Check for updates

OPEN ACCESS

EDITED BY Eytan R Barnea, BioIncept, LLC, United States

REVIEWED BY Vera Tocci, Magna Græcia University, Italy Rawaa A. Faris, University of Baghdad, Iraq

*CORRESPONDENCE Yutong Li liyutong@gdmu.edu.cn Shan Jiang 578588030@qq.com

RECEIVED 26 November 2024 ACCEPTED 03 April 2025 PUBLISHED 01 May 2025

CITATION

Li Y, Lin X, Zou K, Du J, Li Q, Zhong L and Jiang S (2025) Blood biochemical landscape and new insights into clinical decision-making for polycystic ovary syndrome in Chinese women: a prospective cohort study. *Front. Endocrinol.* 16:1534733. doi: 10.3389/fendo.2025.1534733

COPYRIGHT

© 2025 Li, Lin, Zou, Du, Li, Zhong and Jiang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Blood biochemical landscape and new insights into clinical decision-making for polycystic ovary syndrome in Chinese women: a prospective cohort study

Yutong Li^{1,2*}, Xiufeng Lin³, Ke Zou¹, Jing Du³, Qingni Li³, Linkun Zhong² and Shan Jiang^{3*}

¹The First Clinical College, Guangdong Medical University, Zhanjiang, Guangdong, China, ²Department of General Surgery, Zhongshan City People's Hospital, Zhongshan, Guangdong, China, ³Reproductive Center, Boai Hospital of Zhongshan, Zhongshan, Guangdong, China

Introduction: The Polycystic ovary syndrome (PCOS), a prevalent endocrine disorder affecting women's reproductive and metabolic health, faces diagnostic challenges due to heterogeneous clinical presentations and the absence of reliable biomarkers. This study investigates the role of Glucosaminyl (N-acetyl) transferase 2 (GCNT2) in modulating sex hormone-binding globulin (SHBG) and its potential as a therapeutic target in PCOS pathophysiology.

Methods: A prospective cohort of 103 PCOS patients treated with oral contraceptives (2021–2024) was established. Bidirectional Mendelian randomization (MR) was employed to assess genetic associations and causal relationships between PCOS and SHBG. Molecular docking studies evaluated cryptotanshinone's binding affinity to key proteins (COL1A1, COL4A2, COL6A2) in the PI3K/Akt pathway. GCNT2's regulatory effects on collagen synthesis and extracellular matrix pathways. Pharmacokinetic profiling validated therapeutic viability.

Results: Bidirectional MR revealed significant genetic associations (P < 0.001) and causal links between PCOS and SHBG, implicating GCNT2 as a key modulator. Cryptotanshinone exhibited strong binding affinity to PI3K/Akt signaling pathway proteins and favorable pharmacokinetic properties. Enrichment analyses highlighted GCNT2's role in collagen biosynthesis (*FDR* < 0.05) and extracellular matrix regulation.

Discussion: This study identifies GCNT2 as a critical mediator of PCOS pathophysiology through SHBG modulation and collagen remodeling.

Cryptotanshinone emerges as a promising therapeutic candidate, targeting PI3K/ Akt signaling pathway with high specificity. These findings advance the understanding of PCOS mechanisms and provide a foundation for biomarkerdriven diagnostics and precision therapeutics. Further validation in clinical trials is warranted to translate these insights into practice.

KEYWORDS

cryptotanshinone, glucosaminyl (N-acetyl) transferase 2, SHBG, polycystic ovary syndrome, Mendelian randomization (MR)

1 Introduction

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder affecting women of reproductive age globally, with a prevalence of up to 33% (1, 2). According to recent studies, there are approximately 1.55 million cases of PCOS among women worldwide (3). The characteristics of PCOS include hyperandrogenism, hyperinsulinemia, and disrupted adipokine secretion, with patients often experiencing β -cell dysfunction (4). Low birth weight and fetal exposure to androgens may contribute to the development of PCOS phenotypes, predisposing individuals to metabolic complications. PCOS poses the multifaceted challenge for patients and healthcare providers, with emerging evidence indicating that its metabolic impact extends beyond reproductive symptoms. PCOS is closely associated with insulin resistance, affecting nearly 70% of patients, which contributes to hyperinsulinemia and widespread metabolic dysregulation. These disturbances markedly increase the risk of type 2 diabetes, cardiovascular diseases, and metabolic syndrome in affected women, highlighting the urgent need for targeted and effective therapeutic strategies (5, 6). Consequently, PCOS presents multifaceted challenges for both patients and healthcare providers (7, 8). As research advances in unraveling the molecular mechanisms of PCOS, it is evident that a comprehensive understanding of its pathophysiology is crucial for designing precise and effective therapeutic interventions.

Glucosaminyl (N-acetyl) transferase 2 (GCNT2) is an essential glycosyltransferase enzyme that mediates the biosynthesis of complex carbohydrates, particularly through the modification of glycoproteins and glycolipids. It catalyzes the transfer of Nacetylglucosamine residues to specific substrates, thereby modulating the structural and functional characteristics of glycoproteins, which are integral to processes such as cellular signaling, adhesion, and immune regulation (9). The glycosylation patterns mediated by GCNT2 are essential for protein functionality, influencing their stability, cellular localization, and molecular interactions. Notably, GCNT2 has been identified as a key regulator of sex hormone-binding globulin (SHBG), a glycoprotein that modulates the bioavailability of sex hormones, including testosterone and estradiol, thereby playing a critical role in hormonal homeostasis (8, 10). Dysregulation of SHBG is associated with several endocrine disorders, including PCOS, which is characterized by hormonal imbalances and metabolic dysfunction (2, 11, 12). Studies have indicated that alterations in GCNT2 expression can lead to changes in SHBG levels, further implicating GCNT2 in the pathophysiology of PCOS and its associated symptoms, such as insulin resistance and infertility (13). Furthermore, GCNT2 has been shown to play a role in the PI3K/Akt pathway, which is crucial for cell survival, proliferation, and metabolism (14, 15). Investigating the specific mechanisms through which GCNT2 influences SHBG levels and other key pathways could provide new insights into the development of targeted therapies for PCOS and related conditions.

This study is designed to address three key objectives: (1) the identification of novel blood-based diagnostic biomarkers for PCOS through comprehensive analysis of clinical blood biochemical indicators, with a specific focus on the regulatory role of GCNT2, and the evaluation of their diagnostic precision and clinical applicability, (2) the investigation of the effects of estradiolcyproterone acetate (EE-CA) oral contraceptives on serum SHBG levels in PCOS patients, with an emphasis on potential interactions with GCNT2 expression, and (3) the discovery of candidate genes, particularly GCNT2, for early detection and targeted therapeutic interventions in PCOS. To this end, we have established a welldefined training cohort comprising 103 PCOS patients, systematically analyzing their blood biochemical profiles and employing advanced machine learning techniques to assess the specificity and predictive capacity of these indicators for PCOS diagnosis. Bidirectional two-sample Mendelian randomization (MR) will be utilized to explore risk associations of target proteins, including GCNT2, based on identified blood biochemical markers, aiming to uncover potential therapeutic targets and elucidate their molecular mechanisms. Additionally, we will leverage AI-Driven Drug Design and computational biology approaches to develop low-toxicity, high-efficacy lead compound targeting GCNT2 and associated proteins, thereby paving the way for personalized and precision-based treatment strategies for PCOS patients.

2 Methods

2.1 Materials and sources of clinical samples

Blood samples were collected from 103 PCOS patients receiving treatment with Diane-35 at Zhongshan Boai Hospital (Third Class, Category A). Diane-35, produced by Bayer Weimar GmbH & Co. KG (Bayer, National Drug Approval Number HJ20170210, Germany), was administered at a dosage of one tablet per day. Each tablet contains a combination of EE-CA and ethinylestradiol, which are essential components for managing the symptoms associated with PCOS; Insulin (INS) measurement kit (YHLO, Shenzhen, China); D-dimer measurement kit (PUSHKANG, Hunan, China); Elecsys Müller Tube (AMH) Hormone Plu (Roche Diagnostics, USA); Elecsys SHBG (Roche Diagnostics, USA); Access TSH (3rd IS) (Beckman Coulter, USA); Access hFSH (Beckman Coulter, USA); Access hLH (Beckman Coulter, USA); Access Testosterone (TESTO) (Beckman Coulter, USA); Total Cholesterol (CHOL) Assay Kit (Gcell, Beijing, China); High-Density Lipoprotein Cholesterol (HDL-C) Assay Kit (Gcell, Beijing, China); Low-Density Lipoprotein Cholesterol (LDL-C) Assay Kit (Gcell, Beijing, China); Triglycerides (TG) Assay Kit (Gcell, Beijing, China); Glucose Kit (GLU) (Medicalsystem, Zhejiang, China); Homocysteine (Hcy, HCY) Determination Reagent Kit (ERKN, Zhejiang, China).

2.2 Detection instruments

Fully automated coagulation analyzer (Sysmex Europe, CS5 100, Kobe, Japan); fully automated biochemical immunoassay system (Roche Diagnostics, COBAS[®] 8000, USA); fully automated chemiluminescent immunoassay analyzer (Beckman Coulter, UniCel DxI 800 Access, USA); fully automated biochemical analyzer (Beckman Coulter, AU5831/AU5821, USA); chemiluminescence detector (YHLO, iFlash 3000-A, Hunan, China).

2.3 Study cohort construction

The study included 103 PCOS patients who visited our hospital between March 2021 and July 2024. These patients primarily presented with oligomenorrhea or amenorrhea, dysfunctional uterine bleeding, or hyperandrogenemia. Type A PCOS patients typically exhibit more pronounced metabolic alterations and represent the most prevalent subtype (16). In this study, all 103 enrolled patients were newly diagnosed with Type A PCOS.

PCOS is diagnosed based on the Rotterdam Consensus criteria when at least two of the following three conditions are met: menstrual irregularities (oligomenorrhea or amenorrhea), hyperandrogenism (clinically or biochemically confirmed), and specific ovarian characteristics (at least 12 follicles of 2-9 mm in each ovary or ovarian volume >10 ml). This method facilitates

accurate diagnosis and informs subsequent clinical management (17). Exclusion criteria comprised individuals who had used hormonal medications, vitamins, metformin, or homocysteineelevating agents within the preceding six months, as well as those with documented deficiencies in folate, vitamin B12, or vitamin B6.

Patients' weight and height were recorded to calculate the body mass index (BMI = weight (kg)/height (m)²). A BMI of 25.1-30 kg/ m² indicated overweight, while a BMI above 30 kg/m² signified obesity. Waist and hip circumferences were measured at the narrowest and widest points, respectively. The waist-to-hip ratio (WHR) was then calculated, with a WHR > 0.72 considered abnormal. All patients took a combined oral contraceptive containing 35 µg of EE and 2 mg of CA (Diane-35, Bayer, Germany) for 21 days each month, followed by a 7-day drug-free interval, for a total of three cycles. Assessments were performed before treatment and following the completion of the third cycle, evaluating various parameters such as hormone levels (including follicle-stimulating hormone [FSH], luteinizing hormone [LH], TESTO, and free testosterone [FAI]), lipid profiles (CHOL, LDL-C, HDL-C and TG), as well as homocysteine, folate, and serum insulin levels. Blood samples were obtained during the early follicular phase (days 3-5) of spontaneous or progesteroneinduced bleeding, collected between 8:00 and 9:00 AM after an overnight fast. The samples were kept on ice, promptly centrifuged, and the serum was stored at -80°C for future analysis.

2.4 Detection of biochemical markers and metabolic indicators

INS levels were measured using a chemiluminescence detector YHLO iFlash 3000-A. D-dimer was quantified via the PUSHKANG latex immunoturbidimetric method. AMH and SHBG were assessed using the Roche COBAS[®] 8000 system, employing electrochemiluminescence. The levels of TSH, FSH, LH and TESTO were measured using the Beckman Coulter UniCel DXI 800 Access system, which utilizes chemiluminescent assays. GLU levels were determined using the Meikang Biotechnology kit, while HCY was measured via the enzyme cycle method with the ERKN kit. All analyses were performed using fully automated systems to ensure accuracy and efficiency in the laboratory. CHOL, HDL-C, LDL-C, TG and transaminases were analyzed with the Beckman Coulter AU5831/AU5821 biochemical analyzers.

2.5 Selection of diagnostic and therapeutic predictive markers for PCOS using machine learning and decision models

Statistical analysis of clinical characteristics, biochemical markers, and metabolic features of PCOS patients was performed using R 4.3.2, focusing on indicators that significantly changed following EE-CA treatment (P < 0.05). LASSO regression and random forest (RF) algorithms were employed to train a cohort of 103 PCOS patients, facilitating the selection of biomarkers for

diagnosis and treatment outcome prediction. Additionally, a decision curve analysis (DCA) was conducted to compare the predictive efficacy of these novel biomarkers against established clinical and biochemical markers [LH/FSH, TESTO, and FAI [FAI = $100 \times (TESTO/SHBG)$] (18)], thereby aiding in the identification and validation of new biomarkers for PCOS.

2.6 Data source

The overview of data sources, genetic instrument selection, and statistical analysis in this study is illustrated in Figure 1. The summary data for PCOS was obtained from a large genome-wide association study (GWAS) conducted by Tyrmi JS et al., which included 797 cases and 140,558 controls (19). The summary data for SHBG was obtained from the UK Biobank, comprising serum SHBG levels from up to 189,473 women of European ancestry (20). PCOS participants were diagnosed based on the criteria established by the National Institutes of Health, the Rotterdam criteria, or self-reported diagnoses (21).

Plasma protein quantitative trait loci (pQTL) data were obtained from a comprehensive integrative study conducted by Ferkingstad et al., which detailed pQTL information for 4,907 plasma proteins from 35,559 Icelandic participants (22). To pinpoint cis-pQTLs, we established the following criteria: (1) pQTLs must demonstrate genome-wide significance ($P < 5 \times 10^{-8}$); (2) pQTLs should be situated outside the major histocompatibility complex region; (3) independence is required, defined by linkage disequilibrium (LD) with $r^2 < 0.001$; (4) pQTLs must show ciseffects, meaning they fall within a 1000 kb range of the corresponding protein-coding sequences; (5) pQTLs should not be considered weak instruments (*F-statistic* > 100); and (6) Single nucleotide polymorphisms (SNPs) that were palindromic or had missing data were excluded. In total, we identified 6,238 cis-acting SNPs linked to 1,729 proteins. To further validate our findings, we obtained blood expression quantitative trait locus (eQTL) datasets from eQTLGen (https://www.eqtlgen.org/). This dataset includes cis-eQTL information for 16,987 genes derived from 31,684 blood samples of European ancestry (19).

The independent validation cohort for PCOS patients was sourced from the GEO database (https://www.ncbi.nlm.nih.gov/), specifically GSE87435 dataset, which includes 36 individuals diagnosed with PCOS. Molecular 3D structures of plasma proteins and natural compounds were retrieved from the UniProt (https://www.uniprot.org/), RCSB PDB (https://www.rcsb.org/), and PubChem (https://pubchem.ncbi.nlm.nih.gov/). All data utilized in this study were derived from existing publications or publicly available databases that had obtained ethical approval and informed consent, thus negating the need for additional ethical clearance.

2.7 MR analysis

Plasma proteins were treated as exposures, with SHBG and PCOS as outcomes. When a plasma protein had a single pQTL as an



instrumental variable (IV), we assessed the causal effect using the Wald ratio (WR) method (23). In instances with two or more IVs, we employed inverse variance weighting (IVW) (24). Statistical results were expressed as odds ratios (OR) with 95% confidence intervals (95% CI), using a nominal significance threshold of P < 0.05. To minimize false positives, a false discovery rate (FDR) correction was applied, defining significant findings as those with FDR < 0.05. If a protein showed a nominally significant association that lost significance after correction, it was classified as suggestively significant.

MR analyses were conducted using the "TwoSampleMR" package within R 4.3.2. To assess the robustness of our findings, we utilized Cochran's Q test and the MR-Egger intercept test to evaluate potential heterogeneity and horizontal pleiotropy after applying the IVW method (with P > 0.05 indicating the absence of these issues) (25). For the IVW analysis, both fixed-effect and random-effect models were used; the fixed-effect model applied when no significant heterogeneity was detected, while the random-effect model was used in its presence.

We employed MR-PRESSO to identify outliers, specifically potential pleiotropic SNPs, and to estimate causal effects after their exclusion (26). Leave-one-out analysis was conducted to determine if influential SNPs drove the causal effect, recalculating MR estimates by omitting one instrument at a time (27). Additionally, using the "phenoscanner" tool, we performed phenotype scanning to assess whether the identified pQTLs were associated with other traits and exhibited pleiotropic effects (28, 29). pQTLs were considered pleiotropic if they met two criteria: (1) associations that achieved genome-wide significance ($P < 5 \times 10^{-8}$); (2) known associations with established risk factors for SHBG and PCOS.

2.8 Functional enrichment analysis of target proteins and their regulation in PCOS

We constructed a receiver operating characteristic (ROC) model using a cohort of PCOS patients from the GEO database to evaluate the diagnostic efficacy of potential risk plasma proteins related to SHBG and PCOS. We examined the differential mRNA expression levels of risk plasma proteins in PCOS versus normal ovarian tissues, identifying those with the greatest contribution to PCOS risk for further analysis, including genome-wide coexpression gene analysis (logFC > 1, P < 0.05). Protein-protein interaction (PPI) networks were constructed using the STRING platform (https://string-db.org/), and chromosomal co-localization analysis was performed for the identified risk proteins. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted to elucidate the molecular mechanisms through which the target plasma proteins regulate processes in PCOS. All analyses and visualizations were carried out using R 4.3.2.

2.9 High-throughput virtual screening (HTVS) of natural compounds targeting the identified plasma proteins

Natural compounds potentially targeting the identified plasma proteins were sourced from the TCM-Suit, HERB, and TCMSID database (http://tcm-suite.aimicrobiome.cn/, http://herb.ac.cn/, https://tcm.scbdd.com/). These compounds were imported into the Discovery Studio 2019 (DS) for molecular conformation expansion. GCNT2, identified as the most promising plasma protein biomarker for PCOS, had its 3D structure imported into DS. The protein amino acid sequence was imported into SWISS-MODEL to construct the protein crystal structure in a human physiological environment (30). We performed protein residue modifications and active pocket predictions under the CHARMM36 force field in a human aqueous environment, conducting both rapid and precise molecular docking. The docking energies and intermolecular interaction energies were calculated, allowing us to select the optimal conformation for each compound. The best conformations were then visualized using PyMOL to generate 3D interaction models. Additionally, we validated the results from DS using AutoDock Vina, calculating the protein-ligand binding energies through Vina docking.

2.10 Analysis of the pharmacokinetic and toxicological properties of lead compounds

The structures of lead compounds processed in DS were imported into Swiss-ADME (http://www.swissadme.ch/), VenomPred 2.0 (31) and VirtualTaste (https://insilico-cyp.charite.de/) to evaluate their ADME, toxicokinetic (including hepatotoxicity, neurotoxicity, mutagenicity, hormonal toxicity, strong irritancy, and acute oral toxicity), and palatability analysis. Subsequently, the collected data were visualized using R 4.1.3.

2.11 The fast fourier transform algorithm supports protein-protein molecular docking

We employed a protein-protein docking technique based on the FFT algorithm, implemented through DS. This approach leverages the mathematical properties of FFT to convert the spatial coordinates of protein structures into frequency space, facilitating efficient searches for optimal binding orientations between protein complexes. Specifically, the FFT algorithm processes molecular surface interactions in the frequency domain, enabling rapid identification of potential binding sites while significantly reducing computational time and maintaining high accuracy in predicting protein interactions. In implementation process, we first constructed three-dimensional models for both the target protein and the ligand protein, converting them into a format suitable for FFT calculations. Subsequently, we discretized the molecular surfaces to construct an interaction matrix and performed convolution operations in the frequency domain to filter out the binding modes with the lowest energy from a vast array of possible conformations. Finally, the generated docking models were validated using energy scoring functions and further refined through molecular dynamics simulations.

3 Results

3.1 Analysis of the correlation between clinical and hematological indicators in patients with PCOS before and after treatment

All participating women completed the study, with each patient exhibiting clear signs of PCOS during ultrasound examinations. Among the patients, 22 were found to have folate metabolism disorders (21.4%), while 27 experienced adverse outcomes following artificial embryo transfer (IUI/IVF) (26.2%). 19 patients encountered pregnancy complications (18.4%), 13 presented with signs of hyperandrogenism (12.6%), and 10 had hyperinsulinemia (9.7%). 10 patients exhibited hypercholesterolemia (9.7%), and 9 were diagnosed with hyperhomocysteinemia (HHCY) (8.7%). Following treatment with oral contraceptives, patients exhibited significant improvements in serum levels of LH, SHBG, HCY, LH/ FSH ratio, and TESTO (P < 0.05). Notably, SHBG emerged as a promising novel biomarker for assessing the diagnosis and treatment efficacy in PCOS patients (P < 0.0001) (Table 1).

3.2 Changes in clinical and hematological indicators in PCOS patients before and after treatment with oral contraceptives

Although only 32 patients (31.1%) had a waist circumference greater than 88 cm, central obesity, as estimated by the WHR, was observed in 13 patients (40.6%). Among the participants, 35 had normal weight (34%), while 64 were classified as obese (62.1%). Throughout the treatment period, no significant changes were noted in BMI or WHR. All patients maintained normal levels of vitamin B12 and folate at both the beginning and end of the study. During treatment, TESTO levels decreased from 0.72 ± 0.26 to 0.62 \pm 0.23 ng/ml (*P* = 0.0025), and the FAI dropped significantly from 11.7 \pm 4.9 to 4.7 \pm 3.6 (P < 0.0001). In contrast, SHBG) levels significantly increased from 41.36 \pm 26.70 to 114.02 \pm 57.43 (P < 0.0001). Additionally, the levels of HCY and the LH/FSH ratio showed substantial improvement, decreasing from 9.88 \pm 2.11 to 8.92 ± 1.84 (P = 0.0015) and from 1.48 ± 0.81 to 1.22 ± 0.63 (P = 0.0167), respectively. The data indicate that oral contraceptive treatment significantly improved SHBG levels in both obese and non-obese patients, with values rising from 34.09 ± 24.40 to 115.35 \pm 55.56 and from 94.90 \pm 25.47 to 132.73 \pm 67.28 (P < 0.0001) TABLE 1 Endocrinological and metabolic characteristics of PCOS subjects at baseline and post-treatment (n = 103).

	Baseline	Post-treatment	P-value
Age (year)	29.44 ± 3.74	1	/
WHR	0.85 ± 0.06	1	/
D-D (mg/L FEU)	0.40 ± 1.18	1	/
AMH (ng/ml)	7.19 ± 4.28	1	/
HDL-C	1.46 ± 0.31	/	/
LDL-C	3.27 ± 0.70	1	/
TSH (mU/L)	2.14 ± 1.38	1	/
CHOL (mmol/L)	5.13 ± 0.80	/	/
Systolic	114.50 ± 10.43	1	/
Diastolic	77.10 ± 7.69	1	/
BMI (kg/m ²)	24.80 ± 3.64	24.37 ± 3.72	NS
FSH (mIU/ml)	6.47 ± 1.28	6.40 ± 1.55	NS
LH (mIU/ml)	9.53 ± 5.44	7.43 ± 3.28	0.0018
GLU (mmol/L)	5.43 ± 0.58	5.28 ± 0.50	NS
SHBG	41.36 ± 26.70	114.02 ± 57.43	< 0.0001
INS (IU/ml)	16.66 ± 9.76	17.31 ± 8.24	NS
TG (mmol/L)	1.46 ± 0.94	1	/
HCY (µmol/L)	9.88 ± 2.11	8.92 ± 1.84	0.0015
TESTO (ng/ml)	0.72 ± 0.26	0.62 ± 0.23	0.0025
LH/FSH	1.48 ± 0.81	1.22 ± 0.63	0.0167

(Table 1). Pearson and Spearman correlation analyses revealed a strong association between SHBG levels before and after treatment and factors such as BMI, TESTO, GLU, INS, and HCY (R > 0.2, P < 0.05) (Supplementary Tables S1, S2).

3.3 Machine learning screening of PCOSspecific diagnostic and treatment efficacy assessment indicators

The combined analysis using LASSO and RF identified LH, LH/ FSH ratio, TESTO, SHBG, and HCY as potential specific metabolic markers for PCOS, with AUC values of 0.841, 0.820, 0.619, 0.841, and 0.637, respectively. Our clinical decision model demonstrates a reliable capacity to assess treatment efficacy in PCOS patients, achieving an AUC of 0.637 (95% CI: 0.535-0.737) (Figures 2A, S1A, B). With an AUC greater than 0.7 indicating high specificity for PCOS, LH, LH/FSH ratio, and SHBG are considered the most promising potential biomarkers for the condition (Figures 2B–D). Model training on the cohort of PCOS patients before and after oral contraceptive treatment was conducted to assess the diagnostic efficacy of blood biochemical markers. The results indicated that prior to treatment, only INS demonstrated diagnostic specificity for PCOS (AUC = 0.602). After treatment with oral contraceptives,



SHBG, INS, TESTO, and GLU (AUC > 0.6) were identified as effective markers for evaluating treatment outcomes. Both models successfully predicted the treatment efficacy in PCOS patients receiving oral contraceptives (AUC > 0.6) (Figures 3A, B).

3.4 Risk regression analysis identifies metabolic markers for prognostic risk in PCOS patients

TESTO is already recognized as a serological diagnostic marker for PCOS in clinical practice, as well as a prognostic indicator for adverse pregnancy outcomes in PCOS patients (32, 33). Using TESTO as a risk indicator in regression analysis for PCOS (34). Univariate regression analysis indicated that SHBG serves as a protective factor for PCOS (HR = 1.01, 95% CI: 1.01-1.02, P < 0.0001) (Supplementary Table S3). Multivariate regression analysis further revealed that FSH, LH, and LH/FSH are risk factors for PCOS (P < 0.05), while SHBG continues to be identified as a protective factor (HR = 1.01, 95% CI: 1.01-1.02, P < 0.0001) (Supplementary Table S4). These findings suggest that SHBG is a potential independent prognostic risk factor for PCOS. Additionally, the LH/FSH ratio is a significant clinical indicator for treatment decision-making and the risk of complications in PCOS, demonstrating high diagnostic efficacy (35, 36). By employing the LH/FSH ratio as a clinical decision parameter (Clinic model), we were able to compare the diagnostic efficacy of other metabolic markers associated with PCOS. DCA was



07

performed to evaluate the clinical utility of TESTO, LH, SHBG, and INS in treatment decision-making. The results revealed that, prior to oral contraceptive treatment, SHBG and INS exhibited higher diagnostic efficacy for PCOS compared to the LH/FSH ratio. Following treatment, TESTO, SHBG, and INS demonstrated superior decision-making effectiveness over the LH/FSH ratio (Figures 4A-D). These findings highlight the high specificity of SHBG for the clinical diagnosis and treatment decision-making of PCOS, positioning it as a promising novel biomarker for the condition.

3.5 The association of SHBG with complications and adverse outcomes in PCOS

Univariate regression analysis of the association between SHBG and pregnancy outcomes, folate metabolism status, and embryo transfer outcomes in PCOS patients before and after treatment revealed that SHBG is an independent risk factor for pregnancy outcomes, folate metabolism disorders, and overall pregnancy outcomes prior to treatment (HR: 1.01, 95% CI: 1.01-1.07, P = 0.004; HR: 1.02, 95% CI: 1.01-1.04, P = 0.008; HR: 1.03, 95% CI: 1.01-1.06, P = 0.004) (Supplementary Table S5). These findings indicate that SHBG has significant diagnostic value for the prognosis of PCOS patients before treatment, although its relevance diminishes post-treatment. This suggests that SHBG may serve as a potential therapeutic target for PCOS.

3.6 Comparison of the decision-making efficacy between common clinical indicators for PCOS and SHBG

FAI is a biochemical marker for high-risk complications and adverse reproductive outcomes in clinical PCOS, demonstrating significant clinical value (37, 38). The DCA indicates that both SHBG and FAI exhibit comparable diagnostic capabilities for assessing outcomes after oral contraceptive treatment and predicting reproductive outcomes in PCOS. However, prior to treatment, SHBG demonstrated significantly superior clinical decision-making ability compared to FAI (P = 0.042) (Figures 5A–C, Supplementary Table S6).

3.7 Causal effects between SHBG and genetic plasma proteins in PCOS

Through two-sample bidirectional MR analysis, filtering for IVW *P-values* > 0.05 identified GCNT2, PIGN, KREMEN1, GCDH, CD93, CCDC77, and HSD17B13 as genetic risk proteins associated with PCOS and SHBG levels (Supplementary Figure S2). Notably, CD93 (OR: 1.138, 95% CI: 1.024-1.265, P = 0.016), GCNT2 (OR: 1.177, 95% CI: 1.058-1.310, P = 0.003), HSD17B13 (OR: 1.134, 95% CI: 1.024-1.255, P = 0.015), and PIGN (OR: 1.097, 95% CI: 1.014-1.186, P = 0.021) showed a strong association with the risk of PCOS. In contrast, CCDC77, GCDH, and KREMEN1 were identified as protective factors for the condition (Supplementary Figure S4).



FIGURE 4

Comparison of the decision-making efficacy between metabolic biomarkers and existing clinical markers for PCOS patients before and after treatment. (A–D). Decision curves for assessing treatment outcomes in PCOS patients, comparing TESTO, SHBG, LH, and INS with the established clinical biochemical marker LH/FSH.



Sensitivity analyses using the leave-one-out method, along with tests for horizontal pleiotropy, indicated no evidence of genetic pleiotropy, horizontal diversity or heterogeneity affecting the results (Supplementary Figures S6, S7). Additionally, results from MR-Egger and MR-PRESSO further corroborated these findings. In the GSE87435 cohort, the risk plasma proteins were assessed, yielding a model AUC of 0.931 (95% CI: 0.778-1.000). The AUCs for CD93, GCDH, GCNT2, and PIGN were 0.653, 0.750, 0.847, and 0.569, respectively (Supplementary Figure S3A). Analysis of the expression levels of these risk proteins revealed significant differences in the mRNA expression of GCDH and GCNT2 between PCOS and normal ovarian tissues (P < 0.05) (Supplementary Figure S3C). These findings suggest that SHBG is likely to influence the expression levels of GCNT2, thereby mediating adverse prognoses in PCOS. All risk proteins are located outside the sex chromosomes, with GCNT2, GCDH, PIGN, and HSD17B13 situated on chromosomes 6, 19, 18, and 4, respectively (Supplementary Figure S5B).

3.8 GCNT2 regulates PI3K/AKT signaling in PCOS

In the GSE87435 cohort, an integrative transcriptomic analysis based on GCNT2 expression levels identified 53 differentially expressed genes (DEGs) associated with variations in GCNT2 expression (logFC > 1, P < 0.05) (Figure 6A). PPI analysis revealed that COL1A1, COL4A2, COL5A1, and COL6A2 are central genes within the network (Supplementary Figure S5A). GO enrichment analysis indicated that GCNT2 is linked to collagen synthesis, the composition of ECM structural components, and connective tissue development. Furthermore, KEGG analysis demonstrated that GCNT2 positively regulates several signaling pathways, including ECM-receptor interaction, the AGE-RAGE signaling pathway in diabetic complications, and the PI3K-Akt signaling pathway (Supplementary Figures S3B, C). Notably, COL1A1, COL4A2, and COL6A2 were identified as key downstream factors of GCNT2 that modulate the PI3K-Akt signaling pathway.

3.9 Virtual screening of natural compounds targeting GCNT2

Through screening a natural product compound library, we identified 76 candidate molecules that target GCNT2, selecting the top six lead compounds for further analysis. Vina docking results indicated that Hexose, Meletin, Resveratrol, Cryptotanshinone, Dehydrodieugenol, and Berberine all exhibited specific binding to GCNT2 with relatively low binding energies of -4.7, -7.6, -6.9, -8.0, -7.0, and -8.0 kcal/mol, respectively (Table 2). CDOCKER analysis revealed that Meletin, Resveratrol, Cryptotanshinone, and berberine demonstrated low interaction energies and stable binding in both the CHARMM36 force field and physiological aqueous environments (Table 2). Notably, Meletin and Cryptotanshinone achieved LibdockScores > 100, indicating exceptional binding affinity and target specificity. The ligand interactions with GCNT2 primarily rely on hydrogen bonding, with Cryptotanshinone emerging as the most promising candidate for GCNT2-targeted therapy. It forms stable hydrogen bonds with key side chain amino acid residues HIS-344, LYS-367, HIS-104, and TYR-161 (Figure 6B). Comparative analysis highlighted that the side chain residues ALA-265, ARG-166, HIS-104, and TYR-266 of GCNT2 exhibit high druggability and low molecular binding energies, with their free hydroxyl and carbonyl groups capable of forming stable hydrogen bonds with the lead compounds.

3.10 Pharmacokinetic and toxicokinetic analysis of *Cryptotanshinone*

The analysis of ADME properties *in vivo* indicated that all candidate compounds comply with Lipinski's rule of five, demonstrating high synthetic accessibility (< 5) and drug-likeness



stacking, alkyl interactions, and π -alkyl interactions.

(> 50%). With the exception of hexose, all compounds exhibited favorable gastrointestinal absorption. *Resveratrol*, *Cryptotanshinone*, *Berberine*, and *Dehydrodieugenol* showed high permeability across the blood-brain barrier (BBB) and displayed significant inhibition of cytochrome P450 enzymes, making them less susceptible to rapid degradation. Notably, resveratrol and *Dehydrodieugenol* are non-P-glycoprotein substrates, which reduces the degradation of the drugs during transport. All candidate compounds exhibited high solubility in both aqueous and lipid environments, with *Cryptotanshinone* demonstrating low molecular polarity and flexibility, suggesting that its pharmacophore structure is stable (Supplementary Figures S8A– F). Toxicokinetic parameters and functional group analyses revealed that *Resveratrol, Berberine, Dehydrodieugenol*, and *Meletin* possess considerable hepatotoxicity, neurotoxicity, mutagenicity, hormonal toxicity, strong irritancy, and acute oral toxicity (> 50%). In contrast, *Cryptotanshinone* and hexose exhibited lower biological toxicity (Figure 7A). These findings suggest that *Cryptotanshinone* is a promising candidate for targeting GCNT2 as a low-toxicity, high-efficacy natural compound for the treatment of PCOS. Toxicological functional group analysis indicated no significant toxicological risks associated with the core structure of *Cryptotanshinone*, further supporting its safety profile (Figure 7B).

Candidates	LibdockScore	CDOCKER_energy	Interaction_energy	Vina score
Hexose	65.942	-9.077	-13.472	-4.7
Meletin	105.018	-34.196	-40.543	-7.6
Resveratrol	78.600	-24.433	-34.118	-6.9
Cryptotanshinone	109.774	-37.819	-41.233	-8.0
Berberine	85.994	-20.259	-35.629	-8.0
Dehydrodieugenol	97.444	-10.174	-42.314	-7.0

TABLE 2 Molecular docking screening of potential drugs targeting GCNT2 for PCOS.



3.11 Prediction of the direct effects of *Cryptotanshinone* on the PI3K/Akt signaling pathway

Through molecular docking studies, we investigated the interaction between Cryptotanshinone and the core proteins regulated by GCNT2 in the PI3K/Akt pathway, specifically COL1A1, COL4A2, and COL6A2. The results indicated that Cryptotanshinone can form stable complexes with COL1A1 and COL6A2, exhibiting low binding energies of -9.1 and -9.2 kcal/mol, respectively, along with high target specificity (LibdockScore > 100) (Supplementary Table S7). These interactions are primarily maintained through hydrogen bonding (Supplementary Figures S10A-C). Given that PCOS patients require long-term medication, we also assessed the palatability of the candidate compounds to enhance patient compliance. Predictions indicated that hexose has a 92% probability of being bitter, meletin 72%, resveratrol 68%, while Cryptotanshinone shows a 93% probability of being sweet. Berberine is predicted to be bitter with 100% confidence, and dehydrodieugenol with 91% (Supplementary Figures S9A-F). Consequently, these findings suggest that

patients are likely to have good adherence to *Cryptotanshinone* due to its favorable taste profile.

3.12 GCNT2/SHBG axis modifies PI3K/Akt signaling by N-glycosylation

Using ZDOCK 2.3.2f and ZDOCK 3.0.2f with IRaPPA reranking techniques, we investigated the N-glycosylation modification of SHBG by GCNT2 and identified the relevant binding sites. The results demonstrated that GCNT2 interacts with SHBG through the side chain residues CYS-188 and VAL-358, targeting the protein residues LEU-357 and VAL-16 for Nglycosylation modification. Following this modification, SHBG is transformed into a glycoprotein, which aligns with its role in androgen transport (Figures 8A, B). GCNT2 mediates glycosylation modifications on the core protein COL4A2 of the PI3K/Akt signaling pathway through the side chain residues THR-632 and LYS-692, specifically targeting the side chain residues TRP-5 and SER-13 of the A chain. Additionally, GCNT2 regulates the PI3K/Akt signaling pathway by N-glycosylating several critical



FIGURE 8

Revealing the N-glycosylation modification of SHBG protein by GCNT2 using FFT algorithms and deep learning models. (A, B) Construction of the GCNT2-SHBG complex using ZDOCK 2.3.2f and ZDOCK 3.0.2f with IRaPPA re-ranking algorithms. The stick model represents the side chain amino acid residues of the proteins, with green dashed lines indicating hydrogen bond interactions. The brown protein represents the receptor GCNT2, while the blue protein represents the ligand SHBG.

11



residues on the B and C chains of COL6A2, including LYS-1563 (GLN-130, GLU-157, HIS-127), ASN-1659 (GLN-168, LEU-167), and HIS-1666 (LYS-155, GLU-175) (Figures 9A, B).

4 Discussion

PCOS is a common endocrine disorder that affects 6-10% of women of reproductive age, manifesting as irregular menstrual cycles, hyperandrogenism, and the presence of polycystic ovaries (7, 39). The pathophysiology of PCOS is complex and multifactorial, involving intricate interactions between genetic, environmental, and metabolic factors that give rise to its diverse clinical manifestations. Current diagnostic frameworks, largely based on the Rotterdam criteria, frequently lead to underdiagnosis or misdiagnosis, underscoring the syndrome's inherent heterogeneity and the challenges in its accurate identification (1, 8). Existing biomarkers, such as elevated LH and TESTO, exhibit limitations in sensitivity and specificity (40-42). Furthermore, conventional treatment modalities, including



hormonal contraceptives and insulin sensitizers, may fail to adequately address the underlying metabolic disturbances associated with PCOS (43). Recent research has shed light on the intricate molecular mechanisms underlying PCOS, particularly the pivotal roles of the PI3K/Akt signaling pathway and glycosylation processes (44, 45). These insights highlight the urgent need for innovative diagnostic and therapeutic strategies that not only enhance our understanding of PCOS but also pave the way for personalized treatment approaches tailored to individual patient profiles.

The GCNT2/SHBG axis plays a crucial role in the hormonal regulation and metabolic health of women with PCOS. Recent studies have shown that variations in GCNT2 expression significantly influence levels of SHBG, which in turn modulates the bioavailability of androgens (46). Elevated levels of androgens, often observed in PCOS patients, are associated with numerous metabolic disturbances, including insulin resistance and dyslipidemia (47). This regulatory axis not only underscores the pathophysiological links between hyperandrogenism and metabolic dysfunction but also positions GCNT2 as a potential therapeutic target for restoring metabolic balance (48). Our findings demonstrate that variations in GCNT2 expression significantly influence SHBG levels, thereby modulating the bioavailability of androgens. Elevated androgen levels, commonly observed in PCOS patients, are closely linked to metabolic disturbances such as insulin resistance and dyslipidemia. By elucidating the mechanisms through which GCNT2 regulates SHBG and subsequent androgen dynamics, we provide critical insights into the hormonal dysregulation that underpins PCOS pathophysiology.

Furthermore, alterations in N-glycosylation patterns associated with GCNT2 have significant implications for metabolic dysregulation in PCOS. Our research demonstrates that changes in glycosylation affect insulin signaling pathways, exacerbating insulin resistance: a hallmark of PCOS (49). The aberrant glycosylation of proteins can disrupt normal cellular functions, leading to impaired glucose metabolism and increased adiposity (50). By integrating the roles of the GCNT2/SHBG axis and Nglycosylation, we reveal a complex interplay that contributes to the metabolic phenotype of PCOS. These findings suggest that targeting glycosylation processes may offer new avenues for therapeutic intervention, potentially mitigating the metabolic complications often faced by women with PCOS. Overall, our study emphasizes the importance of the GCNT2/SHBG axis and N-glycosylation in understanding and addressing the multifaceted nature of PCOS, paving the way for more effective, personalized treatment strategies.

The PI3K/Akt signaling pathway plays a crucial role in the metabolic and reproductive dysfunctions associated with PCOS. Our findings indicate that dysregulation of this pathway significantly contributes to insulin resistance and impaired ovarian function, both of which are hallmark features of PCOS. Specifically, we observed that aberrant activation of the PI3K/Akt pathway is linked to increased androgen production and disrupted follicle development, further aggravating the clinical manifestations of the

syndrome (51). This connection underscores the pathway's critical role in bridging metabolic disturbances and hormonal imbalances, suggesting that targeted interventions aimed at modulating this signaling cascade may offer promising therapeutic options for managing PCOS-related complications. Importantly, our research highlights how GCNT2 regulates the PI3K/Akt signaling pathway through N-glycosylation modifications. Changes in GCNT2 expression can alter the glycosylation patterns of key signaling molecules, impacting their function and stability. This modification has profound implications for insulin signaling and ovarian responsiveness, further linking GCNT2 to the metabolic dysfunction observed in PCOS. By influencing the PI3K/Akt pathway via N-glycosylation, GCNT2 serves as a critical mediator that not only affects androgen levels but also addresses the underlying insulin resistance that characterizes the syndrome. This intricate interplay suggests that targeting GCNT2 could provide a dual therapeutic strategy: enhancing ovarian function while simultaneously improving metabolic health. Therefore, understanding the role of GCNT2 in modulating the PI3K/Akt pathway through glycosylation presents an exciting avenue for developing novel interventions aimed at alleviating the multifaceted challenges faced by women with PCOS (52).

Natural products have demonstrated the advantages of high efficiency, low toxicity, and easy synthesis in the treatment of diseases (53). Cryptotanshinone, bioactive compound derived from Salvia miltiorrhiza, exhibits considerable therapeutic potential for various metabolic and endocrine disorders, particularly PCOS. Recent research demonstrates that Cryptotanshinone enhances insulin sensitivity by activating the PI3K/Akt signaling pathway, a crucial mechanism often disrupted in PCOS, facilitating improved glucose uptake in peripheral tissues (13). Additionally, this compound displays anti-androgenic effects, evidenced by its ability to reduce androgen production in ovarian cells, thereby alleviating hyperandrogenism symptoms commonly experienced by affected women. Cryptotanshinone possesses antiinflammatory properties, effectively downregulating proinflammatory cytokines, which play a significant role in the pathophysiology of PCOS (54). Our findings also indicate that Cryptotanshinone may modulate glycosylation processes, particularly influencing the GCNT2/SHBG axis, further enhancing metabolic health (55). In preclinical studies, treatment with Cryptotanshinone restored normal ovarian morphology and function, indicating its potential to improve fertility outcomes in women with PCOS (56).

From a pharmacokinetic perspective, *Cryptotanshinone* has shown favorable absorption and distribution characteristics, with minimal toxicity reported, making it a promising candidate for clinical applications (57). Molecular docking studies demonstrate that *Cryptotanshinone* exhibits strong interactions with key proteins implicated in PCOS-related pathways, offering mechanistic insights into its therapeutic potential. These findings suggest that *Cryptotanshinone* effectively targets hormonal imbalances and metabolic dysfunction, providing a dual therapeutic approach that enhances ovarian function and improves metabolic health. Further exploration of its synergistic effects with other therapeutic agents holds promise for developing integrated treatment strategies, advancing the management of PCOS (58).

Future research should focus on validating *Cryptotanshinone* associated biomarkers and therapeutic candidates to enhance their clinical applicability in PCOS management. Longitudinal studies are crucial to assess the long-term efficacy, safety, and potential cumulative effects of *Cryptotanshinone* and related natural compounds.

5 Limitations

While our study highlights the therapeutic potential of *Cryptotanshinone* in PCOS management through modulation of the GCNT2/SHBG axis, several limitations must be acknowledged. First, the preclinical nature of our research, particularly the reliance on MR analyses, necessitates caution in extrapolating these results to clinical settings. Although MR provides robust evidence for causal inference, potential biases such as horizontal pleiotropy or population stratification may confound the observed genetic associations. Future studies should incorporate sensitivity analyses and multivariable MR approaches to address these limitations.

Our sample size, while sufficient for initial analyses, may limit the generalizability of the results, particularly in capturing the full spectrum of PCOS heterogeneity. Larger, multi-center cohorts with diverse ethnic backgrounds are needed to confirm these findings. Additionally, the study's focus on the GCNT2/SHBG axis may not fully capture the complexity of PCOS pathophysiology, which involves intricate metabolic, endocrine, and inflammatory interactions.

Individual variations in treatment response, influenced by genetic, epigenetic, and lifestyle factors, were not extensively addressed. Future research should incorporate pharmacogenomic analyses and longitudinal designs to develop personalized treatment strategies. Moreover, while molecular docking and ADME analyses suggest *Cryptotanshinone* as a promising candidate, experimental validation in preclinical models and clinical trials is essential to translate these findings into therapeutic applications.

Finally, reliance on existing data may overlook emerging biomarkers or pathways relevant to PCOS. Future studies should explore novel therapeutic targets and evaluate synergistic effects of *Cryptotanshinone* with existing therapies. Despite these limitations, our findings provide a robust foundation for further investigation into *Cryptotanshinone* as the potential therapeutic agent, particularly in targeting the GCNT2 pathway to improve PCOS management. We hope these insights will inspire future research to address the identified gaps and advance personalized treatments for PCOS.

6 Conclusion

This study elucidates the critical role of the GCNT2/SHBG axis in regulating N-glycosylation and the PI3K/Akt signaling pathways, offering novel insights into PCOS pathophysiology (Figure 10). Using a well-defined patient cohort and MR, we established causal relationships that underscore GCNT2 as a key biomarker. GCNT2mediated alterations in N-glycosylation modulate SHBG levels, insulin sensitivity, and ovarian function, directly linking metabolic disruptions to PCOS clinical manifestations. Moreover, the dysregulation of the PI3K/Akt pathway emerged as a central mechanism translating these metabolic changes into endocrine dysfunction, highlighting a sophisticated regulatory network governing metabolic and reproductive health.

Advanced machine learning techniques identified novel biomarkers, enhancing diagnostic precision and paving the way for personalized treatment strategies. Protein-protein docking studies revealed *Cryptotanshinone* as a promising therapeutic candidate, with potential to modulate the GCNT2/SHBG axis and influence the PI3K/Akt pathway. This integrative approach deepens our understanding of PCOS molecular mechanisms and underscores the therapeutic potential of targeting these pathways. Collectively, this research establishes a foundation for precision medicine in PCOS, aiming to refine treatments and improve patient outcomes.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving humans were approved by the medical ethics committees of Zhongshan Boai Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YL: Conceptualization, Project administration, Writing – original draft, Writing – review & editing. XL: Data curation, Formal Analysis, Resources, Writing – review & editing. KZ: Software, Validation, Writing – review & editing. JD: Methodology, Visualization, Writing – review & editing. QL: Methodology, Visualization, Writing – review & editing. LZ: Formal Analysis, Investigation, Writing – original draft. SJ: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported

by the Social Science and Technology Research Major Project of Zhongshan (2021B3011), the Graduate Education Innovation Plan Project of Guangdong Province (GDMU2023355, GDMU2023048) and the Medical Research Program of Zhongshan Municipal Health Bureau (2022A020423).

Acknowledgments

The authors sincerely appreciate the generous support received from the Reproductive Center, Boai Hospital of Zhongshan (Third Class, Category A).

Conflict of interest

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

References

 Lizneva D, Suturina L, Walker W, Brakta S, Gavrilova-Jordan L, Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome. *Fertil Steril*. (2016) 106:6–15. doi: 10.1016/j.fertnstert.2016.05.003

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

3. Liu J, Wu Q, Hao Y, Jiao M, Wang X, Jiang S, et al. Measuring the global disease burden of polycystic ovary syndrome in 194 countries: Global Burden of Disease Study 2017. *Hum Reprod.* (2021) 36:1108–19. doi: 10.1093/humrep/deaa371

4. Macut D, Bjekić-Macut J, Rahelić D, Doknić M. Insulin and the polycystic ovary syndrome. *Diabetes Res Clin Pract*. (2017) 130:163-70. doi: 10.1016/j.diabres.2017.06.011

5. Faris RA, Mahdi ZF, Al-Layla AMT. Preliminary study of the insulin growth factor binding protein-3 (IGFBP3) level in Iraqi women with breast cancer. *AIP Conf Proc.* (2021) 2372(1):030003. doi: 10.1063/5.0065368

 Faris RA, Mahdi ZF, Al-Tu'ma MMK, Al-Tu'ma FJ. Surface enhanced Raman spectroscopy based sensitive and specific detection of vitamin D3, glycated hemoglobin, and serum lipid profile of breast cancer patients. *AIP Conf Proc.* (2023) 2830:020012. doi: 10.1063/5.0157207

7. Dumesic DA, Oberfield SE, Stener-Victorin E, Marshall JC, Laven JS, Legro RS. Scientific statement on the diagnostic criteria, epidemiology, pathophysiology, and molecular genetics of polycystic ovary syndrome. *Endocr Rev.* (2015) 36:487–525. doi: 10.1210/er.2015-1018

8. Stener-Victorin E, Teede H, Norman RJ, Legro R, Goodarzi MO, Dokras A, et al. Polycystic ovary syndrome. *Nat Rev Dis Primers*. (2024) 10:27. doi: 10.1038/s41572-024-00511-3

9. Stuck MW, Conley SM, Naash MI. Retinal degeneration slow (RDS) glycosylation plays a role in cone function and in the regulation of RDS-ROM-1 protein complex formation. *J Biol Chem.* (2015) 290:27901–13. doi: 10.1074/jbc.M115.683698

10. Kostopoulou E, Anagnostis P, Bosdou JK, Spiliotis BE, Goulis DG. Polycystic ovary syndrome in adolescents: pitfalls in diagnosis and management. *Curr Obes Rep.* (2020) 9:193–203. doi: 10.1007/s13679-020-00388-9

11. Bendlová B, Zavadilová J, Vanková M, Vejrazková D, Lukásová P, Vcelák J, et al. Role of D327N sex hormone-binding globulin gene polymorphism in the pathogenesis of polycystic ovary syndrome. *J Steroid Biochem Mol Biol.* (2007) 104:68–74. doi: 10.1016/j.jsbmb.2006.10.002

12. Imran RE, Faris RA. The utility of SHBG assay in clinical biochemistry for Iraqi women. *Res J Biotechnol.* (2024) 19:35–9. doi: 10.25303/1907rjbt035039

13. Li ZW, Kuang Y, Tang SN, Li K, Huang Y, Qiao X, et al. Hepatoprotective activities of Antrodia camphorata and its triterpenoid compounds against CCl(4)-induced liver injury in mice. *J Ethnopharmacol.* (2017) 206:31–9. doi: 10.1016/j.jep.2017.05.020

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2025.1534733/ full#supplementary-material

14. Perez M, Chakraborty A, Lau LS, Mohammed NBB, Dimitroff CJ. Melanomaassociated glycosyltransferase GCNT2 as an emerging biomarker and therapeutic target. *Br J Dermatol.* (2021) 185:294–301. doi: 10.1111/bjd.v185.2

15. Liu Z, Hong Z, Ma H, Yu D, Qu P. Key factors mediated by PI3K signaling pathway and related genes in endometrial carcinoma. *J Bioenerg Biomembr*. (2020) 52:465–73. doi: 10.1007/s10863-020-09854-4

16. Rebar RW, Keator CS. Polycystic ovary syndrome: consider the entire spectrum. *Fertil Steril.* (2024) 121:934–6. doi: 10.1016/j.fertnstert.2024.02.008

17. Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop groupRevised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod.* (2004) 19:41–7. doi: 10.1093/humrep/deh098

18. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab.* (1999) 84:3666–72. doi: 10.1210/jcem.84.10.6079

19. 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. doi: 10.1093/humrep/deab250

20. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med.* (2020) 26:252-8. doi: 10.1038/s41591-020-0751-5

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

22. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet.* (2021) 53:1712–21. doi: 10.1038/s41588-021-00978-w

23. Zhang C, Qin F, Li X, Du X, Li T. Identification of novel proteins for lacunar stroke by integrating genome-wide association data and human brain proteomes. *BMC Med.* (2022) 20:211. doi: 10.1186/s12916-022-02408-y

24. 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. doi: 10.1007/s10654-015-0011-z

25. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* (2015) 44:512–25. doi: 10.1093/ije/dyv080

26. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet.* (2018) 50:693–8. doi: 10.1038/s41588-018-0099-7

27. 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. doi: 10.1097/EDE.000000000000559

28. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics*. (2019) 35:4851–3. doi: 10.1093/bioinformatics/btz469

29. Li Y, Song X, Huang Y, Zhou S, Zhong L. Genetic associations of plasma metabolites with immune cells in hyperthyroidism revealed by Mendelian randomization and GWAS-sc-eQTLs xQTLbiolinks analysis. *Sci Rep.* (2025) 15 (1):1377. doi: 10.1038/s41598-025-85664-1

30. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* (2018) 46:W296-w303. doi: 10.1093/nar/gky427

31. Di Stefano M, Galati S, Piazza L, Granchi C, Mancini S, Fratini F, et al. VenomPred 2.0: A novel in silico platform for an extended and human interpretable toxicological profiling of small molecules. *J Chem Inf Model.* (2024) 64:2275–89. doi: 10.1021/acs.jcim.3c00692

32. Cena H, Chiovato L, Nappi RE. Obesity, polycystic ovary syndrome, and infertility: A new avenue for GLP-1 receptor agonists. *J Clin Endocrinol Metab.* (2020) 105:e2695–2709. doi: 10.1210/clinem/dgaa285

33. Hirschberg AL. Approach to investigation of hyperandrogenism in a postmenopausal woman. *J Clin Endocrinol Metab.* (2023) 108:1243–53. doi: 10.1210/ clinem/dgac673

34. Valdimarsdottir R, Wikström AK, Kallak TK, Elenis E, Axelsson O, Preissl H, et al. Pregnancy outcome in women with polycystic ovary syndrome in relation to second-trimester testosterone levels. *Reprod BioMed Online*. (2021) 42:217–25. doi: 10.1016/j.rbmo.2020.09.019

35. Gao R, Liao X, Huang W, Zeng R, Qin L, Bai P. Follicular-phase gnRH agonist protocol is another choice for polycystic ovary syndrome patients with lower LH/FSH and lower AMH levels without increasing severe OHSS risk. *Front Endocrinol (Lausanne).* (2022) 13:905263. doi: 10.3389/fendo.2022.905263

36. Le MT, Le VNS, Le DD, Nguyen VQH, Chen C, Cao NT. Exploration of the role of anti-Mullerian hormone and LH/FSH ratio in diagnosis of polycystic ovary syndrome. *Clin Endocrinol (Oxf)*. (2019) 90:579–85. doi: 10.1111/ cen.2019.90.issue-4

37. Deshmukh H, Papageorgiou M, Wells L, Akbar S, Strudwick T, Deshmukh K, et al. The effect of a very-low-calorie diet (VLCD) vs. a moderate energy deficit diet in obese women with polycystic ovary syndrome (PCOS)-A randomized controlled trial. *Nutrients.* (2023) 15(18):3872. doi: 10.3390/nu15183872

38. Rahmatnezhad L, Moghaddam-Banaem L, Behrouzi Lak T, Shiva A, Rasuli J. Free androgen index (FAI)'s relations with oxidative stress and insulin resistance in polycystic ovary syndrome. *Sci Rep.* (2023) 13:5118. doi: 10.1038/s41598-023-31406-0

39. Stener-Victorin E, Padmanabhan V, Walters KA, Campbell RE, Benrick A, Giacobini P, et al. Animal models to understand the etiology and pathophysiology of polycystic ovary syndrome. *Endocr Rev.* (2020) 41(4):bnaa010. doi: 10.1210/endrev/bnaa010

40. Diamanti-Kandarakis E, Dunaif A. Insulin resistance and the polycystic ovary syndrome revisited: an update on mechanisms and implications. *Endocr Rev.* (2012) 33:981–1030. doi: 10.1210/er.2011-1034

41. Dunaif A. Insulin resistance and the polycystic ovary syndrome: mechanism and implications for pathogenesis. *Endocr Rev.* (1997) 18:774–800. doi: 10.1210/er.2011-1034

42. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, et al. The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertil Steril.* (2009) 91:456–88. doi: 10.1016/j.fertnstert.2008.06.035

43. Diedrich K, Fauser BC, Devroey P, Griesinger G. The role of the endometrium and embryo in human implantation. *Hum Reprod Update.* (2007) 13:365–77. doi: 10.1093/humupd/dmm011

44. Eleftheriadis T, Pissas G, Crespo M, Filippidis G, Antoniadis N, Liakopoulos V, et al. The effect of anti–HLA class I antibodies on the immunological properties of human glomerular endothelial cells and their modification by mTOR inhibition or GCN2 kinase activation. *Mol Med Rep.* (2021) 23(5):355. doi: 10.3892/mmr.2021.11994

45. Wu H, Xu L, Ballantyne CM. Dietary and pharmacological fatty acids and cardiovascular health. *J Clin Endocrinol Metab.* (2020) 105:1030-45. doi: 10.1210/ clinem/dgz174

46. Bhargava A, Arnold AP, Bangasser DA, Denton KM, Gupta A, Hilliard Krause LM, et al. Considering sex as a biological variable in basic and clinical studies: an endocrine society scientific statement. *Endocr Rev.* (2021) 42:219–58. doi: 10.1210/endrev/bnaa034

47. Van Steijvoort E, Chokoshvili D, Cannon JW, Peeters H, Peeraer K, Matthijs G, et al. Interest in expanded carrier screening among individuals and couples in the general population: systematic review of the literature. *Hum Reprod Update*. (2020) 26:335–55. doi: 10.1093/humupd/dmaa001

48. Wang J, Xu H, Yuan J, Guo C, Hu F, Yang W, et al. association between obstructive sleep apnea and metabolic abnormalities in patients with hypertrophic cardiomyopathy. *J Clin Endocrinol Metab.* (2021) 106:e2309–21. doi: 10.1210/clinem/dgab015

49. Jiang H, Li Z, Yang W, Sun Y, Yan F, Sun Q, et al. Region-specific disturbed iron redistribution in Cushing's disease measured by magnetic resonance imaging-based quantitative susceptibility mapping. *Clin Endocrinol (Oxf)*. (2022) 97:81–90. doi: 10.1111/cen.14692

50. Salehi M, Gastaldelli A, DeFronzo R. Prandial hepatic glucose production during hypoglycemia is altered after gastric bypass surgery and sleeve gastrectomy. *Metabolism.* (2022) 131:155199. doi: 10.1016/j.metabol.2022.155199

51. Drucker DJ. Coronavirus infections and type 2 diabetes-shared pathways with therapeutic implications. *Endocr Rev.* (2020) 41(3):bnaa011. doi: 10.1210/endrev/bnaa011

52. McLean BA, Wong CK, Campbell JE, Hodson DJ, Trapp S, Drucker DJ. Revisiting the complexity of GLP-1 action from sites of synthesis to receptor activation. *Endocr Rev.* (2021) 42:101–32. doi: 10.1210/endrev/bnaa032

53. Li Y, Miao J, Liu C, Tao J, Zhou S, Song X, et al. Kushenol O regulates GALNT7/ NF-κB axis-mediated macrophage M2 polarization and efferocytosis in papillary thyroid carcinoma. *Phytomedicine*. (2025) 138:156373. doi: 10.1016/ j.phymed.2025.156373

54. Li SS, Wang LL, Liu M, Jiang SK, Zhang M, Tian ZL, et al. Cannabinoid CB₂ receptors are involved in the regulation of fibrogenesis during skin wound repair in mice. *Mol Med Rep.* (2016) 13:3441–50. doi: 10.3892/mmr.2016.4961

55. Greenwood EA, Yaffe K, Wellons MF, Cedars MI, Huddleston HG. Depression over the lifespan in a population-based cohort of women with polycystic ovary syndrome: longitudinal analysis. *J Clin Endocrinol Metab.* (2019) 104:2809–19. doi: 10.1210/jc.2019-00234

56. Ueno S, Berntsen J, Ito M, Uchiyama K, Okimura T, Yabuuchi A, et al. Pregnancy prediction performance of an annotation-free embryo scoring system on the basis of deep learning after single vitrified-warmed blastocyst transfer: a single-center large cohort retrospective study. *Fertil Steril.* (2021) 116:1172–80. doi: 10.1016/j.fertnstert.2021.06.001

57. Chen W, Lu Y, Chen G, Huang S. Molecular evidence of cryptotanshinone for treatment and prevention of human cancer. *Anticancer Agents Med Chem.* (2013) 13:979–87. doi: 10.2174/18715206113139990115

58. Zhang QQ, Huang WQ, Gao YQ, Han ZD, Zhang W, Zhang ZJ, et al. Metabolomics reveals the efficacy of caspase inhibition for saikosaponin D-induced hepatotoxicity. *Front Pharmacol.* (2018) 9:732. doi: 10.3389/fphar.2018.00732

Glossary

PCOS	Polycystic ovary syndrome	eQTL	Expression quantitative trait loci
GCNT2	Glucosaminyl (N-acetyl) transferase 2	LD	Linkage disequilibrium
SHBG	Sex hormone-binding globulin	SNPs	Single nucleotide polymorphisms
MR	Mendelian randomization	WR	Wald ratio
EE-CA	Estradiol-cyproterone acetate	IV	Instrumental variable
AIDD	Artificial intelligence-driven drug design	IVW	Inverse variance weighting
WHR	Waist-to-hip ratio	OR	Odds ratios
FSH	Follicle-stimulating hormone	95%CI	95% confidence intervals
LH	Luteinizing hormone	FDR	False discovery rate
TESTO	Total testosterone	ROC	Receiver operating characteristic
FAI	Free testosterone	GO	Gene Ontology
LDL-C	Low-density lipoprotein cholesterol	KEGG	Kyoto Encyclopedia of Genes and Genomes
HDL-C	High-density lipoprotein cholesterol	HTVS	High-throughput virtual screening
TG	Triglycerides	DS	Discovery Studio 2019
total-C/CHOL	Total cholesterol	ADME	Absorption, distribution, metabolism and excretion
АМН	Anti-Müllerian Hormone	FFT	Fast fourier transform
RF	Random forest	ННСҮ	Hyperhomocysteinemia
DCA	Decision curve analysis	INS	Insulin
GWAS	Genome-wide association study	DEGs	Differentially expression genes
pQTL	Protein quantitative trait loci	BBB	Blood-brain barrier.