BMC Medicine

RESEARCH

Open Access

The PCOS Phenotypes in Unselected Populations (P-PUP) study: participant clinical features and data harmonization on analysis of individual participant data

Asmamaw Demis Bizuneh[1], Sylvia Kiconco[1], Arul Earnest[2], Mahnaz Bahri Khomami[1], Raja Ram Dhungana[1], Ricardo Azziz[3], Larisa V. Suturina[4], Xiaomiao Zhao[5], Alessandra Gambineri[6], Fahimeh Ramezani Tehrani[7], Bulent O. Yildiz[8], Jin Ju Kim[9], Liangzhi Xu[10], Christian Chigozie Makwe[11], Helena J. Teede[1,12], Anju E. Joham[1,12†] and Chau Thien Tay[1,12*†]

Abstract

Background Polycystic ovary syndrome (PCOS) is a multifaceted condition with diagnostic challenges and clinical heterogeneity across populations. Research priorities include enhanced accuracy in defining cut-offs for diagnostic features. Here, we aim to describe participant clinical features and data harmonization in the international PCOS Phenotype in Unselected Populations (P-PUP) study.

Methods We searched EMBASE and Medline (Ovid) from 1990 to October 2, 2020, in population-based, medically unbiased study cohorts. Included studies had ≥ 300 participants, directly assessed PCOS-related features, and provided Individual Participant Data (IPD). Risk of bias was assessed using the AXIS tool. Data integrity was ensured via cross-referencing, identifying outliers/implausible data, and variable harmonization. Reporting follows PRISMA-IPD guidelines, summarizing findings with frequencies and proportions.

Results The study included 9979 reproductive-age women from 12 studies across eight countries (China, Iran, Italy, Nigeria, Russia, South Korea, Turkey, and the USA), representing 11 ethnicities. Ovulatory dysfunction was variably recorded, from mean menstrual cycle length, minimum or maximum cycle length, number of cycles per year, or urinary progesterone measurements. Clinical hyperandrogenism was assessed via modified Ferriman–Gallwey (mFG) scores, with a few also including acne and alopecia. Biochemical hyperandrogenism thresholds varied (95th, 97.5th, or 98th percentile of healthy controls). Polycystic ovary morphology was assessed via transvaginal, transabdominal, or transrectal approaches. Harmonization adhered to International PCOS Guidelines for ovulatory dysfunction, ethnicity-specific cut-offs for hirsutism (via *k*-means clustering), and 95th percentile thresholds for biochemical hyperandrogenism. PCOS prevalence ranged from 3.3 to 19.8% in the original studies and was 11.0% overall after harmonization.

†Anju E. Joham and Chau Thien Tay are equal contributors and have joint last authorship.

*Correspondence:
Chau Thien Tay
jillian.tay@monash.edu
Full list of author information is available at the end of the article

Conclusions The P-PUP study offers an unprecedented, ethnically diverse, medically unbiased population-based cohort, an extraordinarily valuable tool to enhance knowledge and research in PCOS. However, variability in data collection methods and definitions of PCOS diagnostic features across studies limited the ability to fully integrate data for analysis. Despite these limitations, we optimized harmonization in this IPD, and the findings provided valuable insights into the challenges of data harmonization and established a foundation for future collaborative research. Future research should focus on standardizing data collection, establishing normative cut-offs based on true natural groupings, and linking diagnostic clusters to outcomes in diverse populations.

Protocol registration CRD42021267847.

Keywords Data harmonization, Androgens, Polycystic ovary syndrome, PCOS, Hirsutism, Acne, Female pattern hair loss, Polycystic ovary, Phenotype, Individual participant data, IPD

Background

Polycystic ovary syndrome (PCOS) is a complex endocrinopathy with significant long-term reproductive, metabolic, dermatological, and psychological health impacts [1]. According to the 2023 International PCOS Guidelines criteria, diagnosis in adults is based on the presence of at least two of the following three criteria: (1) oligo-/anovulation (OA); (2) clinical or biochemical hyperandrogenism (HA); and (3) either polycystic ovary morphology (PCOM) on ultrasound or elevated serum anti-Müllerian hormone (AMH) levels [2–5]. In adolescents, the diagnosis requires both OA and HA [2–5]. It is estimated that 10–13% of women worldwide are affected by PCOS, and the global incidence is reportedly increasing, making PCOS a major public health concern [2–6]. PCOS and its associated health implications generate substantive health [7, 8] and economic burdens for individuals, societies, and governments [9–12].

Despite increasing research and publications on PCOS [13], the pathogenesis remains unclear and there are significant persisting challenges in diagnosis. This is attributable to its multifaceted and heterogeneous clinical presentation, the lack of gold-standard diagnostic tests, ethnic variation, and the fact that most studies are carried out in medically biased (clinically based) populations [14–19]. It is also due to the multiple components of the diagnostic criteria, which are all continuous variables [20–22], the paucity of data on normative cut-off thresholds for each diagnostic component within unselected, diverse populations [23, 24], and the use of arbitrary diagnostic cut-offs across studies [25–31]. This contributes to delayed or inaccurate diagnosis of PCOS, hindering timely and appropriate interventions to limit complications and treat the condition [12, 32–35].

These challenges underscore the importance of consolidating data from diverse populations for a more comprehensive understanding of the condition [16]. While traditional systematic reviews and meta-analyses represent an improvement over single-center studies, they are limited by aggregated data from original studies with varying methodologies, sample sizes, and study populations. Conversely, the PCOS Phenotype in Unselected Populations (P-PUP) study, is a large-scale, international individual participant data (IPD) meta-analysis designed to address identified challenges, allowing harmonization of raw data from individual studies, standardization of cut-offs and analyses across studies and examination of individual-level variables to advance knowledge on PCOS etiology, diagnosis and clinical features in unselected community-based populations [36].

This report on the P-PUP study aims to (a) outline the selection criteria and study identification process for the study, (b) describe the clinical features of the populations included, and (c) detail the data harmonization process for PCOS diagnostic criteria and features within the study.

Methods

Protocol and registration

The protocol for the P-PUP study was registered in PROSPERO (ID: CRD42021267847) and published previously [36]. This study was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses of Individual Participant Data guideline [37].

Selection of studies

The methodology for the selection and identification of research groups and collaborators has been described previously [36]. In summary, a systematic search of the EMBASE and Medline (Ovid) databases, along with a manual Google search, was conducted from 1990 to October 2, 2020, to identify eligible unselected population-based cohort studies. Key search terms included PCOS, ovulatory dysfunction (OD), hirsutism, polycystic ovary morphology (PCOM), and unselected populations, among others. A full search strategy for Medline (Ovid) is in Additional file 1: Table S1. The PCOS research

community (RA and HJT) was consulted to identify any potentially ongoing or unpublished studies.

Cross-sectional or longitudinal cohort studies were considered eligible if they were published after 1990 (after a change in diagnostic criteria), included a sample size of 300 or more unselected women of any age, ethnicity, or weight, and directly assessed features related to PCOS [36]. These features of PCOS included measurement of ovulatory dysfunction, clinical or biochemical hyperandrogenism, and PCOM. No language restrictions were imposed. Studies that relied on self-reported or International Classification of Diseases (ICD) codes for PCOS status without the assessment of PCOS features were excluded. Exclusion criteria at the participant level included pregnant and lactating women, women with natural/surgical menopause, ovarian pathology, and women meeting the criteria for exclusion of other conditions in the diagnosis of PCOS. Additionally, women receiving contraceptive hormonal treatment within 3 months of assessment were excluded from the androgen analysis.

Titles and abstracts were initially screened by a single reviewer (SK), and then full-text articles were retrieved and assessed in duplicate (SK and ADB) with any discrepancies resolved by a third reviewer (HJT). The corresponding and/or lead authors of the identified studies were contacted and requested to contribute to deidentified IPD for integration and analysis. For authors who did not respond, two reminder emails were sent.

Data collection processes

Data and relevant dictionaries were transferred to the ISO 27001-certified Secure File Transfer Protocol (SFTP) [38] which restricts access and control to the authorized data custodian and researchers. The key variables collected aligned with the PCOS core outcomes set [39] and included sociodemographic factors (age and race/ethnicity); generic outcomes (weight, height, body mass index (BMI)), metabolic outcomes (waist circumference (WC), hip circumference (HC), waist-to-hip ratio, homeostasis model assessment for insulin resistance (HOMA-IR), lipid profiles, blood pressure, and fasting blood glucose (FBG)); dermatological (hirsutism as measured by modified Ferriman-Gallwey (mFG score) and reproductive outcomes (androgen assays, menstrual cycle length, number of menstrual cycles per year, age of menarche, and PCOM features including antral follicle count (AFC) and ovarian volume (OV)). Any data discrepancies or concerns were discussed and resolved with the original study investigators.

Risk of bias assessment

We evaluated the risk of bias in the included studies providing IPD using the AXIS appraisal tool for cross-sectional studies [40]. The appraisal was conducted by two independent reviewers (ADB and RRD), and any discrepancies were resolved through consensus. This tool consists of 20 items that assess various aspects such as the study's aim, design, sample size, outcome variable measurement, statistical methods, response rate, result consistency, discussion and conclusion, limitations, and ethical approval, with responses categorized as "yes," "no," or "don't know" [41]. Overall quality was rated as low (score 1-7), medium (score 8-14), or high (score 15-20) [41].

Data harmonization and analysis

Data synthesis involved rigorous procedures to ensure data integrity and harmonization across diverse datasets. All laboratory measurements were standardized to the International System of Units (SI) for consistency. Outliers were identified using scatter plots and box plots, applying a criterion where values exceeding > 5 or > 10 standard deviations were flagged [42]. Invalid or implausible data points, such as diastolic blood pressure (DBP) surpassing systolic blood pressure (SBP), and height < 120 cm were identified, based on consensus clinical judgment and predefined criteria. Definitions of outliers, invalid data points, and rules for variable harmonization were documented and collaboratively discussed with original study investigators to achieve consensus on data handling (Additional File 1: Table S4).

For the harmonized definition of PCOS, we adhered to the International PCOS Guideline criteria, requiring two of three key features: OD, clinical or biochemical hyperandrogenism, and PCOM. The specific criteria for each feature were defined as follows:

- OD: Menstrual cycles numbering fewer than eight or exceeding 17 per year, and/or cycle lengths shorter than 21 days or longer than 35 days.
- Clinical hyperandrogenism: Defined by hirsutism quantified through the modified Ferriman-Gallwey (mFG) scoring system. We analyzed ethnic distribution patterns of mFG scores using descriptive statistics of means and medians. Subsequently, we standardized the data (mean = 0, SD = 1) and performed a *k*-means cluster analysis ($k = 2$), with cluster determination guided by the average silhouette method, following our previously published methodology [43].
- Biochemical hyperandrogenism: Defined using study-specific 95th percentile cut-offs for first-line androgen tests (TT, FT, cFT, and FAI), supplemented by secondary markers (A4 and DHEAS) when first-line tests yielded negative results. This approach

accounts for variations in assay methods and reference ranges across studies.

- Polycystic ovarian morphology (PCOM): Given the considerable heterogeneity in imaging modalities, ultrasound approaches, and absence of individual participant data (IPD), harmonization of PCOM data proved unfeasible; thus, we relied on each study's original definitions and categorizations for this criterion.

This study undertook a one-stage approach for analysis and only included women of reproductive age between 18 and 45 years. Continuous variables were reported as means with standard deviations (SD) or medians with interquartile ranges (IQR), as appropriate for the distribution of data. Categorical variables were reported as frequencies and proportions. Available case analysis was employed, and we did not opt to undertake multiple imputations for any missing data because they could not be random. Stata 18 software was used for data management and analysis.

Results

Study characteristics and risk of bias

The literature search identified 6507 publications, from which 1437 duplicates were removed. After screening the titles and abstracts, 342 studies underwent full-text review of which 313 studies were excluded (Additional File 1: Table S2). Finally, 29 studies met the eligibility criteria and their authors were contacted. Of these, the authors of 19 studies, encompassing 31,580 participants, did not respond. Of the 10 authors who responded, seven expressed willingness to collaborate and completed the data-sharing process, whereas three studies involving a combined total of 17,439 participants were unable to do so [44–46]. Additionally, two ongoing studies were identified through other researchers, and two authors contributed more than one study from different populations. The entire P-PUP study dataset encompassed 12 population-based medically unbiased representative cohorts [26, 31, 47–56] and 12,513 participants drawn from PCOS research groups spanning eight countries internationally (Fig. 1). A detailed assessment of the included studies' methodological quality appraisal is presented in Additional File 1: Table S3. All of the included studies had a low risk of bias (high quality) [26, 31, 47–56].

Table 1 presents the characteristics of the final included studies. The entire P-PUP study dataset included studies from China ($n=2$; 4645 women) [48, 52], Iran ($n=3$; 2517 women) [49, 53, 54], Italy ($n=1$; 519 women) [50], Russia ($n=1$; 2695 women), South Korea ($n=1$; 499 women) [26], Turkey ($n=1$; 392 women) [51], the USA ($n=2$; 806 women) [55, 56], and Nigeria ($n=1$; 440 women) [47]. All studies were cross-sectional in design with sample sizes ranging from 392 to 3000.

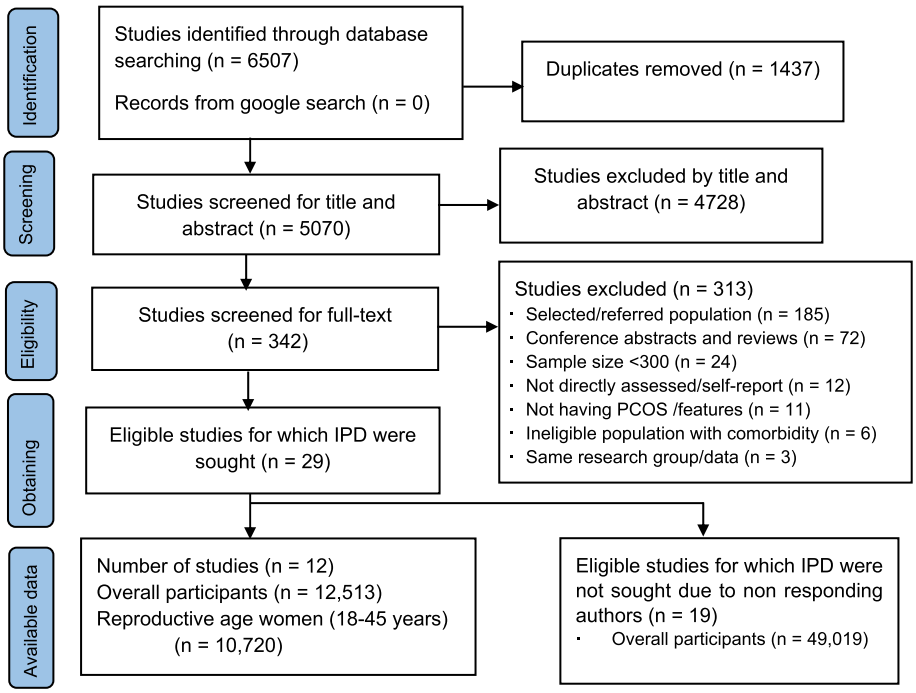

Fig. 1 PRISMA IPD flow diagram for P-PUP study selection

Table 1 Characteristics of the included studies

Studies	Country	Study design	Studies sample size	Sample size (18–45 years)	Inclusion criteria	Exclusion criteria	Ovulatory dysfunction definition	Clinical hyperandrogenism	Biochemical hyperandrogenism definition for selected healthy controls	Assay method	PCOM definition (approach and frequency)	Quality score
Zhao et al. 2011 [52]	China	Cross-sectional	2988	3000	Women aged 19–45 years	Pregnant women	<8 cycle/year or cycle length >35 days	mFG ≥8 or acne	Above 95th centile; TT: 3.08 nmol/L, FAI: 6.74, A4: 17.67 nmol/L.	CLIA	≥12 follicles (TVUS: 6–9 or 6–13 MHz and TAUS: 5–6 or 3–7 MHz)	18
Zhuang et al. 2014 [48]	China	Cross-sectional	1645	1369	Women aged 18–44.	Women who were not residents in Chengdu, girls or women who in the previous 3 months took gonadal hormones or drugs which could affect ovarian function.	Menstrual cycle ≥35 days and for more than 3 months in a year ovulation is absent	mFG ≥6 with or without acne	Two standard deviations above the normal (above 97.7th percentile) for FT >3.2 pg/mL (>0.0111 nmol/L).	RIA	≥12 follicles and/or ≥10 cm^3 per ovary (TVUS: 7 MHz and TAUS: 3.5 MHz)	18
Tehrani et al. 2011 [54]	Iran	Cross-sectional	1000	1000	Women aged 18–45 years	Pregnant, menopausal, and those who had undergone hysterectomy or bilateral oophorectomy	Cycle length >35 days or P4 <4 ng/ml	mFG ≥8, acne, or alopecia	Above 95th centile; TT: 0.89 ng/mL (3.09 nmol/L), FAI: 5.39, DHEAS: 179 µg/dL (4.85 µmol/L), A4: 2.9 ng/mL (10.12 nmol/L).	EIA	N/A	19
Tehrani et al. 2011 [53]	Iran	Cross-sectional	915	915	Women aged 18–45 years	Pregnant, menopausal, and those who had undergone hysterectomy or bilateral oophorectomy	Cycle length >35 days.	mFG ≥8	Above 95th centile; TT: 0.88 ng/mL (3.05 nmol/L), FAI: 5.47, DHEAS: 246 µg/dL (6.67 µmol/L), A4: 2.3 ng/mL (8.03 nmol/L).	EIA	≥12 follicles and/or ≥10 cm^3 per ovary (TVUS: 5 MHz and TAUS: 3.5 MHz)	19

Table 1 (continued)

Studies	Country	Study design	Studies sample size	Sample size (18–45 years)	Inclusion criteria	Exclusion criteria	Ovulatory dysfunction definition	Clinical hyperandrogenism	Biochemical hyperandrogenism definition for selected healthy controls	Assay method	PCOM definition (approach and frequency)	Quality score
Tehrani et al. 2014 [49]	Iran	Cross-sectional	602	602	Women aged 18–45 years	Pregnant, menopausal, and those who had undergone hysterectomy or bilateral oophorectomy	Amenorrhea, cycle length <21 days or >35 days	mFG ≥8	Above 95th centile; TT: 0.89 ng/mL (3.09 nmol/L), FAI: 5.39, DHEAS: 179 µg/dL (4.85 µmol/L), A4: 2.9 ng/mL (10.12 nmol/L).	EIA	≥12 follicles and/or ≥10 cm^3 per ovary (TVUS: 5 MHz and TAUS: 3.5 MHz)	19
Gambineri et al. 2013 [50]	Italy	Cross-sectional	519	186	Students aged 18–19 years	Those who refused to participate in the study	>6 cycles with length >35 days/year or lack of menses for 3 consecutive months	mFG ≥8 or alopecia	Above 97.5th percentile of the reference interval for TT: 0.45 ng/mL (1.57 nmol/L)	LC-MS/MS	N/A	18
Makwe et al. 2023 [47]	Nigeria	Cross-sectional	440	440	Women aged 18–45 years	Women < 18 or > 45 years, pregnant, lactating, postmenopausal, and those who had undergone hysterectomy and/or bilateral oophorectomy.	Cycle length <26 days or >35 days	mFG ≥6	N/A	N/A	≥20 follicles and/or ≥10 cm^3 per ovary (TVUS: 4–10 (6.7 and 8.5) MHz and TAUS: 1–5 MHz)	20

Table 1 (continued)

Studies	Country	Study design	Studies sample size	Sample size (18–45 years)	Inclusion criteria	Exclusion criteria	Ovulatory dysfunction definition	Clinical hyperandrogenism	Biochemical hyperandrogenism definition for selected healthy controls	Assay method	PCOM definition (approach and frequency)	Quality score
Suturina et al. 2022 [31]	Russia	Cross-sectional	1134	1550	Women aged 18–45 years who provided written informed consent	Current pregnancy or lactation, history of hysterectomy and/or bilateral oophorectomy, endometrial ablation, and/or uterine artery embolization and non-compliance with all study procedures and requirements	<21 or >35 days or <8 cycles per year	mFG ≥5	Above 98th centile; TT: 73.9 ng/dL (2.56 nmol/L) for White Russians) and 41 ng/dL (1.42 nmol/L) for Asian Russians and Mixed Russians, FAI: 6.9 for White Russians and 2.9 for Asian Russians and Mixed Russians, DHEAS: 355 µg/dL (9.62 µmol/L) for all Russians.	LC-MS/MS	≥12 follicles and/or ≥10 cm^3 per ovary (TVUS: 5–8 MHz and TAUS: 2.5–5 MHz)	17
Kim et al. 2011 [26]	South Korea	Cross-sectional	499	499	Women aged 18–40 years	Pregnant women	<8 cycles/year or cycle length >35 days or amenorrhea	mFG ≥6	Above 95th centile; TT: 0.68 ng/mL (2.36 nmol/L), FAI: 5.36.	RIA	≥12 follicles and/or ≥10 cm^3 per ovary (TVUS and TAUS:5–9 MHz)	17
Yildiz et al. 2012 [51]	Turkey	Cross-sectional	392	389	Women aged 18–45 years	Women <18 or >45 years, pregnant, postmenopausal, and those who had undergone hysterectomy or bilateral oophorectomy	Cycle length ≤23 days or ≥35 days or P4 ≤5 ng/ml	mFG ≥6	Above 95th centile; TT: 54.7 ng/dL (1.9 nmol/L), FAI: 4.94, DHEAS: 3257.4 ng/mL (8.83 µmol/L), A4: 2.97 ng/mL (10.37 nmol/L)	CLIA, RIA	≥12 follicles and/or ≥10 cm^3per ovary (TVUS: 5–9 and TAUS: 2–7 MHz)	17

Table 1 (continued)

Studies	Country	Study design	Studies sample size	Sample size (18–45 years)	Inclusion criteria	Exclusion criteria	Ovulatory dysfunction definition	Clinical hyperandrogenism	Biochemical hyperandrogenism definition for selected healthy controls	Assay method	PCOM definition (approach and frequency)	Quality score
Knochenhauer et al. 1998 [56]	USA	Cross-sectional	411	383	Women aged 18–45 years	Oophorectomised or those receiving hormonal treatment were excluded for circulating androgen levels.	≤8 cycles/year or cycle >35 days	mFG ≥6	Above 95th centile; TT: 84.7 ng/dl (2.94 nmol/L), cFT: 0.75 ng/dL (0.026 nmol/L), DHEAS: 2459 ng/ml (6.64 µmol/L), A4: 2496 pg/ml (8.73 nmol/L)	RIA, RIA after extraction	N/A	17
Azziz et al. 2004 [55]	USA	Cross-sectional	395	388	Women aged 18–45 years	Women <18 or >45 years, pregnant, menopausal, and those who had undergone hysterectomy or bilateral oophorectomy	≤8 cycles/year or cycle length <26 or >35 days or day 22–24 P4 <4 ng/ml	mFG ≥6	Above 95th centile; TT: 84.7 ng/dl (2.94 nmol/L), cFT: 0.75 ng/dL (0.026 nmol/L), DHEAS: 2459 ng/ml (6.64 µmol/L), A4: 2496 pg/ml (8.73 nmol/L)	RIA, RIA after extraction	N/A	17

The reported laboratory values were converted to the SI unit using the following conversion factors: for A4 3.49 from ng/mL to nmol/L and 0.00349 from pg/mL to nmol/L; for TT 3.47 from ng/mL to nmol/L and 0.0347 from ng/dL to nmol/L; for FT 0.003467 from pg/mL to nmol/L and 0.03467 from ng/dL to nmol/L; for DHEAS 0.00271 from ng/mL to µmol/L and 0.0271 µg/dL to µmol/L, N/A: Data not available, Quality was assessed using the AXIS tool and rated as high: 15–20, medium: 8–14, and low: 1–7

A4 androstenedione, *FAI* free androgen index, *FT* free testosterone, *mFG score* modified Ferriman-Gallwey score, *TT* total testosterone, *DHEAS* dehydroepiandrosterone sulfate, *P4* progesterone, *CLIA* chemiluminescence immunoassay, *EIA* enzyme immunoassay, *LC-MS/MS* liquid chromatography-tandem mass spectrometry, *RIA* radioimmunoassay, *TVAS* transvaginal ultrasound, *TAUS* transabdominal ultrasound

Participant selection after data harmonization

Of the total 12,513 participants in the P-PUP study, 2046 were excluded due to missing data on age ($n=6$), age < 18 years ($n=609$), age > 45 ($n=1178$), natural/surgical menopause ($n=81$), current pregnancy/lactation ($n=40$), and ovarian pathology ($n=48$). Additionally, 488 participants were excluded due to abnormal follicle-stimulating hormone (FSH > 25 IU/L), prolactin (> 67.2 µg/L), 17-hydroxyprogesterone (17-OHP > 30 nmol/L), or thyroid-stimulating hormone (TSH > 5 µU/mL) (Fig. 2, Additional File 1: Table S4). A total number of 9979 participants was included in the final analysis following data harmonization. The predominant ethnicities represented in reproductive-age women were 40.1% Asian Han Chinese, 23.8% White Iranians, and 9.5% White Russians (Fig. 3).

Assessment of ovulatory dysfunction

All studies collected information on ovulatory dysfunction (OD) through the assessment of menstrual cycles, but the data collection methods were heterogeneous. Four studies collected data about OD in categorical form [48, 52, 55, 56], with differing cut-offs: fewer than 21 days, 21–35 days, or greater than 35 days [52]; 35 days or less or greater than 35 days [56]; less than 26 days and greater than 35 days [55]; and another reported it as a binary with regular (21–35 days) or irregular (fewer than 21 days, greater than 35 days or irregular bleeding) [48].

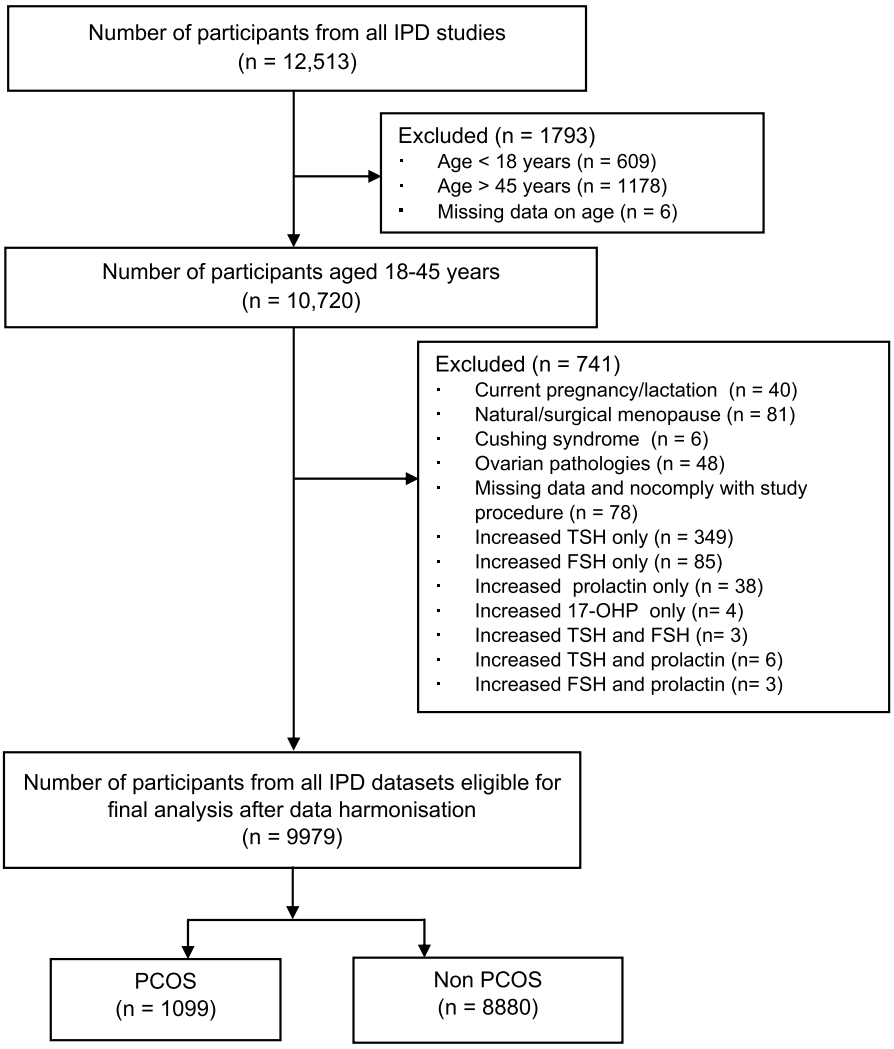

Fig. 2 Flow chart for study participant selection. Note: TSH, prolactin, and 17-OHP were not consistently collected across the total populations in the studies, as some studies collect the data on women suspected of having PCOS

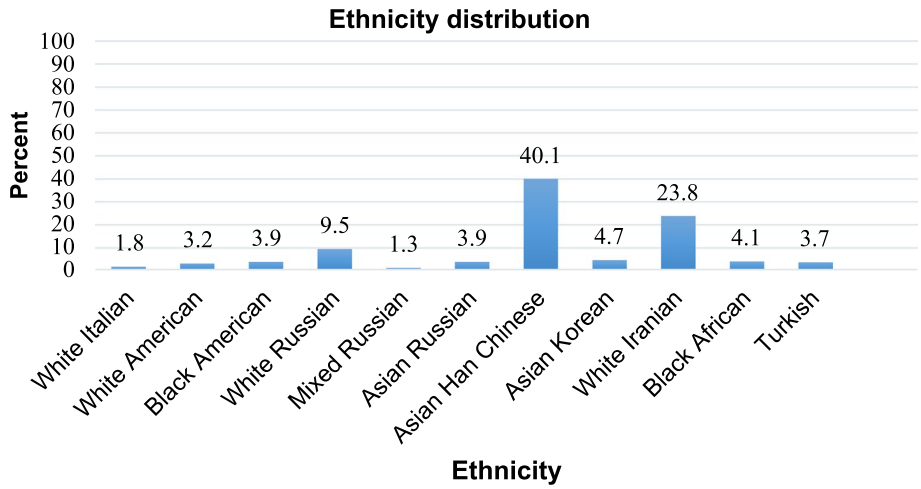

Fig. 3 Ethnicity distribution of the study participants

Six studies collected actual menstrual cycle length [26, 31, 47, 49, 53, 54], seven studies collected the number of menstrual cycles per year [31, 47, 49–51, 53, 54], and one study recorded the minimum and maximum menstrual cycle lengths, in addition to the overall menstrual cycle length [31]. One study also reported minimum and maximum cycle lengths and cycle numbers per year in addition to the categorical menstrual cycle status, with nearly 60% missing data [48]. Except for one study [31] which had 4% (62/1550) missing data, the remaining studies provided complete data on OD (Additional file 1: Table S5).

The definitions of OD varied across studies and are summarized in Table 1. Most studies defined OD based on menstrual cycle length or the number of menstrual cycles per year, with some including progesterone levels to further assess ovulation status. OD was most commonly defined by using an upper limit of menstrual cycle length greater than 35 days [31, 47–49, 52–56], with another study using a threshold of greater than 34 days [51]. The lower limit for menstrual cycle length also varied, with three studies using less than 21 days [31, 48, 49, 53], one study using less than 24 days [51], one using less than 26 days [55], and another one using less than 25 days [47]. For the number of menstrual cycles per year, four studies defined OD as less than eight menstrual cycles per year [26, 31, 47, 52], while two studies used the cut-off of fewer than nine cycles per year [55, 56]. Additionally, three studies incorporated a lack of menses for three consecutive months as part of their OD definition [26, 49–51, 54, 55]. Mid-luteal progesterone level was utilized by three studies as part of their assessment of OD [51, 54, 55]. The reported prevalence of OD in the included studies ranged from 9.2 to 31.9%, with an overall aggregated prevalence of 18.2%. For data harmonization, we defined OD as fewer than eight or more than 17 menstrual cycles per year and/or a menstrual cycle length of less than 21 days or greater than 35 days, in alignment with the 2023 International PCOS Guidelines [2–5]. Based on this definition, the prevalence of OD across the included studies ranged from 3.7 to 31.9%, with an overall prevalence of 15.3% (Table 2).

Assessment of clinical hyperandrogenism

All included studies assessed clinical hyperandrogenism, using hirsutism as a primary criterion [26, 31, 47–56]. Hirsutism was assessed directly by a health professional, including gynecologists, midwives, physicians, nurses, endocrinologists, and general practitioners using the mFG score. Hirsutism was defined by an mFG score of 6 or higher in six studies [26, 47, 48, 51, 55, 56], 8 or higher in five studies [49, 50, 52–54], and 5 or higher in one study [31].

All studies reported acne [26, 31, 47–56], with nine providing IPD [26, 47, 49, 50, 52–56], and three incorporating acne into their definition of clinical hyperandrogenism [48, 52, 54]. One study employed the Burke-Cunliffe acne grading method, assessed by a physician [50]; six studies employed the standard acne lesion assessment of the face, back, and chest, assessed by gynecologists, midwives, physicians, and general practitioners [31, 47, 49, 52–54]; two studies mentioned that no specific acne scoring system was used, with assessments conducted by nurses [55, 56], while another two studies did not provide details on the acne scoring tools used, with assessments conducted by physicians [48, 51]. One study used self-reported acne without a specific scoring system [26]. Among studies that included acne

Table 2 Prevalence of ovulatory dysfunction

Studies	Prevalence			
	Original data *n* (%)		Harmonized data *n* (%)	
	Yes	No	Yes	No
Zhao et al. 2011 [52]	421 (14.0)	2579 (86.0)	519 (17.8)	2400 (82.2)
Zhuang et al. 2014 [48]	436 (31.9)	932 (68.1)	431 (31.9)	919 (68.1)
Tehrani et al. 2011a [53]	140 (14.0)	857 (86.0)	114 (11.7)	861 (88.3)
Tehrani et al. 2011b [54]	170 (18.6)	745 (81.4)	146 (19.2)	614 (80.8)
Tehrani et al. 2014 [49]	72 (12.0)	529 (88.0)	67 (13.4)	432 (86.6)
Gambineri et al. 2013[a] [50]	N/R	N/R	13 (7.1)	169 (92.9)
Makwe et al. 2023 [47]	67 (15.3)	371 (84.7)	15 (3.7)	394 (96.3)
Suturina et al. 2022 [31]	342 (24.4)	1061 (75.6)	81 (6.3)	1209 (93.7)
Kim et al. 2011 [26]	70 (14.1)	427 (85.9)	26 (6.2)	392 (93.8)
Yildiz et al. 2012 [51]	60 (15.4)	329 (84.6)	17 (4.4)	370 (95.6)
Knochenhauer et al. 1998 [56]	33 (9.2)	325 (90.8)	33 (9.2)	325 (90.8)
Azziz et al. 2004 [55]	75 (20.5)	291 (79.5)	52 (14.2)	313 (85.8)
Overall	1886 (18.2)	8446 (81.8)	1514 (15.3)	8398 (84.7)

N/R not reported

[a] very small sample size. For data harmonization, we defined OD as fewer than eight or more than 17 menstrual cycles per year and/or a menstrual cycle length of less than 21 days or greater than 35 days, in alignment with the 2023 International PCOS Guidelines

in their clinical hyperandrogenism definition, a score of two or higher was used in one study [52], six or higher in another [54], and the cut-off was not specified in one study [48].

Six studies reported on the presence of female pattern hair loss (FPHL) [31, 47, 50, 52, 54, 55], with four providing IPD [47, 50, 54, 55], and two incorporated FPHL into their definition of clinical hyperandrogenism [50, 54]. The presence of FPHL was assessed by gynecologists, physicians, nurses, and general practitioners using the Ludwig scale in three studies [31, 50, 52], while three other studies provided no details on the FPHL assessment tool used [47, 54, 55]. FPHL was reported in a categorized form in two studies with no cut-offs provided [50, 54].

For data harmonization, we defined clinical hyperandrogenism using hirsutism only in alignment with the 2023 International PCOS Guidelines which highlights hirsutism as the strongest predictor of hyperandrogenism. We used ethnicity-specific mFG score cut-offs derived from *k*-means ($k=2$) cluster analysis in our recent publication [43]. The thresholds were set at mFG score ≥ 8 for Iranians, ≥ 7 for White Italians and Black Africans, ≥ 5 for Asian Han Chinese, White Russians, Turkish, and Black Americans, and ≥ 4 for White Americans, Asian Koreans, Asian Russians, and Mixed Russians [43]. Based on these cut-offs, the prevalence of clinical hyperandrogenism ranged from 8.1 to 34.2% in individual included studies, with an overall aggregate prevalence of 14.2% (Table 3).

Assessment of biochemical hyperandrogenism

Biochemical hyperandrogenism was assessed in all included studies except for one study done in Nigeria [47]. Seven studies assessed all participants [31, 48–51, 53, 56], three studies [26, 54, 55] assessed selective subgroups, and the remaining one study collected blood samples from a random half of the population [52]. Among the studies that assessed androgen levels in selective subgroups, Tehrani et al. assessed only participants with either hirsutism (mFG score ≥ 8) or menstrual disorder (26% of the sample, 260/1000) [54]; Azziz et al. assessed only participants with menstrual dysfunction (48.7% of the sample, 189/388) [55]; and Kim et al. assessed participants who consented to provide blood samples (32.1% of the sample, 160/499) [26].

Of the included studies, 11 reported total testosterone (TT) [26, 31, 48–54, 56], seven reported calculated free androgen index (FAI) [26, 31, 49, 51–54], eight reported androstenedione (A4) [49–56], and nine reported dehydroepiandrosterone sulfate (DHEAS) [26, 31, 49–51, 53–56]. The assay method used included enzyme immunoassay (EIA) [49, 53, 54], liquid-chromatography tandem mass spectrometry (LC–MS/MS) [31, 50], chemiluminescent immunoassay (CLIA) [31, 51, 52], radioimmunoassay (RIA) after extraction [55, 56], and RIA [26, 48, 51, 55, 56] (Table 1). Three studies reported free testosterone (FT) [48, 55, 56]: two reported calculated FT (cFT), derived from TT measured by an in-house RIA after serum extraction and SHBG binding

Table 3 Prevalence of clinical hyperandrogenism

Studies	Prevalence			
	Original data *n* (%)		Harmonized data *n* (%)	
	Yes	No	Yes	No
Zhao et al. 2011 [52]	325 (10.8)	2675 (89.2)	304 (10.4)	2615 (89.6)
Zhuang et al. 2014 [48]	175 (12.8)	1192 (87.2)	205 (16.9)	1009 (83.1)
Tehrani et al. 2011a [53]	234 (23.5)	763 (76.5)	168 (17.2)	807 (82.8)
Tehrani et al. 2011b [54]	200 (21.9)	715 (78.1)	170 (22.4)	590 (77.6)
Tehrani et al. 2014 [49]	171 (28.4)	430 (71.6)	133 (26.7)	366 (73.3)
Gambineri et al. 2013[a] [50]	40 (21.5)	146 (78.5)	62 (34.2)	119 (65.8)
Makwe et al. 2023 [47]	43 (9.8)	395 (90.2)	28 (6.9)	377 (93.1)
Suturina et al. 2022 [31]	103 (7.3)	1298 (92.7)	105 (8.1)	1183 (91.6)
Kim et al. 2011 [26]	49 (9.8)	449 (90.2)	74 (15.3)	409 (84.7)
Yildiz et al. 2012 [51]	40 (10.3)	349 (89.7)	65 (16.8)	322 (83.2)
Knochenhauer et al. 1998 [56]	30 (8.4)	328 (91.6)	41 (11.9)	303 (88.1)
Azziz et al. 2004 [55]	25 (6.8)	341 (93.2)	41 (11.4)	318 (88.6)
Overall	1435 (13.7)	9081 (86.3)	1396 (14.2)	8418 (85.8)

[a] very small sample size. For data harmonization, we defined clinical hyperandrogenism using ethnicity-specific modified Ferriman-Gallwey (mFG) score cut-offs derived from *k*-means ($k = 2$) cluster analysis with the thresholds set at mFG score ≥8 for White Iranians, ≥7 for White Italians and Black Africans, ≥5 for Asian Han Chinese, White Russians, Turkish, and Black Americans, and ≥4 for White Americans, Asian Koreans, Asian Russians, and Mixed Russians [43]

activity by equilibrium dialysis using Sephadex G-25 and [3H]T as the ligand, and approximations for binding to albumin [55, 56], while one study reported directly measured FT utilizing RIA [48].

Biochemical hyperandrogenism was defined using the selected control subgroups as above 95th percentile in eight studies [26, 49, 51–56], above 97.5th percentile in one study [50], above 97.7th percentile (equivalent to 2 standard deviations above the mean in normally distributed data) in another study [48], and above 98th percentile in one study [31] (Table 1). The prevalence of biochemical hyperandrogenism varied widely across studies, ranging from 5.6 to 30.0%, with an overall prevalence of 17.4%. In this study, we adapted the 95th percentile cut-off specific to each study for harmonization to account for variations in assay methods and reference ranges across studies. We used first-line tests (TT, FT, cFT, and FAI) and second-line tests (A4 and DHEAS) when first-line tests were negative, as per the 2023 International PCOS Guidelines. This approach revealed a biochemical hyperandrogenism prevalence ranging from 11.0 to 30.3%, with an overall prevalence of 20.4%. The overall prevalence remained nearly the same at 20.2% when studies assessing selective subgroups of populations were excluded (Table 4).

Assessment of polycystic ovary morphology (PCOM)

Eight included studies assessed all participants for PCOM. One study provided PCOM as a binary variable [49], OV data was provided by six studies [26, 31, 47, 48, 51, 52], while AFC was provided by seven studies [26, 31, 47, 48, 51–53]. The imaging modalities used included transvaginal ultrasound with a band frequency ranging from 5 to 13 MHz, and transabdominal ultrasound with a band frequency ranging from 1 to 9 MHz. Additionally, transrectal ultrasound was utilized in two studies [26, 52]. Unfortunately, the IPD in the included studies did not specify the ultrasound approach, either transvaginal, transabdominal, or transrectal.

Except for Makwe et al. which defined PCOM as having an AFC count of ≥ 20 and/or OV ≥ 10cm^3 [47], all other included studies used older criteria, defining PCOM as having an AFC count of ≥ 12 and/or OV ≥ 10cm^3 [26, 31, 48, 49, 51–53]. The reported prevalence of PCOM varied widely across studies, ranging from 6.6 to 39.9%, with an overall prevalence of 19.2% (Table 5). Due to the heterogeneity in imaging modalities, ultrasound approaches, and lack of IPD, data harmonization for PCOM was not feasible, and thus we relied on each study's original definitions and categorization.

Prevalence of polycystic ovary syndrome (PCOS)

Of the included study, eight defined PCOS using the Rotterdam criteria [26, 31, 47–49, 51–53], while the remaining four used the National Institutes of Health (NIH 1990) criteria and did not include ultrasound data for PCOM [50, 54–56]. The reported prevalence ranged from 3.3 to 19.8%, with an overall prevalence of 11.7% using each study's original definition. For data harmonization, we defined PCOS by applying the International

Table 4 Prevalence of biochemical hyperandrogenism

Studies	Prevalence				Harmonized androgen cut-offs (95th percentile)
	Original data *n* (%)		Harmonized data *n* (%)		
	Yes	No	Yes	No	
Zhao et al. 2011 [52]	268 (17.6)	1258 (82.4)	250 (17.3)	1197 (82.7)	TT: 3.08 nmol/L, FAI: 6.74, and A4: 17.67 nmol/L
Zhuang et al. 2014 [48]	77 (5.6)	1293 (94.4)	164 (12.1)	1187 (87.9)	TT: 2.95 nmol/L and FT: 3.2 pg/mL (0.010366 nmol/L)
Tehrani et al. 2011a[a] [53]	71 (28.5)	178 (71.5)	64 (27.8)	166 (72.2)	TT: 0.89 ng/mL (3.09 nmol/L), FAI: 5.39, DHEAS: 179 µg/dL (4.85 µmol/L), A4: 2.9 ng/mL (10.12 nmol/L)
Tehrani et al. 2011b [54]	169 (21.3)	623 (78.7)	169 (26.0)	482 (74.0)	TT: 0.88 ng/mL (3.05 nmol/L), FAI: 5.47, DHEAS: 246 µg/dL (6.67 µmol/L), A4: 2.3 ng/mL (8.03 nmol/L)
Tehrani et al. 2014 [49]	160 (30.0)	372 (70.0)	133 (30.3)	306 (69.7)	TT: 0.89 ng/mL (3.09 nmol/L), FAI: 5.39, DHEAS: 179 µg/dL (4.85 µmol/L), A4: 2.9 ng/mL (10.12 nmol/L)
Gambineri et al. 2013[b] [50]	N/R	N/R	20 (11.0)	161 (89.0)	TT: 1.66 nmol/L and A4: 5.78 nmol/L
Makwe et al. 2023 [47]	N/A	N/A	N/A	N/A	N/A
Suturina et al. 2022 [31]	232 (19.7)	947 (80.4)	313 (29.0)	766 (71.0)	TT: 54.77 ng/dL (1.90 nmol/L) for White Russians), 34.33 ng/dL (1.19 nmol/L) for Asian Russians, and 43.78 ng/dL (1.52 nmol/L) for Mixed Russians, FAI: 4.73 for White Russians, 2.11 for Asian Russians, and 2.97 for Mixed Russians, DHEAS: 314.6 µg/dl (8.53 µmol/L) for all Russians.
Kim et al. 2011[a] [26]	16 (10.0)	144 (90.0)	20 (12.8)	136 (87.2)	TT: 0.68 ng/mL (2.36 nmol/L), FAI: 5.36, DHEAS: 10.55 µmol/L
Yildiz et al. 2012 [51]	68 (18.2)	305 (81.8)	67 (18.2)	302 (81.8)	TT: 54.7 ng/dL (1.9 nmol/L), FAI: 4.94, DHEAS: 3257.4 ng/mL (8.83 µmol/L), A4: 2.97 ng/mL (10.37 nmol/L)
Knochenhauer et al. 1998 [56]	37 (20.4)	144 (79.6)	37 (20.4)	144 (79.6)	TT: 84.7 ng/dl (2.94 nmol/L), cFT: 0.75 ng/dL (0.026 nmol/L), DHEAS: 2459 ng/ml (6.64 µmol/L), A4: 2496 pg/ml (8.73 nmol/L)
Azziz et al. 2004[a] [55]	30 (26.6)	83 (73.4)	30 (26.8)	82 (73.2)	TT: 84.7 ng/dl (2.94 nmol/L), cFT: 0.75 ng/dL (0.026 nmol/L), DHEAS: 2459 ng/ml (6.64 µmol/L), A4: 2496 pg/ml (8.73 nmol/L)
Overall	1128 (17.4)	5348 (82.6)	1267 (20.4)	4929 (79.7)	n/a

N/A data not available, *N/R* not reported, *n/a* not applicable

[a] Studies assessing a selective group of participants for biochemical hyperandrogenism

[b] very small sample size. We calculated the 95th percentile from selected healthy controls for TT in Zhuang et al. (2014), A4 in Gambineri et al. (2013), and DHEAS in Kim et al. (2011); the original authors did not use these androgens in their classifications. Additionally, we calculated the 95th percentile for TT in Gambineri et al. (2013). The reported laboratory values were converted to the SI unit using the following conversion factors: for A4 3.49 from ng/mL to nmol/L and 0.00349 from pg/mL to nmol/L; for TT 3.47 from ng/mL to nmol/L and 0.0347 from ng/dL to nmol/L; for FT 0.003467 from pg/mL to nmol/L and 0.03467 from ng/dL to nmol/L; for DHEAS 0.00271 ng/mL to µmol/L and 0.0271 µg/dL to µmol/L. For data harmonization, we utilized the 95th percentile cut-off specific to each individual study population, applying first-line tests (TT, FT, and FAI) and incorporating A4 and DHEAS when first-line tests were negative, in accordance with the 2023 International PCOS Guidelines

PCOS Guidelines criteria, requiring two out of three key features: OD, hyperandrogenism, and PCOM. We defined OD as fewer than eight or more than 17 menstrual cycles per year and/or a menstrual cycle length of less than 21 days or greater than 35 days; we defined hyperandrogenism as either clinical (limited to ethnicity-specific hirsutism scores) or biochemical (using study-specific 95th percentile cut-offs for first-line androgen tests, supplemented by secondary markers when first-line tests were negative), in accordance with the 2023 International PCOS Guidelines recommendations. Post harmonization, the prevalence of PCOS ranged from 3.3 to 18.3%, with an overall prevalence of 11.0% (Table 6).

Discussion

The P-PUP study represents the first international IPD-based meta-analysis involving unselected (medically unbiased) women from the general population, assessed for PCOS features, offering a unique and diverse cohort to explore PCOS characteristics across heterogeneous populations. This comprehensive dataset includes a large sample size of 9979 women aged 18–45 years from 11 primary ethnicities, eight countries and four continents. Our analysis revealed significant heterogeneity in data collection methods and the definitions used for PCOS diagnostic features, particularly in the assessment of ovulatory dysfunction. This variability contributed to a wide range of reported prevalence of PCOS across original studies, from 3.3 to 19.8% (11.7% overall). After data harmonization, the PCOS prevalence ranged from 3.3 to 18.3%, with an overall prevalence of 11.0%.

Ovulatory dysfunction (OD) is a key component in the International Guidelines PCOS diagnostic criteria, focused on menstrual cyclicity, and irregular cycle length and number per year. The International PCOS Guidelines [2–6], defined irregular cycles as > 90 days for any one cycle in women 1-year post menarche, < 21 or > 45 days in women one to less than 3 years post menarche, < 21

Table 5 Prevalence of polycystic ovary morphology (PCOM)

Studies	Prevalence			
	Original data *n* (%)		Harmonized data *n* (%)	
	Yes	No	Yes	No
Zhao et al. 2011 [52]	567 (19.4)	2348 (80.6)	546 (19.2)	2290 (80.8)
Zhuang et al. 2014 [48]	91 (6.6)	1278 (93.4)	90 (6.7)	1260 (93.3)
Tehrani et al. 2011a [53]	N/A	N/A	N/A	N/A
Tehrani et al. 2011b [54]	156 (17.0)	759 (83.0)	132 (17.4)	628 (82.6)
Tehrani et al. 2014 [49]	118 (19.6)	483 (80.4)	99 (19.8)	400 (80.2)
Gambineri et al. 2013 [50]	N/A	N/A	N/A	N/A
Makwe et al. 2023 [47]	151 (34.5)	287 (65.5)	139 (34.0)	270 (66.0)
Suturina et al. 2022 [31]	285 (20.5)	1104 (79.5)	266 (20.8)	1013 (79.2)
Kim et al. 2011 [26]	91 (23.6)	295 (76.4)	89 (23.9)	283 (76.1)
Yildiz et al. 2012 [51]	155 (39.9)	234 (60.1)	154 (39.8)	233 (60.2)
Knochenhauer et al. 1998 [56]	N/A	N/A	N/A	N/A
Azziz et al. 2004 [55]	N/A	N/A	N/A	N/A
Overall	1614 (19.2)	6788 (80.8)	1515 (19.2)	6377 (80.8)

For data harmonization, we used the original authors' categorization for PCOM, which was defined as an antral follicle count (AFC) ≥12 follicles and/or ovarian volume (OV) ≥10 cm^3, except for Makwe et al. (2023), where PCOM was defined as AFC ≥20 follicles and/or OV ≥10 cm^3

PCOM polycystic ovary morphology, *N/A* data not available

Table 6 Prevalence of polycystic ovary syndrome (PCOS)

Studies	Prevalence			
	Original data *n* (%)		Harmonized data *n* (%)	
	Yes	No	Yes	No
Zhao et al. 2011 [52]	346 (11.5)	2654 (88.5)	360 (12.3)	2559 (87.7)
Zhuang et al. 2014 [48]	153 (11.2)	1217 (88.8)	169 (12.5)	1182 (875)
Tehrani et al. 2011a[a] [53]	86 (8.6)	911 (91.4)	63 (6.5)	912 (93.5)
Tehrani et al. 2011b [54]	139 (15.2)	776 (84.8)	120 (15.8)	640 (84.2)
Tehrani et al. 2014 [49]	84 (14.0)	517 (86.0)	73 (14.6)	426 (85.4)
Gambineri et al. 2013[a] [50]	N/R	N/R	6 (3.3)	176 (96.7)
Makwe et al. 2023 [47]	38 (8.7)	400 (91.3)	15 (3.7)	394 (96.3)
Suturina et al. 2022 [31]	168 (14.4)	1001 (85.6)	151 (11.7)	1140 (88.3)
Kim et al. 2011 [26]	57 (11.4)	441 (88.6)	43 (8.9)	440 (91.1)
Yildiz et al. 2012 [51]	77 (19.8)	312 (80.2)	71 (18.3)	316 (81.7)
Knochenhauer et al. 1998[a] [56]	12 (3.3)	346 (96.7)	12 (3.3)	346 (96.7)
Azziz et al. 2004* [55]	19 (5.2)	347 (94.8)	16 (4.4)	349 (95.6)
Overall	1179 (11.7)	8922 (88.3)	1099 (11.0)	8880 (89.0)

PCOS polycystic ovary syndrome. N/R: not reported

[a] NIH criteria. For data harmonization, we defined PCOS using two out of the three key features—OD, hyperandrogenism (clinically defined based on hirsutism and/or biochemically), and PCOM—in studies reporting all three key features, and OD and hyperandrogenism in studies where PCOM was not reported. Secondary causes that may mimic PCOS, such as hyperprolactinemia, thyroid dysfunction, and non-classical adrenal hyperplasia, were excluded

or > 35 days or < 8 cycles per year in women 3 years postmenarche to perimenopause, while acknowledging the need for further research [57]. We identified significant variations in menstrual cycle data collection, especially prior to the International PCOS Guidelines. Categorical and continuous data was collected, including average cycle length in days, average cycle number per year, minimum and maximum cycle length and cycle number per year. OD cut-offs were variable across studies. Here, we harmonized OD data aligned with Guideline recommendations, reporting a prevalence of 3.7% to 31.9% or 15.5% overall. Future studies should collect standardized

continuous data on cycle length (in days) and cycle number per year, with clear explanations of cycle length as the number of days from the first day of one period to the first day of the next, to limit confusion with the duration of menstrual bleeding. While acknowledging the inherent variability of cycles in PCOS and limitations of self-reported data for large-scale epidemiology studies, standardized data collection and definitions would substantially minimize error while ensuring consistency and comparability across studies. The International PCOS Guideline's OD cut-offs were derived from the 95th percentile of menstrual cycle duration, based on variable control populations which likely included women with PCOS [58]. Using this percentile as a diagnostic cut-off, particularly when it is not linked with clinical outcomes, including anthropometric outcomes (weight, waist-hip circumference, BMI), reproductive outcomes (androgens, ovulation frequency, pregnancy complications), and cardiometabolic (insulin resistance, lipid profiles, blood pressure, fasting glucose), represents a crude and overly simplistic approach [58]. Moving forward, applying advanced analytics including cluster analysis to define cut-offs based on well-defined unselected populations and clinical correlates such as association with other PCOS features or with long-term outcomes is a research priority captured in the International PCOS Research Road Map [59].

Hyperandrogenism is an essential PCOS diagnostic feature and manifests clinically as hirsutism, acne and androgenic alopecia. Acne and androgenic alopecia are relatively nonspecific markers of hyperandrogenism, particularly in PCOS [2–5, 16]. While validated scoring systems exist for acne and androgenic alopecia, these are not universally agreed upon for PCOS due to limited research in PCOS populations, ethnic variations, and clinical contexts. In included original studies, acne and alopecia were inconsistently reported, with variable tools and definitions. In contrast, hirsutism is strongly correlated with biochemical hyperandrogenism and is the most reliable clinical sign of hyperandrogenism in PCOS [60, 61]. The modified Ferriman-Gallwey (mFG) visual scale is Guideline recommended [62] and was used in all included original studies. Historically, a cut-off mFG score of ≥ 8 defined hirsutism in PCOS, based on a misinterpretation of original Ferriman and Gallwey data at the 95th percentile of a predominantly white population [63, 64]. In P-PUP IPD we established ethnicity-specific mFG cut-offs via *k*-means cluster analysis, with cut-offs ranging between 4–6 for the majority of populations [43]. This aligns with the International PCOS Guidelines recommendations for an mFG score of 4–6 to diagnose hirsutism [2–5]. Clinical hyperandrogenism ranged from 6.8 to 28.4% (13.7% overall) in original studies with variable mFG cut-offs and use of acne and alopecia. With cluster analysis cut-offs and data harmonization, we report a prevalence of 8.1% to 34.2% (14.2% overall) and aligned to Guidelines, cluster-defined mFG cut-offs are recommended.

Biochemical hyperandrogenism in PCOS focuses on elevated free testosterone, which may be measured directly or calculated primarily through FAI and cFT. Our recent systematic review and meta-analysis informed the International PCOS Guidelines recommendation to use TT and FT as a first-line laboratory test for biochemical hyperandrogenism [65] using LC–MS/MS methods with the best accuracy [66, 67]. In the P-PUP study, only two original studies used LC–MS/MS [31, 50], all others relied on less specific and sensitive immunoassay. Cut-offs for biochemical hyperandrogenism varied, with arbitrary thresholds at the 95th, 97.5th, or 98th percentile of healthy controls. In P-PUP we adapted the 95th percentile cut-off for androgens specific to each study to harmonize data reporting a prevalence from 11.0 to 30.3% in original studies, with an overall prevalence of 20.4%. Further cluster analyses are underway to refine definitions and address persisting limitations.

PCOM is evaluated using pelvic ultrasonography and is characterized by either ovarian enlargement, measured as increased OV, and/or an excess of small antral follicles, either within the entire ovary (FNPO) or a single cross-sectional image (FNPS) [68]. Cut-offs for OV and FNPO vary by ultrasound technology and approach. The 2023 International PCOS Guidelines recommend the transvaginal ultrasound approach, FNPO ≥ 20 in at least one ovary or OV ≥ 10 ml to define PCOM in adults. If older ultrasound technology is used, the AFC count is ≥ 12 in at least one ovary. PCOM is not recommended in adolescents as it lacks specificity [2–5]. In the present analysis, only one study adhered to these recommendations and transabdominal and transrectal ultrasound approaches were also used, with no validated PCOM cut-offs, limiting harmonization and resulting in the retention of original study definitions. Guidelines now clearly recommend a transvaginal ultrasound, and cut-offs for FNPO and OV which need to be applied, along with the inclusion of AMH measurements.

Strengths and limitations

This study has several strengths. Firstly, this IPD is unique and includes a large diverse, international IPD set, with generalizability to those with PCOS. Unselected participants also minimize selection bias, as seen in clinic-based cohorts. Also, the IPD analysis allowed standardization of definitions and cut-offs across studies and aligned with the 2023 International PCOS Guidelines to minimize heterogeneity and enhance comparability. Limitations

include heterogeneity in data collection methods, particularly for ovulatory dysfunction, hyperandrogenism, and PCOM. Some original data could not be harmonized, affecting the comparability of results across populations. Most original studies relied on immunoassays for androgen measurement, which are less accurate than LC-MS/MS. Most studies also used older ultrasound technology and variable approaches and lacked sufficient methodological detail, with reliance on original study PCOM definitions. Although serum AMH is now recommended as an alternative marker for PCOM, this data was not available. Lastly, although the study included participants from four continents and eight countries, including participants from diverse ethnic backgrounds enhanced generalizability, notable geographic and cultural gaps in representation persist, underscoring the need for more globally comprehensive PCOS research.

Conclusions

The P-PUP study brings together an unprecedented cohort of ethnically diverse, unselected population-level data to enable research and advance knowledge in PCOS. In integrating this IPD we have shown a lack of standardized data collection methods across original studies. Cut-offs for PCOS diagnostic features also vary, jeopardizing diagnostic accuracy and consistency and impacting data harmonization. This work directly addresses the high research priorities identified in the PCOS Guidelines research roadmap, which emphasizes creating international data repositories and standardizing research approaches [59]. Future research efforts should prioritize the standardization of data collection methods that align with the International PCOS Guidelines and PCOS core outcome sets [39], while simultaneously refining definitions and establishing accurate normative cut-offs for individual PCOS diagnostic features. These diagnostic thresholds should be based on cluster analysis and natural groupings, identifying diagnostic clusters that are linked to short and long-term outcomes in medically unbiased, unselected, and diverse populations. This dual approach to standardization would facilitate big data integration and the development of machine learning–driven models for precision medicine in PCOS.

Abbreviations

A4	Androstenedione
AFC	Antral follicular count
BMI	Body mass index
cFT	Calculated free testosterone
DBP	Diastolic blood pressure
DHEAS	Dehydroepiandrosterone sulfate
FAI	Free androgen index
FBS	Fasting blood sugar
FT	Free testosterone
HDL	High-density lipoprotein
HOMA-IR	Homeostatic model assessment for insulin resistance
IPD	Individual participant data
LDL	Low-density lipoprotein
mFG	Modified Ferriman-Gallwey score
NIH	National Institute of Health
OV	Ovarian volume
PCOM	Polycystic ovary morphology
PCOS	Polycystic ovary syndrome
P-PUP	PCOS Phenotype in Unselected Populations
SBP	Systolic blood pressure
TT	Total testosterone

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-04221-9.

Additional file 1: Table S1. Search strategy for OVID Medline. Table S2. Studies excluded following full-text review. Table S3. Risk of bias assessment using AXIS tool. Table S4. Harmonized variables with definitions in the P-PUP study dataset. Table S5. Data collection methods for ovulatory dysfunction across studies.

Acknowledgements

We acknowledge all collaborators participating in this individual participant data meta-analysis project for their generous provision of data and scientific expertise. We are profoundly grateful to the many thousands of participants in the individual studies whose contributions made this research possible.

Authors' contributions

Author contributions ADB: involved in study design, data analysis, interpretation of results, and manuscript preparation. CTT: study design, supervision of data analysis, interpretation of results, and manuscript preparation. AEJ: involved in funding, supervision, conception, study design, acquisition of data, supervision of data analysis, interpretation of results, and manuscript preparation. A.E: involved in the supervision of data analysis, interpretation of results, and manuscript preparation. HJT: involved in funding, supervision, conception, study design, interpretation of results, and manuscript writing. SK and RA: involved in conception, study design, interpretation of results, and manuscript writing. MBK, RRD, LVS, XZ, AG, FRT, BOY, JJK, LX, and CCM: were involved in original study design, acquisition of data, interpretation of results and manuscript writing. All the authors have read and approved the final version of the manuscript. The corresponding author attests that all listed authors meet the authorship criteria.

Authors' Twitter handles

Twitter handles: @BahriMahnaz (Mahnaz Bahri Khomami).

Funding

This study was funded by the Australian National Health and Medical Research Council (NHMRC) via the Centre for Research Excellence for Women Health in Reproductive Life (CRE-WHiRL) #1171592. ADB is supported by a PhD scholarship funded by the Monash Graduate Scholarship (MGS) and the Monash International Tuition Scholarship (MITS). CMM received grants from the University of Lagos, Central Research Grant CRC/03/2018, Ferring Pharmaceuticals Grant #2021/2944, and the Foundation for Research and Education Excellence. HJT is supported by NHMRC fellowships #2009326, AEJ is funded by the NHMRC CRE-WhiRL fellowship, CTT is supported by CRE-WHiRL funded by the NHMRC, and RA is supported by funding from the Foundation for Research and Education Excellence.

Data availability

The individual participant data supporting the findings of this study are available at [https://bridges.monash.edu/articles/dataset/The_PCOS_Phenotype_in_Unselected_Populations_Study_P-PUP_/22590316], but restrictions apply. Access requests can be submitted to Anju E Joham (anju.joham@monash.edu) and Chau Thien Tay (jillian.tay@monash.edu), subject to data use agreements

and ethical approvals from the respective institutional review boards of the contributing studies.

Declarations

Ethics approval and consent to participate
Ethical approval for the P-PUP study was obtained from the Monash University Human Research Ethics Committee (HUMREC) (ID: 26938). Additionally, each study within the research project received ethical approval from the relevant Human Research Ethics Committee.

Consent for publication
Not applicable.

Competing interests
AEJ serves as a Board member for Androgen Excess and Polycystic Ovary Syndrome, received honoraria for presentations at educational events at Novo Nordisk, Boehringer and Merck, and is an employee of Monash Health. CTT serves as Chair of CRE-WHiRL ECR Group 2020-2023 and AEPCOS EC-SIG group 2020-2022, as a committee member of the Endocrine Society of Australia's ECR Group. AG serves on the Editorial Board for the European Journal of Endocrinology and on the Advisory Board of the Italian Society of Endocrinology. RA serves as a consultant to May Health, Core Access Surgical Technologies, Spruce Biosciences, and Postera; is an investor in Martin Imaging; received honoraria for speaking from the Davidson-Mestman course, Merck and Stya Paul Oration; and serves on the Editorial Board for the Journal of Clinical Endocrinology and Metabolism, on the Board of Trustees for the Endocrine Society, as a member of the DSMB for grant NCT03625531; ChiCTR180001730 to The First Affiliated of Guangzhou Medical University and grant for the SUPER study to the University of Michigan, and previously served as CEO of the American Society for Reproductive Medicine. All the other authors declare that they have no conflicts of interest.

Author details
[1]Monash Centre for Health Research and Implementation, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC 3168, Australia. [2]School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC 3004, Australia. [3]Departments of Obstetrics and Gynecology, and Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA. [4]Department of Reproductive Health Protection, Scientific Center for Family Health and Human Reproduction Problems, Irkutsk 664003, Russian Federation. [5]Department of Reproductive Medicine, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, 106 Zhongshan 2 Road, Guangzhou 510080, China. [6]Department of Medical and Surgical Science-DIMEC, Division of Endocrinology and Diabetes Prevention and Care, IRCCS Azienda Ospedaliero-Universitaria Di Bologna, Via Massarenti 9, Bologna 40138, Italy. [7]Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, 24 Parvaneh, Yaman Street, Velenjak, Tehran 19395-4763, Iran. [8]Division of Endocrinology and Metabolism, Department of Internal Medicine, Hacettepe University School of Medicine, Ankara 06100, Turkey. [9]Division of Reproductive Endocrinology, Department of Obstetrics and Gynecology, Healthcare System Gangnam Center, Seoul National University Hospital, Seoul 06236, Republic of Korea. [10]Department of Obstetrics and Gynecology, West China Second University Hospital, Sichuan University, Chengdu 610041, China. [11]Department of Obstetrics and Gynaecology, College of Medicine, University of Lagos, Idi-Araba, Lagos, Nigeria, P.M.B. 12003, Surulere, Lagos, Nigeria. [12]Department of Endocrinology and Diabetes, Monash Health, Clayton, Melbourne, VIC 3168, Australia.

Received: 28 January 2025 Accepted: 16 June 2025
Published online: 15 July 2025

References

1. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod. 2016;31(12):2841–55.
2. Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, Costello MF, Boivin J, Redman LM, Boyle JA, 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):2447–69.
3. Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, Costello MF, Boivin J, Redman LM, Boyle JA, et al. Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Eur J Endocrinol. 2023;189(2):G43–64.
4. Teede HJ, Tay CT, Laven J, Dokras A, Moran LJ, Piltonen TT, Costello MF, Boivin J, Redman LM, Boyle JA, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. Hum Reprod. 2023;38(9):1655–79.
5. Teede HJ, Tay CT, Laven J, Dokras A, Moran LJ, Piltonen TT, Costello MF, Boivin JM, Redman LA, Boyle J, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. Fertil Steril. 2023;120(4):767–93.
6. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum Reprod. 2018;33(9):1602–18.
7. Liu J, Wu Q, Hao Y, Jiao M, Wang X, Jiang S, Han L. Measuring the global disease burden of polycystic ovary syndrome in 194 countries: Global Burden of Disease Study 2017. Hum Reprod. 2021;36(4):1108–19.
8. Safiri S, Noori M, Nejadghaderi SA, Karamzad N, Carson-Chahhoud K, Sullman MJ, Collins GS, Kolahi A-A, Avery J. Prevalence, incidence and years lived with disability due to polycystic ovary syndrome in 204 countries and territories, 1990–2019. Hum Reprod. 2022;37(8):1919–31.
9. Azziz R, Marin C, Hoq L, Badamgarav E, Song P. Health care-related economic burden of the polycystic ovary syndrome during the reproductive life span. J Clin Endocrinol Metab. 2005;90(8):4650–8.
10. Riestenberg C, Jagasia A, Markovic D, Buyalos RP, Azziz R. Health care-related economic burden of polycystic ovary syndrome in the United States: pregnancy-related and long-term health consequences. J Clin Endocrinol Metab. 2022;107(2):575–85.
11. Yadav S, Delau O, Bonner AJ, Markovic D, Patterson W, Ottey S, Buyalos RP, Azziz R. Direct economic burden of mental health disorders associated with polycystic ovary syndrome: Systematic review and meta-analysis. Elife. 2023;12:e85338.
12. Teede H, Deeks A, Moran L. Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. BMC Med. 2010;8(1):1–10.
13. Cai M, Ni Z, Yuan Z, Yu J, Zhang D, Yao R, Zhou L, Yu C. Past and present: a bibliometric study on polycystic ovary syndrome. J Ovarian Res. 2023;16(1):42–42.
14. Engmann LMD, Jin SMPH, Sun FMPH, Legro RSMD, Polotsky AJMDM, Hansen KRMDP, Coutifaris CMDP, Diamond MPMD, Eisenberg EMDMPH, Zhang HP, et al. Racial and ethnic differences in the polycystic ovary syndrome metabolic phenotype. Am J Obstet Gynecol. 2017;216(5):493 e491-493.e413.
15. VanHise K, Wang ET, Norris K, Azziz R, Pisarska MD, Chan JL. Racial and ethnic disparities in polycystic ovary syndrome. Fertil Steril. 2023;119(3):348–54.
16. Azziz R, Kintziger K, Li R, Laven J, Morin-Papunen L, Merkin SS, Teede H, Yildiz BO. Recommendations for epidemiologic and phenotypic research in polycystic ovary syndrome: an androgen excess and PCOS society resource. Hum Reprod. 2019;34(11):2254–65.
17. Luque-Ramírez M, Alpañés M, Sanchón R, Fernández-Durán E, Ortiz-Flores AE, Escobar-Morreale HF. Referral bias in female functional hyperandrogenism and polycystic ovary syndrome. Eur J Endocrinol. 2015;173(5):603–10.
18. Lizneva D, Kirubakaran R, Mykhalchenko K, Suturina L, Chernukha G, Diamond MP, Azziz R. Phenotypes and body mass in women with polycystic ovary syndrome identified in referral versus unselected populations: systematic review and meta-analysis. Fertil Steril. 2016;106(6):1510–20 e1512.
19. Ezeh U, Yildiz BO, Azziz R. Referral bias in defining the phenotype and prevalence of obesity in polycystic ovary syndrome. J Clin Endocrinol Metab. 2013;98(6):E1088–96.

20. Zawadski JK, Dunaif A. Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. In: Dunaif A, Givens JR, Haseltine F (eds). Polycystic Ovary Syndrome. Boston: Blackwell Scientific, 1992:377–384.
21. Rotterdam ESHRE ASRM Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum Reprod. 2004;19(1):41–7.
22. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE. Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an androgen excess society guideline. J Clin Endocrinol Metab. 2006;91(11):4237–45.
23. Group REASPCW. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum Reprod. 2004;19(1):41–7.
24. Taieb J, Mathian B, Millot F, Patricot MC, Mathieu E, Queyrel N, Lacroix I, Somma-Delpero C, Boudou P. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. Clin Chem. 2003;49(8):1381–95.
25. DeUgarte CM, Woods K, Bartolucci AA, Azziz R. Degree of facial and body terminal hair growth in unselected black and white women: toward a populational definition of hirsutism. J Clin Endocrinol Metab. 2006;91(4):1345–50.
26. Kim JJ, Chae SJ, Choi YM, Hwang SS, Hwang KR, Kim SM, Yoon SH, Moon SY. Assessment of hirsutism among Korean women: results of a randomly selected sample of women seeking pre-employment physical check-up. Hum Reprod. 2011;26(1):214–20.
27. Li R, Qiao J, Yang D, Li S, Lu S, Wu X, Wei Z. Epidemiology of hirsutism among women of reproductive age in the community: a simplified scoring system. Eur J Obstet Gynecol Reprod Biol. 2012;163(2):165–9.
28. DÖNer N, EkmekÇİ TR. Criticising Modified Ferriman-Gallwey Scroring System in the Evaluation of Hirsutism in 1034 Turkish Women. Turkiye Klinikleri J Med Sci. 2013;33(2):314–20.
29. Zhou Z, Li R, Qiao J. Androgen profile in Chinese women with polycystic ovary syndrome in their reproductive years. Reprod Biomed Online. 2017;35(3):331–9.
30. Ilagan MKCC, Paz-Pacheco E, Totesora DZ, Clemente-Chua LR, Jalique JRK. The Modified Ferriman-Gallwey Score and Hirsutism among Filipino Women. Endocrinol Metab. 2019;34(4):374–81.
31. Suturina L, Lizneva D, Atalyan A, Lazareva L, Belskikh A, Bairova T, Sholokhov L, Rashidova M, Danusevich I, Nadeliaeva I, et al. Establishing Normative Values to Determine the Prevalence of Biochemical Hyperandrogenism in Premenopausal Women of Different Ethnicities from Eastern Siberia. Diagnostics (Basel). 2022;13(1):33.
32. Gibson-Helm M, Teede H, Dunaif A, Dokras A. Delayed diagnosis and a lack of information associated with dissatisfaction in women with polycystic ovary syndrome. J Clin Endocrinol Metab. 2017;102(2):604–12.
33. Wilcock J, Taylor D. Polycystic ovarian syndrome: overdiagnosed and overtreated? Br J Gen Pract. 2018;68(670):243–243.
34. Joham AE, Piltonen T, Lujan ME, Kiconco S, Tay CT. Challenges in diagnosis and understanding of natural history of polycystic ovary syndrome. Clin Endocrinol (Oxf). 2022;97(2):165–73.
35. Dokras A, Saini S, Gibson-Helm M, Schulkin J, Cooney L, Teede H. Gaps in knowledge among physicians regarding diagnostic criteria and management of polycystic ovary syndrome. Fertil Steril. 2017;107(6):1380–6 e1381.
36. Kiconco S, Mousa A, Azziz R, Enticott J, Suturina LV, Zhao X, Gambineri A, Tehrani FR, Yildiz BO, Kim JJ, et al. PCOS Phenotype in Unselected Populations Study (P-PUP): Protocol for a Systematic Review and Defining PCOS Diagnostic Features with Pooled Individual Participant Data. Diagnostics (Basel). 2021;11(11):1953.
37. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF. Group ftP-ID: Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. JAMA. 2015;313(16):1657–65.
38. Jones KH, Ford DV, Ellwood-Thompson S, Lyons RA. The UK Secure eResearch Platform for public health research: a case study. The Lancet. 2016;388:S62.
39. Al Wattar BH, Teede H, Garad R, Franks S, Balen A, Bhide P, Piltonen T, Romualdi D, Laven J, Thondan M, et al. Harmonising research outcomes for polycystic ovary syndrome: an international multi-stakeholder core outcome set. Hum Reprod. 2020;35(2):404–12.
40. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open. 2016;6(12): e011458.
41. Oliver E, Coates A, Bennett JM, Willis ML. Narcissism and Intimate Partner Violence: A Systematic Review and Meta-Analysis. Trauma Violence Abuse. 2024;25(3):1871–84.
42. Barnett V, Lewis T. Outliers in statistical data. 3rd ed. Chichester: Wiley;1994.
43. Bizuneh AD, Joham AE, Tay CT, Kiconco S, Earnest A, Dhungana RR, Suturina LV, Zhao X, Gambineri A, Ramezani Tehrani F, et al. The PCOS Phenotype in Unselected Populations study: ethnic variation in population-based normative cut-offs for defining hirsutism. Eur J Endocrinol. 2025;192(3):228–39.
44. Chen X, Yang D, Mo Y, Li L, Chen Y, Huang Y. Prevalence of polycystic ovary syndrome in unselected women from southern China. Eur J Obstet Gynecol Reprod Biol. 2008;139(1):59–64.
45. Joshi B, Mukherjee S, Patil A, Purandare A, Chauhan S, Vaidya R. A cross-sectional study of polycystic ovarian syndrome among adolescent and young girls in Mumbai. India Indian J Endocrinol Metab. 2014;18(3):317–24.
46. Li R, Zhang Q, Yang D, Li S, Lu S, Wu X, Wei Z, Song X, Wang X, Fu S, et al. Prevalence of polycystic ovary syndrome in women in China: a large community-based study. Hum Reprod. 2013;28(9):2562–9.
47. Makwe CC, Olamijulo JA, Balogun MR, Akinkugbe AO, Samuel A, Udenze K, Bril F, Buyalos RP, Laven JSE, Azziz R. SAT363 The Prevalence And Phenotype Of Polycystic Ovary Syndrome (PCOS) In A Community-Based Population In Sub-Sahara Africa: The Nigeria-PEP Study. J Endocr Soc. 2023;7(Supplement_1).
48. Zhuang J, Liu Y, Xu L, Liu X, Zhou L, Tang L, Kang D, Guo W, He M, Yang F, et al. Prevalence of the Polycystic Ovary Syndrome in Female Residents of Chengdu. China Gynecol Obstet Invest. 2014;77(4):217–223.
49. Tehrani FR, Rashidi H, Khomami MB, Tohidi M, Azizi F. The prevalence of metabolic disorders in various phenotypes of polycystic ovary syndrome: a community based study in Southwest of Iran. Reprod Biol Endocrinol. 2014;12(1):89–89.
50. Gambineri A, Fanelli F, Prontera O, Repaci A, Di Dalmazi G, Zanotti L, Pagotto U, Flacco ME, Guidi J, Fava GA, et al. Prevalence of Hyperandrogenic States in Late Adolescent and Young Women: Epidemiological Survey on Italian High-School Students. J Clin Endocrinol Metab. 2013;98(4):1641–50.
51. Yildiz BO, Bozdag G, Yapici Z, Esinler I, Yarali H. Prevalence, phenotype and cardiometabolic risk of polycystic ovary syndrome under different diagnostic criteria. Hum Reprod. 2012;27(10):3067–73.
52. Zhao XMDPD, Ni RMDPD, Li LLMDPD, Mo YPD, Huang JMD, Huang MMD, Azziz RMDMBAMPH, Yang DMDPD. Defining hirsutism in Chinese women: a cross-sectional study. Fertil Steril. 2011;96(3):792–6.
53. Tehrani FR, Simbar M, Tohidi M, Hosseinpanah F, Azizi F. The prevalence of polycystic ovary syndrome in a community sample of Iranian population: Iranian PCOS prevalence study. Reprod Biol Endocrinol. 2011;9(1):39–39.
54. Tehrani FR, Rashidi H, Azizi F. The prevalence of idiopathic hirsutism and polycystic ovary syndrome in the Tehran Lipid and Glucose Study. Reprod Biol Endocrinol. 2011;9(1):144–144.
55. Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The Prevalence and Features of the Polycystic Ovary Syndrome in an Unselected Population. J Clin Endocrinol Metab. 2004;89(6):2745–9.
56. Knochenhauer ES, Key TJ, Kahsar-Miller M, Waggoner W, Boots LR, Azziz R. Prevalence of the polycystic ovary syndrome in unselected Black and White women of the Southeastern United States : A prospective study. J Clin Endocrinol Metab. 1998;83(9):3078–82.
57. Treloar AE, Boynton RE, Behn BG, Brown BW. The menstrual cycle: Variation of the human menstrual cycle through reproductive life. Obstet Gynecol Surv. 1968;23(1):81–4.
58. Teede H, Tay CT, Azziz R. Response to Letter to the Editor From Rosenfield et al:"Recommendations From the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome." J Clin Endocrinol Metab. 2024;109(10):e1976–7.
59. Teede H, Gibson M, Laven J, Dokras A, Moran L, Piltonin T, Costello M, Mousa A, Joham A, Tay C. International PCOS guideline clinical research priorities roadmap: a co-designed approach aligned with end-user priorities in a neglected women's health condition. EClinicalMedicine. 2024;78:102927–102927.

60. Azziz R, Sanchez LA, Knochenhauer ES, Moran C, Lazenby J, Stephens KC, Taylor K, Boots LR. Androgen Excess in Women: Experience with Over 1000 Consecutive Patients. J Clin Endocrinol Metab. 2004;89(2):453–62.
61. Amiri M, Ramezani Tehrani F, Nahidi F, Bidhendi Yarandi R, Behboudi-Gandevani S, Azizi F. Association between biochemical hyperandrogenism parameters and Ferriman-Gallwey score in patients with polycystic ovary syndrome: A systematic review and meta-regression analysis. Clin Endocrinol. 2017;87(3):217–30.
62. Yildiz BO, Bolour S, Woods K, Moore A, Azziz R. Visually scoring hirsutism. Hum Reprod Update. 2009;16(1):51–64.
63. Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. J Clin Endocrinol Metab. 1961;21(11):1440–7.
64. Ferriman D, Purdie AW. The aetiology of oligomenorrhoea and/or hirsuties: a study of 467 patients. Postgrad Med J. 1983;59(687):17–20.
65. Bizuneh AD, Joham AE, Teede H, Mousa A, Earnest A, Hawley JM, Smith L, Azziz R, Arlt W, Tay CT. Evaluating the diagnostic accuracy of androgen measurement in polycystic ovary syndrome: a systematic review and diagnostic meta-analysis to inform evidence-based guidelines. Hum Reprod Update. 2025;31(1):48–63.
66. Janse F, Eijkemans MJC, Goverde AJ, Lentjes EGWM, Hoek A, Lambalk CB, Hickey TE, Fauser BCJM, Norman RJ. Assessment of androgen concentration in women: liquid chromatography–tandem mass spectrometry and extraction RIA show comparable results. Eur J Endocrinol. 2011;165(6):925–33.
67. Rosner W, Vesper H. Toward Excellence in Testosterone Testing: A Consensus Statement. J Clin Endocrinol Metab. 2010;95(10):4542–8.
68. Dewailly D, Lujan ME, Carmina E, Cedars MI, Laven J, Norman RJ, Escobar-Morreale HF. Definition and significance of polycystic ovarian morphology: a task force report from the Androgen Excess and Polycystic Ovary Syndrome Society. Hum Reprod Update. 2014;20(3):334–52.

Publisher's Note