

PATHOGENETIC INTERRELATIONSHIP BETWEEN INSULIN RESISTANCE AND HYPERANDROGENISM IN POLYCYSTIC OVARY SYNDROME: MOLECULAR MECHANISMS, CLINICAL EVIDENCE, REAL-WORLD DATA, AND TRANSLATIONAL OPPORTUNITIES

Xolboyeva Mashhura Yoqutjon qizi

Email – yoqutjonovnam@gmail.com

Termez University of Economics and Service

Faculty of Medicine

Abstract

Background: Polycystic ovary syndrome (PCOS) is a heterogeneous reproductive–metabolic disorder in which insulin resistance, hyperinsulinemia, androgen excess, altered adipose biology and ovarian dysfunction interact rather than operate as independent abnormalities. Objective: To critically synthesize the mechanistic bidirectionality of insulin resistance and hyperandrogenism and examine how randomized trials, real-world evidence and computational methods may refine disease stratification and treatment evaluation. Methods: A structured, non-systematic evidence synthesis prioritized the 2023 International Evidence-based PCOS Guideline, indexed mechanistic reviews, randomized-trial syntheses and regulatory documents concerning real-world evidence and artificial intelligence. This article does not report new patient-level analyses or a completed PRISMA systematic review. Results: Insulin may amplify luteinizing-hormone-dependent theca-cell steroidogenesis and decrease hepatic sex hormone-binding globulin, increasing circulating bioavailable testosterone. Conversely, androgen excess may alter adipocyte function, body-fat distribution, skeletal-muscle insulin signalling and insulin secretion. Neither mechanistic direction is universally dominant; phenotype, adiposity, age and treatment exposure alter the observable association. The 2023 international guideline estimates PCOS prevalence at approximately 10–13% in reproductive-age women, conditional on diagnostic criteria, and discourages routine clinical insulin-resistance assays because available measures lack reliability. Therapeutic and observational evidence must therefore distinguish metabolic endpoints, reproductive outcomes and patient-centred effects. Real-world registries and cautiously validated machine learning can complement controlled trials, but confounding, assay heterogeneity, selection bias and subgroup underrepresentation complicate causal interpretation. Conclusion: PCOS is more accurately conceptualized as a network of reciprocal endocrine and metabolic feedback

loops than as a single linear insulin-to-androgen pathway. Translational progress requires standardised androgen measurement, longitudinal metabolic phenotyping, prospectively registered causal analyses, equity-conscious algorithms and clinically meaningful outcomes.

Keywords: polycystic ovary syndrome; insulin resistance; hyperinsulinemia; hyperandrogenism; ovarian theca cells; sex hormone-binding globulin; adipose dysfunction; real-world evidence; causal inference; machine learning; precision medicine; pharmacovigilance.

1. Introduction

Clinical and biological framing

PCOS cannot be inferred from an ultrasound image alone: diagnostic criteria require appropriate combinations of androgen excess, ovulatory dysfunction, and polycystic ovarian morphology or, in adults, an eligible anti-Müllerian hormone alternative, after exclusion of competing diagnoses. Phenotypes differ in insulin sensitivity, adiposity and reproductive expression. Women with normal body mass index may have clinically important metabolic abnormalities, while weight gain can accentuate pre-existing susceptibility. The consensus guideline underscores that diagnostic decisions should not depend on an insulin assay routinely offered in commercial panels. This distinction matters when mechanistic biomarkers are translated into clinical pathways [1,2].

Non-linear causation

The conventional account of PCOS begins with insulin resistance and compensatory hyperinsulinemia, proceeding to ovarian androgen synthesis, suppressed hepatic sex hormone-binding globulin and anovulation. Yet longitudinal causal sequence is unsettled: insulin hypersecretion or altered hepatic insulin clearance could precede measurable peripheral resistance in some individuals. Excess androgens may themselves worsen insulin action, providing a feedback explanation for heterogeneity rather than requiring a single initiating lesion. An association between fasting insulin and free testosterone cannot identify which process began first; medication exposure, adiposity and pubertal maturation may confound both [3,4].

Meaning of RWD and RWE

Real-world data (RWD) are routinely generated observations about health and healthcare delivery, including clinical records, claims, registries, pharmacy dispensing

and patient-generated measurements. Real-world evidence (RWE) refers to clinical evidence derived from appropriate analysis of RWD concerning use, benefits or risks of an intervention. The distinction is epistemic: a data repository becomes evidence only through a clearly defined research question, trustworthy measurements, a valid design and transparent inference. The FDA has issued guidance on data fitness, registries, submissions and non-interventional studies; none implies that an uncontrolled association automatically establishes drug effectiveness [7–11].

Efficacy versus effectiveness

Randomised controlled trials (RCTs) usually estimate treatment efficacy under protocol-specified selection, monitoring and adherence. In contrast, effectiveness describes performance under ordinary conditions where discontinuation, access barriers, comorbidities and co-interventions operate. PCOS illustrates why this is clinically consequential: improvement in menstrual regularity, hirsutism, conception, weight, glycaemia and patient-reported wellbeing may differ in magnitude or timing. RCT evidence is indispensable for treatment decisions, but observational follow-up can reveal persistence, treatment switching and infrequent adverse outcomes excluded by short follow-up. Observational emulation must approximate baseline randomization and control biases rather than treating scale as an antidote to confounding [1,5,7,8].

Regulatory and ethical context

WHO guidance requires autonomy, safety, transparency, accountability, inclusion and sustainability for health-related artificial intelligence. FDA oversight applies to qualifying AI-enabled medical-device functions and evaluates safety and effectiveness in the context of intended use. EMA initiatives have developed infrastructure for observational evidence, including the Data Analysis and Real World Interrogation Network (DARWIN EU). These bodies do not endorse deploying unvalidated PCOS prediction algorithms. Their relevance is governance: clinically derived evidence must remain auditable, fair, reproducible and proportionate to risk [12–16].

Why PCOS is a useful translational model

Few common disorders so clearly juxtapose endocrine regulation, metabolic disease and reproductive outcomes. The ovarian theca cell, liver, muscle, pancreas, adipose tissue and hypothalamic–pituitary axis each receive different insulin and androgen signals. This creates a natural test case for multimodal precision medicine, but also a setting where plausible biological stories may outrun actual evidence. This review

therefore separates well-supported clinical recommendations, experimental pathway models, methodological proposals and areas of genuine uncertainty rather than interpreting all mechanistic hypotheses as proven in patients [1–4].

2. Materials and Methods

Design: critical integrative narrative review incorporating methodological appraisal. The evidence base was assembled from identifiable indexed publications and official organizational documents; it was not generated by a prospective study protocol, original database download or formal, reproducible PRISMA workflow. Searches were oriented to PubMed-indexed PCOS guidelines and mechanistic reviews, with contextual consideration of Cochrane Library, Embase, Scopus, Web of Science and publishers such as JAMA, Lancet, NEJM and Nature Medicine; coverage of each commercial database was not independently verified. No invented screening totals, pooled estimates, confidence intervals or heterogeneity statistics are reported.

Eligibility logic: prioritize human PCOS cohorts, experimental studies that identify tissue-specific mechanisms, international clinical practice guidelines, RCTs and systematic reviews, with emphasis on literature published from 2020 through October 2026. Older methodological or seminal sources were retained where indispensable. Exclude articles with unverifiable provenance, abstracts lacking ascertainable endpoints, studies conflating phenotype definitions, and prediction algorithms lacking outcome validation. Evidence is narrated by design, applicability, biological plausibility, confounding risk and transferability. The lack of dual-reviewer screening and preregistration limits claims of comprehensiveness.

Propensity score matching (PSM)

Estimate treatment probability from pretreatment covariates, establish common support, match observations and assess standardized mean differences rather than relying only on significance tests. PSM cannot balance unmeasured endocrine severity or clinician preferences.

Inverse probability weighting (IPW)

Weight treatment groups by inverse propensity probabilities to emulate a balanced pseudo-population. Inspect extreme weights, positivity, time-varying confounding and truncation. In longitudinal PCOS care, treatment decisions often respond to evolving weight and glycaemia.

Cox proportional-hazards regression

Estimate relative hazard for time-to-event endpoints such as incident diabetes under proportional hazards assumptions; check time interactions, competing events and delayed entry. Hazard ratios should not be described as cumulative risk ratios.

Kaplan–Meier and survival analysis

Display event-free survival with correct censoring and risk sets. Switching therapy, loss to follow-up and reproductive status can generate informative censoring; competing-risk curves may be preferable for some outcomes.

Target-trial emulation

Specify eligibility, treatment strategies, assignment time, grace periods, follow-up, outcomes and estimand before analysis. Align time zero for exposure and comparator groups to prevent immortal-time bias.

Causal inference

Use directed acyclic graphs to identify confounders and mediators, distinguish total from direct effects, consider sensitivity analyses for unmeasured confounding, and avoid conditioning on colliders such as specialist referral.

Machine learning and deep learning

Evaluate penalized regression, tree ensembles, neural architectures and temporal models with patient-level splits and external-site testing. Compare with a transparent baseline, calibrate probabilities and assess clinically consequential thresholds.

Natural language processing and LLMs

Extract symptoms, menstrual-cycle narratives, adverse effects and family history from unstructured notes. Require manual adjudication samples, temporal negation handling, privacy protections and error monitoring; generative text is not independently verified clinical evidence.

Research reporting

For future original observational analyses use STROBE and RECORD where relevant; use TRIPOD-compatible predictive reporting and risk-of-bias assessment for models; declare protocol deviations and prespecified subgroups.

3. Results: Integrated Evidence Synthesis

The following results constitute analytical findings from published clinical and mechanistic evidence, not outcomes of newly enrolled patients. Each subsection distinguishes observations that are incorporated into guidelines from hypotheses requiring prospective validation.

3.1. Molecular coupling of insulin signalling and ovarian steroidogenesis

Theca-cell steroidogenic response

Insulin augments the response of ovarian theca cells to luteinizing hormone, favouring activity within the steroidogenic cascade that generates androgen precursors. The pathway is not equivalent to the action of insulin in skeletal muscle glucose transport: selective insulin resistance may allow ovarian and steroidogenic responses to persist despite impaired systemic glucose disposal. The degree to which individual intracellular branches explain phenotype variation in vivo remains uncertain. LH pulsatility, genetic susceptibility, local growth factors and follicular stage provide alternative or interacting influences [3,4].

Hepatic SHBG suppression

The liver connects metabolic status to circulating free-androgen exposure by regulating sex hormone-binding globulin (SHBG). Insulin-linked metabolic perturbation can lower SHBG, increasing the proportion of bioavailable testosterone even without a matching increase in total testosterone. Consequently, apparent hyperandrogenemia depends partly on assay choice and binding-protein biology. In clinical practice, high-quality total testosterone measurement and properly calculated free testosterone are preferred; direct immunoassays may misclassify mild androgen excess. This mechanism demonstrates why insulin and androgen laboratory variables cannot be interpreted in isolation [1,2].

Pancreatic insulin secretion and clearance

Hyperinsulinemia is often treated as a compensation for defective insulin-mediated glucose uptake, but insulin secretion and hepatic clearance are separately regulated.

Individuals with similar sensitivity estimates may experience distinct insulin exposure and divergent ovarian responses. Contemporary scholarship contests a universally fixed temporal ordering between hypersecretion and resistance; different phenotypes may have different proximal drivers. Studies using a single fasting sample are poorly equipped to determine dynamic secretion, clearance or feedback loops, emphasizing a need for longitudinal and physiological measurement [3].

Androgens as metabolic effectors

Androgens can affect adipose distribution, adipocyte differentiation, inflammatory signalling, skeletal-muscle substrate use and aspects of pancreatic function. Human and animal evidence supports biological plausibility but effect sizes vary by tissue, exposure concentration and developmental stage. Obesity should neither be treated as a compulsory intermediary nor excluded as an important amplifier. Strong causal claims require experiments isolating androgen exposure and longitudinal analyses that distinguish an androgen-driven metabolic change from shared antecedents [3,4].

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

3.2. Phenotypes, diagnosis and clinical stratification

Phenotype heterogeneity

PCOS is an umbrella diagnosis rather than a homogeneous metabolic entity. The combination of menstrual disturbance, biochemical or clinical hyperandrogenism and ovarian morphological features affects apparent prevalence and prognosis. The guideline summary cites approximately 10–13% prevalence in reproductive-age women, varying with criteria and setting. It would be erroneous to assign this figure to Uzbekistan or to every age stratum without population-based ascertainment [1,2].

Differential diagnosis

Thyroid dysfunction, hyperprolactinemia, non-classic congenital adrenal hyperplasia, Cushing syndrome where clinically suspected, androgen-secreting tumours and medication effects should be considered according to presentation. Rapid virilization,

marked androgen elevations or sudden severe progression require investigation beyond a routine PCOS label. Menstrual irregularity after menarche also needs age-appropriate interpretation because adult ultrasound thresholds and anti-Müllerian hormone criteria cannot simply be applied to adolescents [1,2].

Insulin resistance testing caveat

Insulin resistance is a key pathophysiological feature, but guideline authors caution that routine clinical measures are inaccurate and do not currently recommend their routine use as a diagnostic test. HOMA-IR may be useful in explicitly designed research, but no universal cutoff can be exported across laboratories and populations. Clinical assessment of glycaemic risk requires established metabolic testing and consideration of diabetes risk factors rather than reliance on fasting insulin as a diagnostic shortcut [1,2].

Outcome prioritization

Fertility, menstrual regularity, hirsutism, acne, androgen biochemistry, dysglycaemia, blood pressure, adiposity, mental health and quality of life describe different outcome domains. Improvement in one domain cannot be assumed to imply improvement in the others. Patient-reported satisfaction and therapeutic burden are particularly relevant when agents are continued for years. A research report should specify the exact endpoint and observation window and distinguish symptomatic benefit from reversal of underlying pathophysiology [1,5].

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

3.3. Real-world clinical data sources and interpretive value

Electronic health records

Longitudinal encounter histories can link PCOS diagnosis, body-mass trajectories, glycaemic investigations, prescriptions and pregnancy outcomes. Yet clinician coding may privilege classic phenotypes and miss patients primarily treated for acne or infertility. A robust case definition should combine diagnosis codes with menstrual and

biochemical information and, when possible, undergo chart adjudication. The date of first recorded diagnosis is not the date of biological onset; conflating them will distort causal models and natural-history estimates [7–10].

Electronic prescribing and pharmacy records

Prescribing datasets clarify the intended clinical strategy, while dispensing histories better approximate access. Neither is a direct measure of medicine ingestion. Combining orders, fills and treatment changes can describe adherence and persistence, but differences in insurance coverage, public supply chains and over-the-counter use must be addressed. For metformin, contraceptives and incretin-based medicines, classification of indication is crucial because the same agent may be used for other conditions [7–10].

Laboratory repositories

Glucose challenge tests, glycated haemoglobin, lipids and hormone results provide objective longitudinal variables when units, reference ranges and assay platforms are harmonized. Testosterone assays are especially vulnerable to interchangeability assumptions; an LC–MS/MS result and a low-range immunoassay value may not be directly comparable. Missing endocrine assays may indicate selective testing of symptomatic patients rather than random absence; interpretation therefore requires missingness diagnostics [1,2].

National registries and claims

Population registries can establish linkage across facilities and assess long-term diabetes, cardiovascular or reproductive events. Claims extend geographic reach but usually provide sparse physiology and may be biased by reimbursement pathways. Registry transparency requires defined inclusion rules, coding validation, linkage accuracy and secure governance. Apparent increases in PCOS incidence can reflect changed diagnostic access or recording rules, not necessarily altered biological disease frequency [7–10].

Pharmacovigilance datasets

Spontaneous safety reports can identify unusual patterns after exposure to treatment, particularly when an event is uncommon. They cannot provide incidence rates or treatment-specific causal risk by themselves because denominator data, duplicate

reports and stimulated reporting may be absent. A reported association must be evaluated against background risk, co-medications, age and pregnancy status before any clinical claim is made [7,9].

Wearables and mobile applications

Step counts, sleep timing, heart-rate metrics and symptom diaries may describe trajectories inadequately captured by periodic appointments. But device adoption is socially patterned, menstrual tracking may be intermittent, and algorithm-generated features may change after software updates. The data support contextual assessment rather than PCOS diagnosis unless instruments are validated for the specific purpose. Research protocols should document sensor model, missingness, adherence and participant consent [12,13].

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

3.4. Drug effectiveness, trial evidence and causal comparison

Lifestyle intervention and metformin

Individualized lifestyle measures remain central across PCOS phenotypes, while metformin is used for selected metabolic indications within guideline recommendations. A 2023 systematic review and meta-analysis evaluated randomized comparisons of metformin with or without lifestyle modification against placebo, underscoring that treatment effects must be appraised outcome by outcome rather than advertised as global syndrome resolution. Gastrointestinal adverse effects and patient preferences affect real-world persistence; randomized efficacy cannot predict adherence without further evidence [1,5].

Antiandrogenic and contraceptive strategies

Combined oral contraceptive pills can address hyperandrogenic symptoms and cycle concerns where appropriate, whereas antiandrogens require pregnancy-prevention consideration because of potential fetal risks. Because these therapies suppress signs of androgen excess through distinct mechanisms, changes in hormone measures during

treatment cannot be equated with elimination of intrinsic PCOS susceptibility. Observational comparative effectiveness studies must control baseline symptom burden and contraindications, since the choice of medication is inherently confounded by indication [1,2].

Incretin-based treatment and metabolic endpoints

GLP-1 receptor agonists may be relevant in selected patients with excess weight or metabolic comorbidities under applicable care pathways. Changes in weight, insulin sensitivity, cycle parameters and androgen indices are related yet distinct; extrapolation of diabetes and obesity trial findings to reproductive PCOS outcomes requires PCOS-specific evidence. Treatment discontinuation, access, tolerability and pregnancy planning complicate extended follow-up, particularly in observational cohorts without reliable fertility-intention documentation [1,2].

Comparators across other diseases

In diabetes, treatment effectiveness is typically measured with glycaemia, metabolic events and complications; cardiovascular medicine uses hard events and risk-factor trajectories; oncology frequently focuses on progression and survival; infectious disease studies emphasize microbiological or clinical resolution; neurology uses disease-specific function or relapse endpoints. These templates illustrate transferable causal methods, not evidence that PCOS medications prevent cancer, infections or neurologic disease. Mechanistically plausible cross-disease models must not silently merge incompatible outcome definitions [7–10].

Target-trial design for PCOS

A realistic target trial could compare initiation of metformin versus non-initiation among eligible adults with a defined metabolic profile, assigning time zero to the same clinical decision point. Prespecify treatment strategies, crossover rules, one-year glycaemic outcomes and relevant tolerability endpoints. A propensity score can balance measured pretreatment predictors; inverse weighting or clone-censor-weight designs address treatment switching under appropriate assumptions. Explicit negative controls and quantitative bias analysis help demonstrate how much residual confounding might alter findings [7–10].

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic

risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

3.5. Artificial intelligence for prediction and evidence generation

Supervised machine learning

A classifier may integrate cycle history, measured androgens, metabolic tests, demographics and treatment history to predict dysglycaemia or loss to follow-up. Predictors collected after the clinical decision must be excluded from a baseline model to prevent leakage. Performance should be evaluated against logistic regression and guideline-based rules, emphasizing calibration, sensitivity in underrepresented phenotypes and net clinical benefit rather than area-under-curve alone. No single published algorithm justifies universal PCOS deployment without external clinical validation [12–16].

Deep learning and multimodal fusion

Neural architectures can handle ultrasound imaging, sequential notes and dense physiological measures, but imaging recognition of polycystic morphology is not equivalent to syndrome diagnosis. Models trained on images selected because a specialist suspected PCOS risk incorporating referral or operator signals rather than disease biology. Multi-centre acquisition variation, image quality and hidden patient overlap are central hazards. A sensible pathway begins with transparent image annotations, independently held-out institutions and clinical reader comparisons [12–16].

Reinforcement learning

Sequential decision algorithms could theoretically suggest treatment adjustments to improve metabolic and reproductive trajectories. Nevertheless, reinforcement learning depends on reliable counterfactual outcomes, well-defined states and safe action constraints. Retrospective healthcare data offer limited exploration of untried actions and may reward recording intensity rather than health. Offline policy evaluation cannot substitute for prospective safety testing; autonomous therapy modification in PCOS would require stringent governance and clinical oversight [12–16].

NLP and large language models

Natural language processing can recover menstrual irregularity, hirsutism descriptions, lifestyle barriers and adverse-effect narratives omitted from structured fields. Large language models may draft phenotype summaries or support chart abstraction; however, hallucinated medication history, imprecise temporal grounding and leakage of sensitive reproductive information create unacceptable risks without review. Validation should compare extracted variables with clinician-adjudicated records and measure subgroup-specific error and abstention quality [12,13].

Clinical decision support

A meaningful clinical decision-support system combines evidence-based thresholds, explicit uncertainties and understandable explanations. It should alert to metabolic screening, review contraindications and support shared decisions, rather than assign causal authority to correlational risk scores. Workflow studies must show that the system changes decisions beneficially without creating alert fatigue or widening gaps between urban specialty care and resource-limited settings. Prospective impact evaluation is a different claim from algorithmic benchmark performance [12–16].

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

3.6. Pharmacovigilance, genetics and precision medicine

Adverse drug reactions

Treatment choices for PCOS require clinically contextualized safety monitoring. Metformin-related gastrointestinal symptoms, contraindications to estrogen-containing regimens, pregnancy-related concerns with antiandrogens and adverse reactions to weight-management agents have different monitoring pathways. Pharmacovigilance signals are hypothesis-generating; signal detection statistics are not event incidence estimates. Causal assessment should consider exposure duration, dechallenge, competing causes and reporting dynamics [1,5,7].

Pharmacogenetics and pharmacogenomics

Genetic heterogeneity may influence insulin action, ovarian steroidogenesis and individual treatment response, but broad genomic profiling is not yet a validated routine method for selecting all PCOS therapies. Discriminative power, ancestry transferability and clinical utility must be demonstrated before genetic associations are used as treatment rules. A marker predicting biochemical response does not automatically predict live birth, glycaemic complications or quality of life. Pharmacogenetic claims should distinguish replicated pharmacodynamic effects from exploratory signals [1,3].

Biomarker hierarchy

Useful biomarker candidates include accurately measured testosterone, SHBG-derived free-androgen assessment, anthropometric variables, glucose tolerance and selected investigational metabolic or inflammatory signatures. Clinical biomarker status depends on analytical validity, clinical validity and evidence that acting on the result improves outcomes. The complexity of insulin physiology makes it unlikely that a single fasting value will universally partition PCOS into causal subtypes. Multi-marker panels require validation that survives changes in age, body composition and treatment exposure [1–4].

Precision medicine as intervention

A precision strategy should begin by defining an actionable patient subgroup, the available treatment choices and an outcome that matters to patients. Clustering phenotypes can be useful descriptively, yet cluster boundaries may be unstable across measurement protocols. A treatment-benefit model must predict the difference between potential outcomes under alternative therapies, not just prognosis under a chosen therapy. Personalization therefore requires robust trial or trial-emulation evidence of heterogeneous treatment effects [7–10].

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

Healthcare digitalization context

For Uzbekistan, the immediate opportunity is to develop reliable reproductive–metabolic care pathways that can eventually be audited across public and private providers. Institutional initiatives for electronic records, digital prescriptions and longitudinal registries require confirmation of their current scope and implementation status through official Ministry of Health documents; this review does not assert a nationwide level of completion. A PCOS registry should use consent-compatible data governance and standard fields for cycle history, androgen results, metabolic status and treatment exposure.

Electronic prescriptions and longitudinal linkage

A structured e-prescription record can document therapeutic initiation, switching and possible drug–drug interactions, but the ability to link prescriptions to dispensing, laboratory results and clinical endpoints depends on legal permissions and software interoperability. Before causal evaluation, health authorities should publish data dictionaries, standardized identifiers, data-retention policies and quality-control procedures. A pilot may be conducted across selected centres; claims about broader deployment must follow public audit data rather than promotional descriptions.

Laboratory and clinical pharmacology capacity

Variation in hormone assays is a serious barrier to comparing PCOS diagnoses across regions. A harmonization programme could prioritize total testosterone assay methods, reference ranges, glucose testing procedures, body-mass units and systematic treatment-reaction documentation. Clinical pharmacology services could coordinate adverse-effect reviews, medication-use evaluation and research governance, particularly for therapies prescribed off-label. The first deliverables should be validation studies of recorded variables rather than complex predictive models.

AI implementation roadmap

Stage one should audit missingness and assay consistency; stage two should establish a deidentified multicentre dataset with independent ethics review; stage three should externally validate a simple, transparent risk model; stage four should prospectively test decision support under clinician oversight. Each stage needs predefined stopping

criteria, adverse-event reporting, public performance summaries and governance for model revision. This incremental approach is more credible than deploying generative models into reproductive care before demonstrating safety and patient benefit.

Evidence interpretation: these observations should be read against differences in PCOS definitions, selection into specialty care, analytical platforms, background metabolic risk and follow-up duration. Mechanistic plausibility alone does not establish a treatment indication, and observational convergence does not remove the need for design-specific bias appraisal. Whenever a specific numeric treatment effect is absent, this synthesis intentionally refrains from supplying an invented estimate.

4. Discussion

Convergence and tension in international evidence

International guidance and recent mechanistic accounts converge on the importance of a reciprocal endocrine–metabolic network, but they differ in the strength of causal language attached to the initial abnormality. A conceptual model centered on insulin resistance is clinically useful; a model that also permits early hyperinsulinemia, altered clearance and androgen-mediated metabolic effects is more faithful to heterogeneity. The evidence does not establish that every woman enters the feedback loop through the same organ. Clinical recommendations should consequently prioritize verified outcomes rather than mechanistic certainty not yet demonstrated [1–4].

Selection, confounding and collider bias

Women referred for specialist care tend to have more severe symptoms or infertility concerns, and treatment choice reflects perceived severity, access and comorbidity. Conditioning on specialty attendance can induce collider bias between biological severity and healthcare-seeking behaviour. Propensity matching cannot repair unrecorded indications, while machine learning can excel at predicting who is tested rather than who truly has disease. Diagnostic and therapeutic models should therefore explicitly represent the care pathway that produced their data [7–10].

Missing data and measurement error

Androgen assays, menstrual records and metabolic investigations are often missing not at random: clinicians order them selectively, and patients facing financial or geographic barriers may not obtain testing. Multiple imputation requires assumptions whose plausibility must be discussed; complete-case analysis may preferentially retain

patients receiving intensive care. Even an apparently high-performing model can be miscalibrated where testosterone assays change. Missingness indicators must not be mistaken for disease-specific biology without validation [1,2,12].

Explainable and ethical AI

Attribution maps or feature importance plots can describe a model without proving causal influence. Explanations must be checked for stability, clinical plausibility and possible encoding of protected or sensitive attributes. WHO emphasizes autonomy, safety, transparency, accountability, inclusiveness and sustainability; compliance requires human review and meaningful consent in reproductive health settings. Performance audits should include age, BMI, geography and clinically appropriate groups while safeguarding privacy [12,13].

Regulatory science and DARWIN EU

FDA RWE guidance focuses on whether the data and design are fit to answer a regulatory question, not whether an enormous observational dataset exists. Its AI-enabled medical-device oversight similarly evaluates intended use and lifecycle concerns. EMA real-world evidence initiatives including DARWIN EU provide an instructive organizational model for distributed observational analyses, but do not confer regulatory approval on a hypothetical PCOS tool. Uzbekistan could adapt governance principles while retaining locally appropriate implementation and legal safeguards [7–16].

Implications for future trials

The most informative prospective work would repeatedly characterize insulin sensitivity, secretion or clearance, ovarian and adrenal androgen production, adiposity distribution and clinical outcomes in explicitly defined phenotypes. Pragmatic randomization and causal trial emulation could test whether treatment benefit differs by baseline metabolic profile. Such designs must state pregnancy intentions and medication exposures, since reproductive decisions substantially affect treatment assignments and censoring. Patient-reported burden deserves equal methodological attention to biochemical surrogates [1,3,5].

5. Conclusion

PCOS arises from reciprocal interactions among insulin resistance, variable insulin secretion and clearance, androgen production, hepatic SHBG regulation, adipose

signalling and the hypothalamic–pituitary–ovarian axis. Insulin resistance is important but neither a universal initiating event nor a sufficiently accurate stand-alone clinical diagnostic measure. The strongest near-term strategy combines guideline-concordant endocrine assessment, individualized metabolic surveillance, patient-prioritized treatment and rigorous appraisal of trial and real-world evidence. Prediction systems should be introduced only after transparent validation, assessment of calibration and subgroup fairness, privacy protections and prospective clinical evaluation. For Uzbekistan, careful standardization of laboratory measures and electronic clinical records is a necessary precursor to credible multi-centre PCOS evidence generation.

Declarations

Funding: Not reported. Conflict of interest: To be completed by the authors. Ethics approval: Not applicable to this critical literature review; no identifiable patient-level data were collected. Data availability: No original dataset was generated. Author contributions: To be supplied. Review limitations: Non-systematic search, no quantitative meta-analysis, and incomplete coverage of paywalled bibliographic services. This manuscript is not an original clinical outcomes study.

6. Extended critical appraisal and methodological research agenda

Temporal ordering of metabolic and ovarian abnormalities

The hypothesis that hyperinsulinemia initiates reproductive dysfunction can only be evaluated if investigators collect measurements before manifestations arise, rather than taking baseline at first diagnosis. Incident PCOS itself is an uncertain endpoint because recognition depends on menstrual maturation and clinician testing. Repeated insulin secretion and clearance assessments, weight changes, steroid profiles and cycle tracking would permit analysis of temporal sequence, but longitudinal associations still require assumptions regarding time-varying shared causes. Pubertal transitions, sleep, glucocorticoid exposure and contraception create changing endocrine environments. A persuasive study would predefine a plausible time scale between metabolic disruption and ovarian response, test alternative lag periods, and report whether androgen levels forecast later insulin changes after prior insulin trajectories are considered. The reverse model should be evaluated symmetrically. Testing only the dominant hypothesis invites confirmation bias, particularly when several interconnected variables are measured with different error rates. Negative findings should not be dismissed solely because the initial model was biologically attractive.

Selective tissue insulin responsiveness

The phrase selective insulin resistance describes a theoretical distinction between insulin actions that remain responsive and those that become less effective. In PCOS, one clinically important observation is that ovarian steroidogenic responsiveness can coexist with impaired peripheral metabolic control. Translating this into a precise molecular model is complicated by differing cell environments and the participation of insulin-like growth factor signalling. Receptor abundance, intracellular adaptor proteins, inflammatory mediators and gonadotropin availability can alter pathway readouts in cultured cells. In-vitro models rarely reproduce the full concentration gradients and endocrine feedback encountered in vivo. Consequently, a single molecular marker derived from ovarian tissue should not be presented as a validated patient-level diagnostic biomarker. Researchers should triangulate cellular perturbation experiments with physiological endocrine challenges and longitudinal patient observations. Findings concordant across these layers deserve greater weight than associations confined to one experimental platform.

Interpreting anthropometry and adipose dysfunction

Body-mass index is convenient but cannot distinguish visceral adipose accumulation from lean mass or ectopic fat. Regional adiposity, inflammation and liver fat may influence insulin metabolism even when total body weight is not strikingly elevated. Studies that statistically adjust for BMI and declare an androgen effect independent of obesity may still leave residual confounding by adipose distribution and developmental exposures. Conversely, controlling for adiposity can remove part of a biologically meaningful mediated effect if androgen excess alters fat accumulation over time. Whether BMI belongs in an adjustment set therefore depends on the causal question and presumed temporal order, not on a universal covariate checklist. Prospective work should measure waist circumference consistently and consider more detailed body-composition measures when feasible. Investigators should report both absolute metabolic risk and measures of association so that a subgroup with a strong relative effect but low baseline risk is not incorrectly portrayed as requiring identical management.

Genetic architecture and environmental modifiers

PCOS susceptibility appears polygenic and environmentally modified; neither a single gene nor a single obesogenic behaviour explains most clinical presentations. Developmental conditions, family history and population ancestry may influence expression, but familial aggregation alone cannot separate inherited variation from

shared exposures. Any genetic risk score requires validation in an ancestry-relevant population before it can be promoted as a treatment-selection instrument. A molecular association between a variant and insulin signalling may provide mechanistic clues while yielding little incremental predictive value beyond ordinary clinical examination. Similarly, an association between environmental exposures and PCOS phenotypes requires careful attention to timing, exposure accuracy and socioeconomic confounding. Claims that a lifestyle choice caused an individual PCOS are therefore unjustified. A balanced explanation can recognize modifiable metabolic risk without attributing blame or presenting untested genomic approaches as established care.

Measurement of circulating androgen excess

Testosterone measurement in women presents difficult analytical conditions because concentrations are lower than those typically encountered in adult men. Laboratory methods differ in cross-reactivity, sensitivity, calibration and reference populations; these differences can produce apparent phenotype shifts without any biological change. The international guideline emphasizes validated mass spectrometry for total testosterone and appropriately estimated or measured free testosterone. Overtreatment or missed diagnoses may follow if an immunoassay value is interpreted as exact. Hormonal contraception further complicates interpretation by altering SHBG and gonadotropin regulation. Ideally, investigators record the analytical platform, time since treatment cessation when clinically appropriate, menstrual context and reference interval. For multicentre RWE, laboratory-standardization exercises are not a cosmetic step but a prerequisite for valid causal comparisons. Without them, a model may identify hospital-specific testing equipment rather than endocrine disease.

Cardiovascular risk without overstating event evidence

PCOS is associated with several recognized metabolic and cardiovascular risk factors, yet elevated risk-factor frequency should not be conflated with proven disease-specific effects on myocardial infarction or cardiovascular mortality in all age groups. Many observational studies follow relatively young cohorts for limited periods, and treatment of dyslipidaemia, obesity or diabetes can modify later event rates. A Cox model yielding a hazard ratio for cardiovascular events requires accurate event adjudication, proportional-hazards assessment, careful age control and sufficient follow-up. Selection of symptomatic patients into care may affect both baseline cardiovascular testing and outcome surveillance. Clinically, standardized risk-factor assessment is warranted, but research writing should separate established guidance from uncertain

long-term absolute risk. This distinction guards against both neglect and exaggerated prediction.

Pregnancy intentions as a time-varying clinical variable

Women with PCOS may change priorities between skin symptoms, menstrual control, infertility care and metabolic prevention. Pregnancy planning strongly affects which medications are acceptable, how menstrual outcomes are interpreted and whether follow-up is continuous. Observational datasets rarely capture these intentions reliably, which can produce apparent differences between treated cohorts that have little to do with pharmacological action. An emulated trial addressing medication persistence must distinguish cessation because of intolerance from cessation because conception is being attempted. Live birth, biochemical pregnancy and ovulation are different estimands; conflating them exaggerates treatment effects. Researchers should consider competing events and pregnancy-related censoring explicitly, rather than treating every termination of follow-up as statistically exchangeable. Shared decision-making should also recognize that not all patients prioritize fertility and that a reproductive diagnosis should not erase mental-health or metabolic concerns.

Comparative evidence for inositol products

Inositol formulations are frequently marketed through mechanistic claims involving insulin-related signalling, but plausibility does not guarantee a clinically meaningful or reproducible effect. A systematic review prepared for the international guideline evaluated the evidence and found uncertainty concerning efficacy across important PCOS outcomes. Products may vary in isomer composition, dose, quality and regulatory status, limiting the transferability of small trials. Observational studies of supplement use are especially prone to healthy-user effects and undocumented co-interventions, because users may simultaneously change diet or take other therapies. Reporting should distinguish study-grade interventions from commercially heterogeneous preparations and should not label all supplements equivalent. A trial should prespecify reproductive and metabolic endpoints, adherence verification, adverse events and clinically credible follow-up; marketing testimonials and within-person pre–post change are insufficient evidence of effectiveness [21].

Comparative effectiveness and the limits of propensity scores

Propensity scores summarize measured covariate information about treatment assignment, not biological similarity in every respect. If women with the most severe

menstrual disturbances preferentially receive one strategy and symptom severity is poorly captured, matched cohorts may still differ in important ways. Transparent reporting should include distributions before and after matching, numbers discarded, overlap, balance for continuous laboratory variables and follow-up after treatment initiation. Weighting can preserve more patients but creates instability when predicted probabilities approach zero or one. The analyst should identify positivity violations rather than mechanically truncate away clinically meaningful subgroups. Where prescribing depends on unmeasured fertility plans, practice-site preference or contraindications, an observational estimate may remain noncausal. Reporting null results with honest uncertainty is more valuable than presenting superficially precise hazard ratios from poorly aligned comparisons.

Transportability and algorithmic fairness

A model developed in a tertiary reproductive centre may overestimate disease severity when moved to primary care, where symptom prevalence, referral thresholds and access to hormone testing are different. Simple demographic adjustment cannot necessarily correct these shifts. Validation should examine calibration intercepts and slopes, decision thresholds, under-diagnosis among patients without overt obesity and performance across geographic or resource strata. Gender-related and reproductive health variables also require careful privacy governance: adding sensitive covariates should be justified by clinical value and consent, not merely by a small improvement in discrimination. Clinical usability depends on whether a model supplies actionable information at the point of care without increasing clinician workload or stigmatizing patients. Periodic reassessment is essential when clinical guidelines, laboratory methods or electronic coding practices change.

Pharmacovigilance beyond automated signal mining

AI-assisted adverse-event detection can organize text reports and identify unexpected medication–event combinations, but frequency of a reporting term is not an estimate of attributable harm. For example, gastrointestinal complaints may be common reasons for metformin discontinuation, yet an adverse-event record may omit dose titration, formulation, co-medications or intercurrent illness. Automated pharmacovigilance should preserve provenance, de-duplicate records and clearly indicate uncertainty in temporal association. Clinical reviewers must assess alternative explanations, potential pregnancy exposure and whether the event meets a case definition. Robust postmarketing studies can sometimes complement spontaneous reports with

identifiable denominators and comparators. A safety signal should generate a testable hypothesis and proportionate clinical response, not a headline that a medication has definitively caused a newly recognized syndrome.

Cross-disease extrapolation: methodological gains and biological hazards

Methodological tools developed in diabetes, cardiovascular medicine, oncology, infectious diseases and neurology offer PCOS investigators valuable examples of longitudinal endpoint definition, registry linkage, adaptive monitoring and treatment-benefit modelling. These domains nevertheless differ sharply in natural history, outcome incidence and acceptable treatment risk. A survival model appropriate for death in oncology may be inefficient when the primary PCOS outcome is symptom improvement over months. Antimicrobial resistance dynamics offer little direct biological analogy for ovarian androgen production, even when machine-learning techniques are shared. The transfer should therefore be methodological rather than evidentiary: importing an analytical approach does not imply that findings or treatments transport across diseases. PCOS-specific endpoints and reproductive context should anchor every study, with other disease areas used as examples of design discipline.

A feasible multicentre protocol for Uzbekistan

A pragmatic initial study could prospectively enroll consecutive adults evaluated for possible PCOS in selected gynaecology and endocrinology settings, documenting both confirmed diagnoses and exclusions. The protocol would record menstrual history, standardized clinical androgen features, laboratory test method, anthropometry, relevant medication use and established metabolic assessments. Investigators should preregister a limited number of outcomes and harmonize variable names before recruitment. A short clinician-completed data form, multilingual patient information and secure pseudonymization would be preferable to collecting excessive unvalidated variables. Missingness and assay differences should be reviewed during pilot enrolment, and the dataset should undergo an external quality audit before predictive modelling. The proposal does not imply that a national PCOS registry currently exists or that the Ministry has endorsed a particular protocol; these would require separate institutional authorization.

Implementation science and health system decision-making

Evidence-based innovations fail when they do not match clinical workflow, staffing and patient access. A national digital-health strategy relevant to PCOS should address the full pathway from first presentation to metabolic follow-up and fertility consultation. A validated laboratory result unavailable to a rural patient does not provide equitable diagnostic benefit. Electronic decision support should be piloted with both specialists and primary-care clinicians, measuring time spent, unnecessary repeat testing, referral patterns, adverse events and patient experience. Procurement criteria ought to include data portability and algorithmic auditing, not merely technical demonstration performance. Training should teach clinicians to distinguish prediction from causal evidence and to interpret missing or unreliable test results. These implementation outcomes are often more consequential than small gains in model discrimination.

Research agenda: high-priority unresolved questions

Several empirical questions remain open: whether early insulin hypersecretion identifies a reproducible causal subgroup; whether androgen lowering can independently improve insulin sensitivity beyond changes in adiposity; which repeated endocrine measures predict meaningful metabolic outcomes; and whether phenotype-tailored treatment exceeds guideline-concordant usual care. High-quality answers will require longitudinal time ordering, analytical standardization, experimental triangulation and sufficiently representative recruitment. Trial emulation can support but not replace randomized intervention where equipoise persists. Investigators should commit to publishing protocols and null findings, reporting effect-size uncertainty and making non-identifying data dictionaries available. Success should be assessed by improvements in patient outcomes and equity, not the number of biomarkers or algorithms produced.

References

1. Teede HJ, Tay CT, Laven JJE, et al. Recommendations From the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. *J Clin Endocrinol Metab.* 2023;108(10):2447-2469. doi:10.1210/clinem/dgad463.
2. Teede HJ, Tay CT, Laven JJE, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. *Hum Reprod.* 2023;38(9):1655-1679. doi:10.1093/humrep/dead156. [Parallel publication of reference 1; not independent evidence.]

3. Reappraising the relationship between hyperinsulinemia and insulin resistance in PCOS. PubMed record PMID:40013621. 2025. Available from: <https://pubmed.ncbi.nlm.nih.gov/40013621/> . [Full author and journal details require final metadata export.]
4. Interface between reproductive and metabolic dysfunction in polycystic ovary syndrome. PubMed record PMID:40299715. 2025. Available from: <https://pubmed.ncbi.nlm.nih.gov/40299715/> . [Full author and journal details require final metadata export.]
5. Melin J, Forslund M, Alesi S, et al. The impact of metformin with or without lifestyle modification versus placebo on polycystic ovary syndrome: a systematic review and meta-analysis of randomized controlled trials. Eur J Endocrinol. 2023;189(2):S37-S63. doi:10.1093/ejendo/lvad098.
6. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Clin Endocrinol (Oxf). 2018;89(3):251-268. doi:10.1111/cen.13795.
7. United States Food and Drug Administration. Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products. Guidance for Industry. 2023. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug>
8. United States Food and Drug Administration. Real-World Evidence: Considerations Regarding Non-Interventional Studies for Drug and Biological Products. Draft guidance. 2024. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-evidence-considerations-regarding-non-interventional-studies-drug-and-biological-products>
9. United States Food and Drug Administration. Data Standards for Drug and Biological Product Submissions Containing Real-World Data. Guidance for Industry. 2023. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/data-standards-drug-and-biological-product-submissions-containing-real-world-data>
10. United States Food and Drug Administration. Real-World Data: Assessing Registries To Support Regulatory Decision-Making for Drug and Biological Products. Guidance. 2023. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-registries-to-support-regulatory-decision-making-for-drug-and-biological-products>

fda-guidance-documents/real-world-data-assessing-registries-support-regulatory-decision-making-drug-and-biological-products

11. United States Food and Drug Administration. Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products. 2022. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/submitting-documents-using-real-world-data-and-real-world-evidence-fda-drug-and-biological-products>

12. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021. ISBN:9789240029200.

13. World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Geneva: WHO; 2024. ISBN:9789240084759.

14. World Health Organization. Artificial Intelligence for Health. Geneva: WHO; 2024. Available from: <https://www.who.int/publications/m/item/artificial-intelligence-for-health>

15. United States Food and Drug Administration. Artificial Intelligence-Enabled Medical Devices. 2026. Available from: <https://www.fda.gov/medical-devices/digital-health-center-excellence/artificial-intelligence-enabled-medical-devices>

16. United States Food and Drug Administration. List of Artificial Intelligence-Enabled Medical Devices. 2026. Available from: <https://www.fda.gov/medical-devices/artificial-intelligence-enabled-medical-devices/list-artificial-intelligence-enabled-medical-devices>

17. United States Food and Drug Administration. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. 2025. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices>

18. World Health Organization. Regulatory considerations on artificial intelligence for health. WHO; 2023. Available from: <https://www.who.int/teams/digital-health-and-innovation/harnessing-artificial-intelligence-for-health>

19. Lord JM, Flight IH, Norman RJ. Metformin in polycystic ovary syndrome: systematic review and meta-analysis. BMJ. 2003;327:951. doi:10.1136/bmj.327.7421.951.

20. An endocrinological perspective on polycystic ovarian syndrome. PubMed record PMID:40147712. 2025. Available from: <https://pubmed.ncbi.nlm.nih.gov/40147712/> . [Metadata pending].
21. European Medicines Agency. Data Analysis and Real World Interrogation Network (DARWIN EU). 2026. Available from: <https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence/data-analysis-real-world-interrogation-network-darwin-eu>
22. Inositol for Polycystic Ovary Syndrome: A Systematic Review and Meta-analysis to Inform the 2023 Update of the International Evidence-based PCOS Guidelines. J Clin Endocrinol Metab. 2024. doi:10.1210/clinem/dgad762. PMID:38163998.
23. European Medicines Agency. Data quality framework for medicines regulation: application to real-world data. 2026. Available from: <https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/data-quality-framework-medicines-regulation>
24. European Medicines Agency. Real-world evidence: regulatory framework and guidance. 2025. Available from: <https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence>
25. European Medicines Agency. DARWIN EU continues expanding its capacity to deliver real-world data studies. 2024. Available from: <https://www.ema.europa.eu/en/news/darwin-eur-continues-expanding-its-capacity-deliver-real-world-data-studies>
26. European Medicines Agency. Data Analysis and Real World Interrogation Network. HMA-EMA Catalogues. Available from: <https://catalogues.ema.europa.eu/network/49628>
27. Comparison of dietary and physical activity behaviors in women with and without polycystic ovary syndrome: a systematic review and meta-analysis of 39 471 women. Hum Reprod Update. 2022. PMID:35639552. Available from: <https://pubmed.ncbi.nlm.nih.gov/35639552/>
28. World Health Organization. Ethics and governance of artificial intelligence for health: executive summary. 2021. ISBN:9789240037403. Available from: <https://www.who.int/publications/i/item/9789240037403>
29. World Health Organization. WHO consultation towards the development of guidance on ethics and governance of artificial intelligence for health. 2021.

30. European Medicines Agency. DARWIN EU Advisory Board meeting, 11 December 2023. Meeting records. 2024. Available from: <https://www.ema.europa.eu/en/events/darwin-eu-advisory-board-meeting-11-december-2023>

31. European Medicines Agency. Data analytics and methods: EMA Annual Report 2022. Available from: <https://www.ema.europa.eu/assets/en/annual-report/2022/data-analytics-and-methods.html>

7. Evidence certainty, reporting transparency and final translational priorities

Evidence certainty and multiplicity

Scientific caution is especially important in the PCOS literature because the same clinical population can yield numerous correlated outcomes: fasting insulin, different indices of insulin sensitivity, total testosterone, free androgen index, cycle frequency, body weight and ovarian morphology. Testing each outcome without a prespecified hierarchy increases the probability of apparently persuasive chance findings. In a review, these results should be contextualized by study design, outcome priority, whether treatment comparisons were randomized, and whether primary analyses were registered. When reports provide only within-group improvement rather than between-group treatment contrasts, a causal conclusion is inappropriate. Likewise, statistical significance cannot repair measurement bias or make a small laboratory difference clinically valuable. Future PCOS trials should separate primary endpoints from exploratory pathway analyses, report adverse outcomes comprehensively and avoid treating post-hoc phenotypic subgroups as definitive responders. Evidence synthesis should preserve the uncertainty of underlying research instead of converting multiple small studies into an overconfident consensus.

Clinical meaning and patient-level communication

An integrated metabolic–reproductive model has practical value only when clinicians can explain it accurately without creating deterministic expectations. A patient with PCOS should not be told that every androgen abnormality results from excess insulin or that body weight alone accounts for her symptoms. The relative contributions of gonadotropins, androgen production, binding proteins, adipose function and genetic predisposition vary. Clinical conversations ought to distinguish risk from diagnosis and

clarify why individual therapies target particular manifestations. For example, cycle control is not a validated substitute for metabolic screening, and improved glucose-related markers cannot be presumed to resolve infertility. Effective care requires reproductive goal assessment, cardiometabolic review, emotional wellbeing considerations and follow-up adapted to access and preferences. This multidimensional approach is consistent with the international guideline emphasis on shared decision-making and avoidance of weight stigma [1,2].

Data infrastructure and reproducible science

Any proposed national or international PCOS analytic network needs a reproducibility architecture. Protocols should define source populations, dates, ascertainment logic, exposure windows, outcome algorithms, clinically meaningful estimands and precommitted sensitivity analyses. Secure data-linkage rules and shared semantic definitions are more important than collecting unstructured records at scale. Distributed analysis, as illustrated by European regulatory evidence infrastructures, can permit harmonized methods while limiting unnecessary movement of identifiable information [21,23]. Independent validation teams should review coding accuracy and endocrine assay comparability before analyses are interpreted. Algorithm developers should release model cards, evaluation plans and documentation of relevant clinical constraints where law and privacy rules allow. Publication of negative validation results would reduce waste and counter selective reporting. These processes create the foundation for reliable evidence rather than simply adding technical sophistication.