidogenesis and insulin-like growth factor binding protein-1 (IGFBP-1) production in human ovarian cells. *Horm Metab Res*. 2010;42:754–7.
6. Du H, Daftary GS, Lalwani SI, Taylor HS. Direct regulation of HOXA10 by 1,25-(OH)₂D₃ in human myelomonocytic cells and human endometrial stromal cells. *Mol Endocrinol*. 2005;19:2222–33.
7. Tyrmi JS, Arffman RK, Pujol-Gualdo N, Kurra V, Morin-Papunen L, Sliz E, et al. Leveraging Northern European population history: novel low-frequency variants for polycystic ovary syndrome. *Hum Reprod*. 2022;37:352–65.
8. Bruni V, Capozzi A, Lello S. The Role of Genetics, Epigenetics and Lifestyle in Polycystic Ovary Syndrome Development: the State of the Art. *Reprod Sci*. 2022;29:668–79.
9. Palomba S, Daolio J, La Sala GB. Oocyte Competence in Women with Polycystic Ovary Syndrome. *Trends Endocrinol Metab*. 2017;28:186–98.
10. Palomba S, Piltonen TT, Giudice LC. Endometrial function in women with polycystic ovary syndrome: a comprehensive review. *Hum Reprod Update*. 2021;27:584–618.
11. Norman RJ, Dewailly D, Legro RS, Hickey TE. Polycystic ovary syndrome. *Lancet*. 2007;370:685–97.
12. Palomba S. Is fertility reduced in ovulatory women with polycystic ovary syndrome? An opinion paper *Hum Reprod*. 2021;36:2421–8.
13. Palomba S, De Wilde MA, Falbo A, Koster MPH, La Sala GB, Fauser BCJM. Pregnancy complications in women with polycystic ovary syndrome. *Hum Reprod Update*. 2015;21:575–92.
14. Bacopoulou F, Kolias E, Efthymiou V, Antonopoulos CN, Charmandari E. Vitamin D predictors in polycystic ovary syndrome: a meta-analysis. *Eur J Clin Invest*. 2017;47:746–55.
15. Lejman-Larysz K, Golar A, Baranowska M, Kozłowski M, Guzik P, Szydłowska I, et al. Influence of Vitamin D on the Incidence of Metabolic Syndrome and Hormonal Balance in Patients with Polycystic Ovary Syndrome. *Nutrients*. 2023;15:2952.
16. Durdiakova J, Ostatnikova D, Celec P. Testosterone and its metabolites—modulators of brain functions. *Acta Neurobiol Exp (Warsz)*. 2011;71:434–54.
17. Hahn S, Haselhorst U, Tan S, Quadbeck B, Schmidt M, Roesler S, et al. Low serum 25-hydroxyvitamin D concentrations are associated with insulin resistance and obesity in women with polycystic ovary syndrome. *Exp Clin Endocrinol Diabetes*. 2006;114:577–83.
18. Latic N, Erben RG. Vitamin D and Cardiovascular Disease, with Emphasis on Hypertension, Atherosclerosis, and Heart Failure. *Int J Mol Sci*. 2020;21:6483.
19. Mesinovic J, Teede HJ, Shorakae S, Lambert GW, Lambert EA, Naderpoor N, et al. The Relationship between Vitamin D Metabolites and Androgens in Women with Polycystic Ovary Syndrome. *Nutrients*. 2020;12:1219.
20. Gallea M, Granzotto M, Azzolini S, Faggian D, Mozzanega B, Vettor R, et al. Insulin and body weight but not hyperandrogenism seem involved in seasonal serum 25-OH-vitamin D₃ levels in subjects affected by PCOS. *Gynecol Endocrinol*. 2014;30:739–45.
21. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. Cardon L, editor. *PLoS Med*. 2007;4:e352.
22. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23:R89–98.
23. EPIC- InterAct Consortium, Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. 2015;30:543–52.
24. Skrivanekova 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.
25. Revez JA, Lin T, Qiao Z, Xue A, Holtz Y, Zhu Z, et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. *Nat Commun*. 2020;11:1647.
26. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, et al. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet*. 2021;53:1097–103.
27. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The trans-ancestral genomic architecture of glycemic traits. *Nat Genet*. 2021;53:840–60.
28. Yengo L, Sidorenko J, Kempner 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.
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.
30. Mortada I. Hyperuricemia, Type 2 Diabetes Mellitus, and Hypertension: an Emerging Association. *Curr Hypertens Rep*. 2017;19:69.
31. Choi HK, McCormick N, Lu N, Rai SK, Yokose C, Zhang Y. Population Impact Attributable to Modifiable Risk Factors for Hyperuricemia. *Arthritis Rheumatol*. 2020;72:157–65.
32. Nakamura K, Sakurai M, Miura K, Morikawa Y, Yoshita K, Ishizaki M, et al. Alcohol intake and the risk of hyperuricaemia: a 6-year prospective study in Japanese men. *Nutr Metab Cardiovasc Dis*. 2012;22:989–96.
33. Greco MFD, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med*. 2015;34:2926–40.
34. 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 genome. *eLife*. 2018;7:e34408.
35. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med*. 2021;40:5434–52.
36. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*. 2019;48:713–27.
37. VanderWeele TJ. *Mediation Analysis: A Practitioner's Guide*. *Annu Rev Public Health*. 2016;37:17–32.
38. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol*. 2007;58:593–614.
39. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512–25.
40. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32:377–89.
41. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40:304–14.
42. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46:1985–98.
43. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator - PubMed. Available from: <https://pubmed.ncbi.nlm.nih.gov/27061298/>. [cited 2023 Feb 15].
44. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology*. 2017;28:30–42.
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;35:4851–3.
46. Vösa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet*. 2021;53:1300–10.
47. The Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res*. 2015;43:D1049–56.
48. Yuan C, Qian ZR, Babic A, Morales-Oyarvide V, Rubinson DA, Kraft P, et al. Prediagnostic Plasma 25-Hydroxyvitamin D and Pancreatic Cancer Survival. *J Clin Oncol*. 2016;34:2899–905.

49. Cunningham TK, Allgar V, Dargham SR, Kilpatrick E, Sathyapalan T, Maguiness S, et al. Association of Vitamin D Metabolites With Embryo Development and Fertilization in Women With and Without PCOS Undergoing Subfertility Treatment. *Front Endocrinol.* 2019;10:13.
50. Shan C, Zhu Y, Yu J, Zhang Y, Wang Y, Lu N, et al. Low Serum 25-Hydroxyvitamin D Levels Are Associated With Hyperandrogenemia in Polycystic Ovary Syndrome: A Cross-Sectional Study. *Front Endocrinol.* 2022;13:894935.
51. Krul-Poel YHM, Koenders PP, Steegers-Theunissen RP, Ten Boekel E, Wee MMT, Louwers Y, et al. Vitamin D and metabolic disturbances in polycystic ovary syndrome (PCOS): a cross-sectional study. Narayanan R, editor. *PLOS One.* 2018;13:e0204748.
52. Menichini D, Facchinetti F. Effects of vitamin D supplementation in women with polycystic ovary syndrome: a review. *Gynecol Endocrinol.* 2020;36:1–5.
53. Yang M, Shen X, Lu D, Peng J, Zhou S, Xu L, et al. Effects of vitamin D supplementation on ovulation and pregnancy in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Front Endocrinol.* 2023;14:1148556.
54. Kinuta K, Tanaka H, Moriwake T, Aya K, Kato S, Seino Y. Vitamin D is an important factor in estrogen biosynthesis of both female and male gonads. *Endocrinology.* 2000;141:1317–24.
55. Kumariya S, Ubba V, Jha RK, Gayen JR. Autophagy in ovary and polycystic ovary syndrome: role, dispute and future perspective. *Autophagy.* 2021;17:2706–33.
56. Texada MJ, Malita A, Rewitz K. Autophagy regulates steroid production by mediating cholesterol trafficking in endocrine cells. *Autophagy.* 2019;15:1478–80.
57. Hu W, Zhang L, Li MX, Shen J, Liu XD, Xiao ZG, et al. Vitamin D3 activates the autolysosomal degradation function against *Helicobacter pylori* through the PDIA3 receptor in gastric epithelial cells. *Autophagy.* 2019;15:707–25.
58. Butts SF, Seifer DB, Koelper N, Senapati S, Sammel MD, Hoofnagle AN, et al. Vitamin D Deficiency Is Associated With Poor Ovarian Stimulation Outcome in PCOS but Not Unexplained Infertility. *J Clin Endocrinol Metab.* 2019;104:369–78.
59. Pal L, Zhang H, Williams J, Santoro NF, Diamond MP, Schlaff WD, et al. Vitamin D Status Relates to Reproductive Outcome in Women With Polycystic Ovary Syndrome: Secondary Analysis of a Multicenter Randomized Controlled Trial. *J Clin Endocrinol Metab.* 2016;101:3027–35.
60. Abuzeid Y. Impact of Vitamin D Deficiency on Reproductive Outcome in infertile anovulatory women with polycystic ovary syndrome: a systematic literature review. *Curr Dev Nutr.* 2020;4:nzaa067_001.
61. Li S, Qi J, Sun Y, Gao X, Ma J, Zhao S. An integrated RNA-Seq and network study reveals that valproate inhibited progesterone production in human granulosa cells. *J Steroid Biochem Mol Biol.* 2021;214:105991.
62. Hrabia A, Kamińska K, Socha M, Grzesiak M. Vitamin D3 Receptors and Metabolic Enzymes in Hen Reproductive Tissues. *Int J Mol Sci.* 2023;24:17074.
63. Xu J, Lawson MS, Xu F, Du Y, Tkachenko OY, Bishop CV, et al. Vitamin D3 Regulates Follicular Development and Intrafollicular Vitamin D Biosynthesis and Signaling in the Primate Ovary. *Front Physiol.* 2018;9:1600.
64. Aghadavod E, Mollaei H, Nouri M, Hamishehkar H. Evaluation of Relationship between Body Mass Index with Vitamin D Receptor Gene Expression and Vitamin D Levels of Follicular Fluid in Overweight Patients with Polycystic Ovary Syndrome. *Int J Fertil Steril.* 2017;11:105–11.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.