A Polygenic and Phenotypic Risk Prediction for Polycystic Ovary Syndrome Evaluated by Phenome-Wide Association Studies

Yoonjung Yoonie Joo,[1] Ky'Era Actkins,[2] Jennifer A. Pacheco,[3] Anna O. Basile,[4] Robert Carroll,[5] David R. Crosslin,[6] Felix Day,[7] Joshua C. Denny,[5] Digna R. Velez Edwards,[5,8] Hakon Hakonarson,[9,10] John B. Harley,[11,12] Scott J Hebbring,[13] Kevin Ho,[14] Gail P. Jarvik,[15] Michelle Jones,[16] Tugce Karaderi,[17] Frank D. Mentch,[9] Cindy Meun,[18] Bahram Namjou,[11] Sarah Pendergrass,[14] Marylyn D. Ritchie,[19] Ian B. Stanaway,[6] Margrit Urbanek,[1] Theresa L. Walunas,[20] Maureen Smith,[3] Rex L. Chisholm,[3] International PCOS Consortium, Abel N. Kho,[20] Lea Davis,[5] and M.Geoffrey Hayes[1,3,21]

[1]Division of Endocrinology, Metabolism, and Molecular Medicine, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, 60611; [2]Department of Microbiology, Immunology, and Physiology, Meharry Medical College, Nashville, Tennessee, 37203; [3]Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, 60611; [4]Department of Biomedical Informatics, Columbia University New York, New York, 10032; [5]Departments of Biomedical Informatics and Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, 37203; [6]Department of Biomedical Informatics and Medical Education, University of Washington School of Medicine, Seattle, Wahington, 98195; [7]MRC Epidemiology Unit, Cambridge Biomedical Campus, University of Cambridge School of Clinical Medicine, Cambridge, CB2 0QQ, United Kingdom; [8]Division of Quantitative Sciences, Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, Tennessee, 37203; [9]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104; [10]Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, 19104; [11]Center for Autoimmune Genomics and Etiology (CAGE), Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229; [12]Department of Pediatrics, University of Cincinnati College of Medicine; US Department of Veterans Affairs, Cincinnati, Ohio 45229; [13]Center for Precision Medicine Research, Marshfield Clinic Research Institute, Marshfield, Wisconsin, 54449, USA; [14]Biomedical and Translational Informatics, Geisinger, Danville, Pennsylvania, 17822; [15]Division of Medical Genetics, Department of Medicine (Medical Genetics) and Genome Sciences, University of Washington Medical School, Seattle, Wahington, 98195; [16]Center for Bioinformatics & Functional Genomics, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California, 90048; [17]The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom; [18]Department of Obstetrics and Gynecology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [19]Department of Genetics, University of Pennsylvania, Philadelphia, Pennsylvania, 19104; [20]Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, 60611; and [21]Department of Anthropology, Northwestern University, Evanston, Illinois 60208

ORCiD numbers: 0000-0001-9506-8742 (Y. Y. Joo); 0000-0003-3802-8183 (R. Carroll); 0000-0001-6994-7110 (M. Urbanek); 0000-0002-4617-3981 (M.G. Hayes).

Context: As many as 75% of patients with polycystic ovary syndrome (PCOS) are estimated to be unidentified in clinical practice.

Objective: Utilizing polygenic risk prediction, we aim to identify the phenome-wide comorbidity patterns characteristic of PCOS to improve accurate diagnosis and preventive treatment.

Design, Patients, and Methods: Leveraging the electronic health records (EHRs) of 124 852 individuals, we developed a PCOS risk prediction algorithm by combining polygenic risk scores (PRS) with PCOS component phenotypes into a polygenic and phenotypic risk score (PPRS). We evaluated its predictive capability across different ancestries and perform a PRS-based phenome-wide association study (PheWAS) to assess the phenomic expression of the heightened risk of PCOS.

Results: The integrated polygenic prediction improved the average performance (pseudo-R^2) for PCOS detection by 0.228 (61.5-fold), 0.224 (58.8-fold), 0.211 (57.0-fold) over the null model across European, African, and multi-ancestry participants respectively. The subsequent PRS-powered PheWAS identified a high level of shared biology between PCOS and a range of metabolic and endocrine outcomes, especially with obesity and diabetes: "morbid obesity", "type 2 diabetes", "hypercholesterolemia", "disorders of lipid metabolism", "hypertension", and "sleep apnea" reaching phenome-wide significance.

Conclusions: Our study has expanded the methodological utility of PRS in patient stratification and risk prediction, especially in a multifactorial condition like PCOS, across different genetic origins. By utilizing the individual genome–phenome data available from the EHR, our approach also demonstrates that polygenic prediction by PRS can provide valuable opportunities to discover the pleiotropic phenomic network associated with PCOS pathogenesis.

Abbreviations: AA, African ancestry; ANOVA, analysis of variance; BMI, body mass index; EA, European ancestry; EHR, electronic health records; eMERGE, electronic Medical Records and Genomics Network; GWAS, genome-wide association study; IBD, identity-by-descent; ICD-CM, International Classification of Diseases, Clinical Modification; LD, linkage disequilibrium; MA, multi-ancestry; MAF, minor allele frequency; NIH, National Institutes of Health; PCA, principal component analysis; PheWAS, phenome-wide association study; PCOS, polycystic ovary syndrome; PPRS, polygenic and phenotypic risk score; PRS, polygenic risk score; RAF, risk allele frequency; ROC, receiving operating characteristic; SNV, single nucleotide variant. (*J Clin Endocrinol Metab* 105: 1–19, 2020)

Key Words: phenome-wide association study, genomic prediction, polygenic risk score, polycystic ovary syndrome

Polycystic ovary syndrome (PCOS) is the most common reproductive metabolic disorder, affecting 5% to 15% of reproductive age women worldwide (1). The estimated cost of diagnosing and treating American women with PCOS is $5.46 billion annually as of 2017 (2,3). In addition to being a major cause of female infertility, the disease is a well-known risk factor for endocrine complications, such as type 2 diabetes, impaired glucose tolerance, and metabolic syndrome before age 40 (4). Monozygotic twin studies of PCOS have suggested that PCOS is highly heritable (h^2 = ~70%) (5) and the genetic architecture is polygenic with complex genetic inheritance pattern (6,7). Despite its clinical importance and high heritability, the underlying genetic etiology of PCOS remains incompletely understood. Genome-wide association studies (GWASs) suggest androgen biosynthesis (8–11), gonadotropin secretion (8) and action (9,10), ovarian aging (12), and metabolic regulation (11,12) are involved in PCOS pathogenesis.

Meta-analysis of European ancestry (EA) GWASs brings the total number of PCOS GWASs loci to 19 in these pathways.

The phenotypic manifestations of PCOS are heterogeneous and exhibit considerable variation across race and ethnicity, further complicating the clinical diagnosis. Currently, it is estimated that up to 75% of women with PCOS remain undiagnosed. in part due to varying diagnostic criteria from the National Institutes of Health (NIH), Rotterdam, or Androgen Excess Society (13–17), which use different combinations of hyperandrogenism, ovulatory dysfunction, and/or polycystic ovarian morphology. Despite shared genetic risk across the criteria (11), the disagreement regarding PCOS phenotypic criteria presents a significant challenge for both clinical practice and research (18,19). The commonalities and differences between the phenotypic characteristics of PCOS may be better understood with an integrative observation of phenome-wide pleiotropies and comorbidities.

Polygenic risk scores (PRSs) built from well-powered GWASs have demonstrated operationalizing potential as biological risk predictors for patient stratification and risk prediction (20–23). PRS represents the cumulative effect of common genetic variation summed per individual into a single risk score, providing an intuitive way to translate GWAS findings into clinically relevant information such as a patient's risk of disease (24,25). From a precision medicine perspective, PRSs hold significant promise especially for a multifactorial condition with complicated clinical manifestations, such as PCOS. However, several practical challenges remain in the equitable translation of PRSs into clinical practice (26,27). For instance, most GWASs have been performed in samples of primarily EA, resulting in PRS statistics that systematically perform worse in populations of different ancestry, including African ancestry (AA) populations. This underperformance is due to a combination of population-specific genetic effects that are undetected in a Eurocentric GWAS, and differences in the patterns of linkage disequilibrium (LD) between populations of differing biogeographic ancestry (28–31). Thus, the evaluation of PRSs from existing GWASs in both European and non-EA samples is a critical step in setting priorities for equitable precision medicine initiatives.

The widespread deployment of Electronic Health Records (EHRs) and the availability of these multidimensional records enables evaluation of PRSs in a research context that mimics a clinical hospital setting. Using these data, the predictive capability of PRSs can be assessed regarding many possible diagnoses that can accumulate during an individual's lifespan (ie, the phenome). The eMERGE (electronic MEdical Records and GEnomics) network is a nationwide consortium of multiple medical institutions that link DNA biobanks to EHRs (32), which is an important resource for determining the clinical utility of genomic findings, and enabling exploration of the range of phenotypes associated with genetic variation (33,34).

The aim of this study is to systematically examine the utility of PRSs derived from a GWAS meta-analysis by the International PCOS Consortium (11) for risk prediction across multiple ancestries and to further characterize the other EHR phenotypes that are clinically associated with PCOS genetic risk in both women and men. This meta-GWAS is the largest to date for PCOS in EA participants, with PCOS diagnosed according to NIH or Rotterdam criteria, or by self-report. Using its summary statistics, we first developed the integrative polygenic and phenotypic risk score (PPRS) for PCOS by combining the patient DNA genotype information and PCOS phenotypic elements from the EHR. Then we tested the predictive utility of the algorithm within EA samples and further evaluated its performance in AA and combined multi-ancestry (MA) participants which included EA, AA, and other ancestries. In addition, we performed a phenome-wide association study (PheWAS) of the PPRS for PCOS to identify the range of phenotypic indicators associated with PCOS and evaluated the predictive characteristics of the PPRS to identify underlying PCOS pathophysiological pathways.

Materials and Methods

PCOS polygenic risk score development

We obtained the full summary statistics of the largest meta-GWAS of PCOS through the International PCOS consortium and developed a PRS for PCOS (11,35) (all supplementary material and figures are located in digital research materials repositories). The GWAS was conducted in 5209 cases and 32 055 controls of EA women who were diagnosed according to either NIH or Rotterdam criteria. All variant positions were converted to NCBI Genome Reference Consortium Human Build 37 (GRCh37) positions and we excluded any entries with missing ORs or risk allele frequency (RAF) information. The RAF of each variant was calculated using PLINK (36), and we excluded the entries which RAF deviates more than 15% than our eMERGE data in order to ensure additional quality control (QC). PRSice software (37) was used to filter any correlated single nucleotide variants (SNVs) in pairwise LD ($r^2 > 0.2$) and constructed a PRS for PCOS by summing the best-guess imputed genotype data of PCOS risk variants in each individual weighted by the reported effect sizes. We used 8 different subsets of PCOS susceptibility SNVs to build the model based on P-value cut-off and compared their predictive accuracy in the following validation step: 5×10^{-8}, 5×10^{-7}, 5×10^{-6}, 5×10^{-5}, 5×10^{-4}, 5×10^{-3}, 5×10^{-2}, and 1 (all SNVs).

PRS/PPRS evaluation and PheWAS discovery cohort

Our cohort included genotypes and clinical diagnosis records of 99 185 individuals collected from 12 EHR-linked biobanks nationwide through the eMERGE consortium (33). After identity-by-descent (IBD) analysis, we removed 8019 related individuals that were not in canonical IBD position or genetically identical individuals near the origins (Z0 > 0.83 and Z1 < 0.1). The cohort was composed of multiple self-reported and 3rd party observed ancestries and we defined them into 3 main genetic ancestral groups using the intersection of self-reported ancestries and principal component analysis (PCA) based k-mean clusters: European (71.7%), African (15.0%), and Asian (1.0%). We excluded any self-reported or genetically Hispanic participants for ancestry-stratified analysis for better homogeneity. Throughout this study, the first 4 principal components (PCs) were used to correct population structure, explaining over 17% of the variances among different genetic origins.

The phenome data of the participants were collected from the EHR including diagnostic records and basic demographic information. The data collection was performed

under local institutional review board approval with informed consent from the patients. Diagnostic information was structured in the format of the International Classification of Diseases, Clinical Modification (ICD-CM) codes, in both the 9th and the 10th editions, and aggregated into a hierarchy of 1711 phenotype codes (phecodes) for a standardized categorical analysis of diseases (Phecode map version 1.2) (38,39). For example, ulcerative colitis (ICD-9 556.0), left-sided ulcerative colitis (ICD-9 556.5), and unspecified ulcerative colitis (ICD-9 556.9) are collated into a single phecode (555.2) for "ulcerative colitis"(40). We excluded 23 individuals under the age of 14, the clinically plausible age for PCOS diagnosis, which is defined as 2 years after the first menstruation. A demographic information of the 91 144 participants after filtering criteria is presented in Table 1.

Genotype data and quality control

The participants provided saliva or blood samples for genotyping, which were genotyped on 78 genotype Illumina or Affymetrix array batches from 12 medical sites. We used the Michigan Imputation Server (41) with the minimac3 missing genotype variant imputation algorithm to impute missing genotypes in our sample based on the Haplotype Reference Consortium (HRC1.1) which includes ~65 000 individuals of diverse ancestry (42). The imputation resulted in a genome-wide set of ~40 million SNVs. We filtered the poorly imputed genetic variants with the r-squared imputation quality threshold (mean variant r-square) less than 0.3, minor allele frequency (MAF) less than 0.05 and genotype call rate lower than 95%, which resulted in 5 760 270 autosomal polymorphic variants for subsequent analysis. The detailed data collection and QC report for the eMERGE network is reported elsewhere (33).

Validation of polygenic risk score

Predictive ability of each prediction model with different PRSs.
We performed logistic regression analysis to demonstrate the prediction ability of the PRS for PCOS diagnosis in the female population of 3 different genetic ancestry cohorts: European (n = 33 869), African (n = 8198), and the entire mixed cohort (n = 49 365). Each cohort was randomly divided into 75% training and 25% testing set to separately calculate the regression statistics and out-of-sample prediction error. Using a generalized linear model, the residuals of PRS after covariate adjustments (first 4 PCs, sites) were obtained and scaled to build the logistic regression model in the training set. Regression coefficients and P-value of the PRS variable and pseudo-R^2 of the 8 different PRS models were measured.

We applied the regression model built out of the training set to measure out-of-sample performance in the testing dataset. We predicted the individuals to be "PCOS cases" if their fitted scores were higher than the average fitted score and calculated the accuracy by comparing with their actual diagnosis records of PCOS. The overall accuracy, sensitivity, and specificity of each model were measured and structured through confusion matrix. The area under the receiving operating characteristic (ROC) curve, the AUC, was also measured for classifier performance of each model.

Stratification ability of each prediction model with different PRSs.
To evaluate the phenotypic stratification ability of the PRS, we divided the cohort into 10 quantiles based on the PRS of each individual and compared the average phenotypic values (eg, proportion of PCOS diagnosed patients, body mass index [BMI], PRS) among the groups. The proportion of PCOS patients in each quantile, average PRS values, and average BMI measures of each individual were analyzed. We also performed independent t-test to assess if the average PRS score differences between PCOS cases and controls were statistically significant.

Performance improvement by the PRS variable.
To estimate the performance of the PRS variable, we built a null regression model without the PRS variable for PCOS prediction (PRS model vs PRS null model). The incremental pseudo-R^2 values according to McFadden (43) were calculated between the PRS models and the null logistic regression with only the first 4 PCs and site variables. The analysis of variance (ANOVA) was performed to examine how significant PRS variable impacts the PCOS diagnosis prediction model compared to the null model.

PRS model:

$$\text{logit}(\text{PCOS diagnosis} = 1) = \beta 0 * \text{PRS} + \beta 1 * \text{Site} + \beta 2 * 4\text{PCs} + \beta 3$$

PRS null model:

$$\text{logit}(\text{PCOS diagnosis} = 1) = \beta 0 * \text{Site} + \beta 1 * 4\text{PCs} + \beta 2$$

Development of prediction algorithms with PRS and PCOS component phenotypes (PPRS)

We built an integrative PPRS with PRS and PCOS component phenotypes in the EHR to maximize the utility of PRS for risk prediction. Additional dichotomous phenotypic variables to each individual from their EHR diagnosis records: hirsutism (ICD9 code 704.1, ICD10 code L68.0), irregular menstruation (ICD9 code 626.4, ICD10 code N92.6), and female infertility (ICD9 code 627, ICD10 code N97.0) were selected, all of which are well-established clinical components of PCOS. A total 908 individuals with hirsutism, 4936 individuals with irregular menstruation, and 422 individuals with female infertility ICD diagnosis codes were identified in the eMERGE consortium database.

Firstly, the logistic regression adjusted for first 4 PCs and sites were examined for their effect coefficients and variable P-values. Psuedo-R^2 of each model was calculated for measuring the improvement over the normal PRS model. ANOVA between the integrative model and normal PRS model were examined.

PPRS model:

$$\text{logit}(\text{PCOS diagnosis} = 1) = \beta 0 * \text{PRS} + \beta 1 * \text{Site} + \beta 2 * 4\text{PCs} + \beta 3 * \text{Hirsutism} + \beta 4 * \text{Irregular menstruation} + \beta 5 * \text{Female infertility} + \beta 6$$

PPRS null model:

Table 1. Demographic and clinical characteristics of discovery cohorts (eMERGE) and replication cohort (BioVU)

Site[a]	N subjects	Sex (female)	Ancestry (EA)	Ancestry (AA)	Age average	Age SD	BMI average	BMI SD	PCOS cases	Hirsutism cases	Irregular menses cases	Female infertility cases
BSCH[b]	862	362 (42.2%)	623	74	N/A	N/A	N/A	N/A	2	5	20	0
CCHMC[b]	5385	2320 (43.2%)	4058	523	8.9	6.7	20.9	6.2	11	24	54	2
CHOP[b]	9528	4376 (46.0%)	4898	4105	9.8	5.3	21.1	6.2	47	39	205	2
Columbia	2029	989 (48.8%)	519	143	56.1	19.8	27.0	5.4	3	4	15	1
Geisinger	2785	1320 (47.5%)	2439	8	62.8	15.7	32.6	8.1	77	48	158	8
Harvard	23 922	13 135 (55.0%)	20 727	1343	55.3	16.5	28.3	5.8	417	322	2284	217
KPW/UW	3225	1829 (56.7%)	2891	109	76.1	8.9	26.4	4.8	2	25	10	18
Mayo Clinic	9307	4672 (50.2%)	6680	17	61.7	15.4	29.3	5.8	48	85	217	17
Marshfield	3725	2255 (60.9%)	3696	2	69.3	11.0	29.6	6.0	6	84	476	43
Mt. Sinai	5765	3362 (58.8%)	702	3643	59.6	10.0	30.6	7.4	51	45	200	15
Northwestern	4719	3913 (82.9%)	2250	301	53.7	14.8	28.7	7.2	65	83	280	51
Vanderbilt[c]	19 892	10 810 (54.4%)	15 902	3371	56.6	17.1	29.4	7.1	220	144	1017	48
All (Discovery cohort)	91 144	49 343	65 385	13 639					949	908	4936	422
VUMC replication sample	33 708	18 096 (54%)	33 708	N/A	55.7	20	28.2	6.8	284	225	4330	48

[a]BSCH = Boston Children's Hospital, CCHMC = Children's Hospital of Philadelphia, KPW/UW = Group Health Cooperative/University of Washington.
[b]Children's hospital with low average age.
[c]No sample overlap with replication cohort (BioVU).

$$\begin{aligned} \text{logit(PCOS diagnosis} = 1) = {}& \beta 0 * \text{Site} + \beta 1 * 4\text{PCs} \\ & + \beta 2 * \text{Hirsutism} \\ & + \beta 3 * \text{Irregular menstruation} \\ & + \beta 4 * \text{Female infertility} + \beta 5 \end{aligned}$$

Phenome-wide analysis

To investigate the potential pleiotropy of PCOS, PCOS components, and other diseases in the EMR phenome, we selected the best performing PRS model that presented a validated predictive accuracy and stratification ability across ancestries based on the examination results above. PheWAS was performed on the mapped 1711 representative EHR phenotypes with a minimum of 30 case patients from the discovery cohort of 91 144 participants after QC criteria. Case group for a given phecode is defined by the presence of at least 1 assignment of the corresponding ICD codes from EHR as defined in the phecode map v1.2. Controls for each phecode are defined by the absence of the same ICD codes that defined cases and the absence of clinically related phenotypes. Based on the assumption that a participant with higher PCOS-PRS conveys greater genetic risk, our main sex-stratified PheWAS interrogated the comorbid networks of high-risk predictive phenotypes for PCOS (PheWAS-1). A total of 49 343 female participants and 41 669 male participants were used for the analysis. Logistic regression was used adjusting for genotyping site and the first 4 PCs of ancestry to correct for population stratification in the MA cohort [logit (Clinical Phenotype = 1 | PRS, Site, 4PCs) = β0 + β1*PRS + β2*Site + β3*4PCs].

In this study, phenome-wide significance refers to either (1) the Bonferroni corrected threshold of $P = 2.9 \times 10^{-5}$ adjusting for multiple testing, which is determined by using $P = .05$ divided by the 1711 phenotypes interrogated, or (2) the false discovery rate significance of 0.05, which is a popular alternative threshold to the stringent Bonferroni correction in reporting PheWAS. Manhattan PheWAS plots of $-\log10(P\text{-value})$ were generated for visual inspection of significant clinical traits. All the analyses were performed in the R statistical software environment (ver 2.1.2).

Sensitivity analysis

We performed several comparative PheWAS in an effort to interrogate different phenome-wide aspects of the PRS in the clinical phenome.

Firstly, to distinguish secondary or symptomatic phenotypes derived from the PCOS-diagnosed patients, we removed the clinical diagnosis records of the 949 individuals with PCOS (phecode 256.4, ICD9 256.4, and ICD10 E28.2) and performed the same PheWAS analysis (PheWAS-2). Additionally, to gauge the contrasting performance of polygenic prediction over a single-variant approach, we performed traditional PheWAS of each genome-wide significant susceptibility loci ($P < 5 \times 10^{-8}$) for PCOS (RAF > 0.05). This analysis aims to compare the clinical phenotypes associated with the cumulative effects of multiple genetic variants on PRS versus a single genetic signal generated by an individual PCOS susceptibility locus. Among 113 genome-wide significant loci ($P < 5 \times 10^{-8}$) for PCOS (see (35)), we filtered the entries with MAF > 0.05 and genotype call rate >0.90 in our discovery cohort and MAF > 0.01 in summary statistics. In total 85 SNVs were selected and used for the subsequent PheWAS analysis (PheWAS-3).

PRS PheWAS replication

To confirm the predictive performance of our PRS algorithm and its effect on clinical phenome, replication analyses were performed at Vanderbilt University Medical Center on an independent genotyped sample of 33 708 European descent individuals (BioVU). The participants were genotyped on the Illumina MEGAEX platform (~2 million markers) and we applied filters for individual call rates <98%, batch effects ($P < 5 \times 10^{-5}$), heterozygosity ($|F^{het}| > 0.2$), and sample relatedness (pihat > 0.2). After imputation with the 1000G reference panel, we excluded any genetic variants with missingness >0.02, certainty <0.9, or imputation info score <0.95. The genetic ancestry of the samples were restricted to only EA, based on comparison with the 1000G European population and a K-means clustering definition. The final samples included 33 708 individuals of European descent genotyped on 5 550 390 SNVs. Using the same PRS generation methodology in discovery samples, we took the identical phenome-wide approach to identify the associated phenotypic networks with PRS among the replication samples. Logistic regression was used adjusting for first 4 ancestry PCs.

Results

Polygenic risk scores for PCOS are normally distributed in European and multi-ancestry participants

In total 5 760 270 autosomal SNVs were considered for the PCOS-PRS construction, which displays the genetic architecture of effect size (beta) by RAF presented in Fig. 1. There was a significant negative correlation between RAF and effect size, which is generally anticipated in common quantitative traits and supports the use of methodology of PRS to explore the extreme of the common polygenic liability spectrum. According to the central limit theorem, PRSs in a large population will show normality when the genetic architecture of the target trait is polygenic, that is, produced by the addition of many genetic variants of small effect (44,45).

PRSs were calculated at 8 different P-value cut-offs from the PCOS GWAS summary statistics (5×10^{-8}, $5 \times 10^{-7}, 5 \times 10^{-6}, 5 \times 10^{-5}, 5 \times 10^{-4}, 5 \times 10^{-3}, 5 \times 10^{-2}, 1$) for all the discovery eMERGE participants (n = 91 144). Each set of scores was adjusted for participant site and the first 4 PCs. All the polygenic scores were evaluated for their predictive performance in the female populations of EA (n = 33 869), AA (n = 8198), and MA cohorts (n = 49 365). The covariate-adjusted PCOS-PRS generally presented a normal distribution across each ancestry cohort (see (46)). PRS models with trimodal or skewed distributions (PRS P-value cut-off: 5×10^{-7}, 5×10^{-6}, 5×10^{-5}), which may be a function of poor representation of risk variants across populations, were not considered for the subsequent phenome-wide analysis.
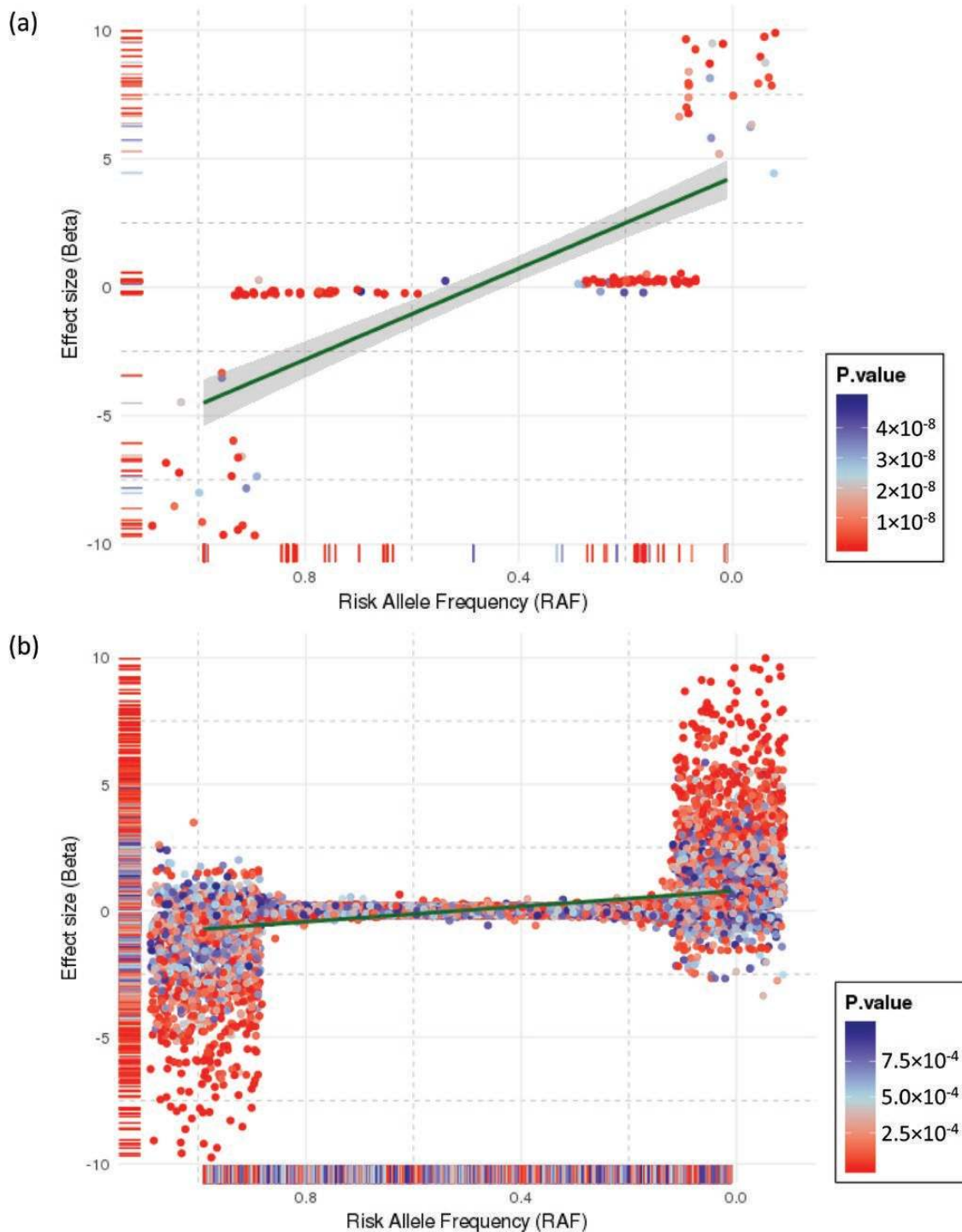
Figure 1. Effect distribution of PCOS susceptibility variants in samples from the International PCOS consortium by risk allele frequency. (A) The 139 PCOS genome-wide significant SNVs ($P < 5 \times 10^{-8}$), and (B) the 120 340 PCOS autosomal SNVs with $P < 0.05$. The dark green line and grey band around it are the linear regression fit and its 95% confidence interval, respectively, between risk allele frequency and effect size (beta).

Validation of PCOS PPRS in European ancestry participants

Predictive ability of each prediction model with different PRSs. In the PRS prediction models using the training set of the female EA cohort (n = 33 869 with 632 PCOS cases), all the coefficient P-values of the PRS variables are statistically significant except for 2 PRS models of SNVs with $P < 5 \times 10^{-7}$ and $P < 5 \times 10^{-6}$ that do not show PRS normality (see (46)). The average odds ratios (ORs) of the significant PRS variable across EA was 1.13 (95% confidence interval [CI 1.04–1.22]) and the average pseudo-R^2 value was 0.044, which indicates 4.4% of the phenotypic variances in the training sample could be explained by PRS (Table 2).

Table 2. Regression results of the PRS/PPRS models in PCOS prediction

PRS/PPRS P-value cut-off	PRS model[a]				PPRS model[b]			
	OR	95% CI	P	R^2	OR	95% CI	P	R^2
EA								
5×10^{-8}	1.14	(1.04–1.25)	4.76×10^{-3}	0.045	1.14	(1.03–1.26)	1.40×10^{-2}	0.232
5×10^{-7}	1.04	(0.95–1.12)	3.78×10^{-1}	0.043	1.04	(0.94–1.13)	3.89×10^{-1}	0.230
5×10^{-6}	1.08	(0.99–1.16)	6.41×10^{-2}	0.044	1.08	(0.99–1.17)	7.26×10^{-2}	0.231
5×10^{-5}	1.10	(1.01–1.19)	2.13×10^{-2}	0.044	1.10	(1.00–1.20)	3.59×10^{-2}	0.231
5×10^{-4}	1.13	(1.03–1.23)	6.12×10^{-3}	0.044	1.11	(1.01–1.22)	2.85×10^{-2}	0.231
5×10^{-3}	1.11	(1.01–1.22)	2.70×10^{-2}	0.044	1.08	(0.97–1.19)	1.45×10^{-1}	0.231
5×10^{-2}	1.16	(1.06–1.27)	2.11×10^{-3}	0.045	1.12	(1.01–1.24)	3.21×10^{-2}	0.231
1	1.15	(1.05–1.27)	3.13×10^{-3}	0.045	1.11	(1.00–1.23)	4.04×10^{-2}	0.231
MA								
5×10^{-8}	1.16	(1.07–1.25)	1.15×10^{-4}	0.040	1.15	(1.06–1.24)	1.19×10^{-3}	0.228
5×10^{-7}	1.08	(1.00–1.16)	4.28×10^{-2}	0.038	1.09	(1.01–1.17)	2.99×10^{-2}	0.227
5×10^{-6}	1.09	(1.02–1.17)	1.60×10^{-2}	0.038	1.10	(1.02–1.18)	1.19×10^{-2}	0.227
5×10^{-5}	1.12	(1.04–1.20)	2.35×10^{-3}	0.039	1.12	(1.04–1.20)	3.67×10^{-3}	0.228
5×10^{-4}	1.12	(1.04–1.21)	1.88×10^{-3}	0.039	1.11	(1.03–1.20)	8.59×10^{-3}	0.228
5×10^{-3}	1.16	(1.08–1.25)	1.25×10^{-4}	0.040	1.13	(1.04–1.23)	2.54×10^{-3}	0.228
5×10^{-2}	1.20	(1.11–1.29)	5.03×10^{-6}	0.041	1.16	(1.07–1.26)	3.81×10^{-4}	0.228
1	1.22	(1.13–1.32)	5.33×10^{-7}	0.041	1.19	(1.09–1.29)	5.91×10^{-5}	0.229
AA								
5×10^{-8}	1.14	(0.96–1.36)	1.42×10^{-1}	0.031	1.15	(0.95–1.39)	1.62×10^{-1}	0.211
5×10^{-7}	1.24	(1.05–1.47)	1.22×10^{-2}	0.034	1.30	(1.08–1.56)	4.63×10^{-3}	0.215
5×10^{-6}	1.25	(1.05–1.48)	9.80×10^{-3}	0.034	1.30	(1.09–1.56)	3.95×10^{-3}	0.216
5×10^{-5}	1.23	(1.04–1.45)	1.71×10^{-2}	0.034	1.27	(1.05–1.52)	1.08×10^{-2}	0.214
5×10^{-4}	1.19	(1.00–1.42)	4.38×10^{-2}	0.032	1.17	(0.97–1.40)	9.82×10^{-2}	0.211
5×10^{-3}	1.18	(0.99–1.41)	6.74×10^{-2}	0.032	1.18	(0.97–1.43)	9.32×10^{-2}	0.211
5×10^{-2}	1.25	(1.05–1.50)	1.23×10^{-2}	0.034	1.17	(0.97–1.42)	1.07×10^{-1}	0.211
1	1.30	(1.09–1.56)	3.33×10^{-3}	0.036	1.26	(1.05–1.53)	1.56×10^{-2}	0.214

Average of the credible models (coefficient P < .05)

PRS model	Average OR	Average R^2	PRS null model[c] R^2	Incremental R^2 over PRS null model
EA	1.13	0.044	0.004	0.041
MA	1.14	0.039	0.004	0.036
AA	1.25	0.034	0.004	0.030

PPRS model	Average OR	Average R^2	PPRS null model[d] R^2	Incremental R^2 over PPRS null model[e]	Incremental R^2 over PRS null model
EA	1.12	0.231	0.193	0.038 (19.6%)	0.228 (61.5-fold)
MA	1.13	0.228	0.201	0.027 (13.2%)	0.224 (58.8-fold)
AA	1.28	0.215	0.193	0.021 (11.0%)	0.211 (57.0-fold)

Abbreviations: OR, odds ratio; CI, confidence interval; R^2, psuedo-R^2.
[a]PRS: PRS + basic covariates (model1 = PCOS ~ PRS + PC1-4 + site).
[b]PPRS: PRS + PCOS component phenotypes + basic covariates (PPRS = PCOS ~ PRS + PC1-4 + site + hirsutism + female infertility + irregular menses).
[c]PRS null model: basic covariates only (null model = PCOS ~PC1-4 + site).
[d]PPRS null model: PCOS component phenotypes + basic covariates (PPRS null model = PCOS ~ PC1-4 + site + hirsutism + female infertility + irregular menses).
[e]Improvement rate: (incremental change in pseudo-R^2 between the model with PRS/PPRS and the null model without PRS/PPRS)/(pseudo-R^2 of the null model without PRS/PPRS).

The regression models built in the training set were then used to predict PCOS case status in the testing dataset. A model including PRS yielded average prediction accuracy of 0.55, sensitivity of 0.55, specificity of 0.76 with an average AUC of 0.72 in the EA participants (Table 3).

Stratification ability of each prediction model with different PRSs. The percentage of PCOS-diagnosed patients increases in higher PRS quantiles, and the individuals in the highest PRS group tend to have higher average BMI. In the genome-wide PRS calculation with SNVs with $P \leq 1$, the average BMI

Table 3.   Average performance of PCOS PRS/PPRS prediction algorithms in the female cohorts of European (n = 33 869), multi-ancestry (n = 49 365) and African (n = 8198) participants

Summary—Average

PRS model[a]	Accuracy	Sensitivity	Specificity	Balanced accuracy	AUC[c]
EA (n = 33 869)	0.551	0.547	0.755	0.651	0.715
MA (n = 49 365)	0.533	0.529	0.736	0.632	0.693
AA (n = 8198)	0.496	0.494	0.590	0.542	0.543
PPRS model[b]					
EA (n = 33 869)	0.873	0.876	0.717	0.797	0.870
MA (n = 49 365)	0.881	0.886	0.640	0.763	0.823
AA (n = 8198)	0.864	0.872	0.522	0.697	0.706

[a]PRS model: PRS + basic covariates model (model1 = PCOS ~ PRS + PC1-4 + site + sex).
[b]PPRS model: PRS + PCOS component phenotypes + basic covariates model (model2 = PCOS ~ PRS + PC1-4 + site + sex + hirsutism + female infertility + irregular menses).
[c]Area under the curve.

of the individuals in highest PRS quantile is 1.1 kg/m^2 higher than the individuals in the lowest PRS group (Cohen's d = 0.16, t-test $P = 1.06 \times 10^{-9}$) (Fig. 2 and Table 4). The finding confirms the positive correlation between elevated generic risk for PCOS, actual PCOS diagnosis, and higher risk for increased BMI. Adding in the count of hyperandrogenism phenotypes (N = 0, 1, 2, 3) did not substantially alter the stratification (data not shown).

The subsequent t-test reveals that PRSs of case patients are significantly higher than the controls in all the nominally significant PRS models with regression $P < 0.05$, implying that higher genetic risk scores indicate higher occurrence of PCOS diagnosis ($P = 2.15 \times 10^{-4}$, 7.75×10^{-4}, 2.43×10^{-4}, 2.51×10^{-5}, 3.12×10^{-5} in PRS model SNVs $P <5 \times 10^{-8}$, 5×10^{-4}, 5×10^{-3}, 5×10^{-2}, 1, respectively) (see (47)).

Performance improvement by the PRS variable. All the PRS models containing PCOS-PRS provided an improved fit over the null model by increasing the estimated explained sum of squares (pseudo-R^2) according to McFadden (43). The average increase of pseudo-R^2 by the PRS variable in EA samples is 0.040, which is a 10-fold improvement (=0.040/0.004) over the null model. The ANOVA P-values of differentiating the PRS models from the null model are all under 1×10^{-31}, which validate the statistical significance of the performance improvement over the null model (Table 2 and (48)).

Evaluation of PRS in multi-ancestry and African ancestry participants

Predictive ability of each prediction model with different PRSs. In the training set of the MA cohort (n = 49 365 with 949 PCOS cases), the coefficient

P-values of all PRS variables remain significant with positive beta coefficients (Table 2; model1). The average OR of PRS is 1.14 (95% CI 1.07–1.21) and the average pseudo-R^2 value is 0.039, indicating that 3.9% of the phenotypic variance in the MA cohort could be explained by the PRS model. In the training set of AA participants (n = 8198 with 172 PCOS cases), the coefficient P-values of PRS variables remain overall significant except for 2 PRS models of SNVs with $P <5 \times 10^{-8}$ and $P <5 \times 10^{-3}$ which may be due to the smaller sample size. Even though the regression P-values of the PRS variable do not show uniform performance in AA compared with EA, the nominally significant PRS models generate a higher effect size in the AA samples than in the other ancestry groups. The average OR of PRS models in the AA is 1.25 (95% CI 1.08–1.42), higher than both the EA (OR 1.13, 95% CI 1.04–1.22) and MA (OR 1.14, 95% CI 1.07–1.21). This is possibly due to the low RAF of PCOS risk variants in AA compared with EA (see (35)).

For the testing dataset, PRS prediction displays an average 0.533 of accuracy, 0.529 of sensitivity, 0.736 of specificity with an average AUC of 0.693 in the MA cohort. The out-of-sample performance in AA yielded an average AUC of 0.543 and showed an overall lower average accuracy (0.496), sensitivity (0.494), and specificity (0.590) than other ancestry groups (Table 3).

Stratification ability of each prediction model with different PRSs. In the MA cohort, the proportion of PCOS patients increases from 1.5% in the lowest quantile to 2.6% in the highest quantile in the PRS calculation of SNVs with $P \le 1$. The average BMI of the participants in the highest PRS quantile is 1.2 kg/m^2 higher (Cohen's d = 0.17, t-test $P = 1.62 \times 10^{-13}$) than the participants in the lowest PRS group (Fig. 2B and (49)).
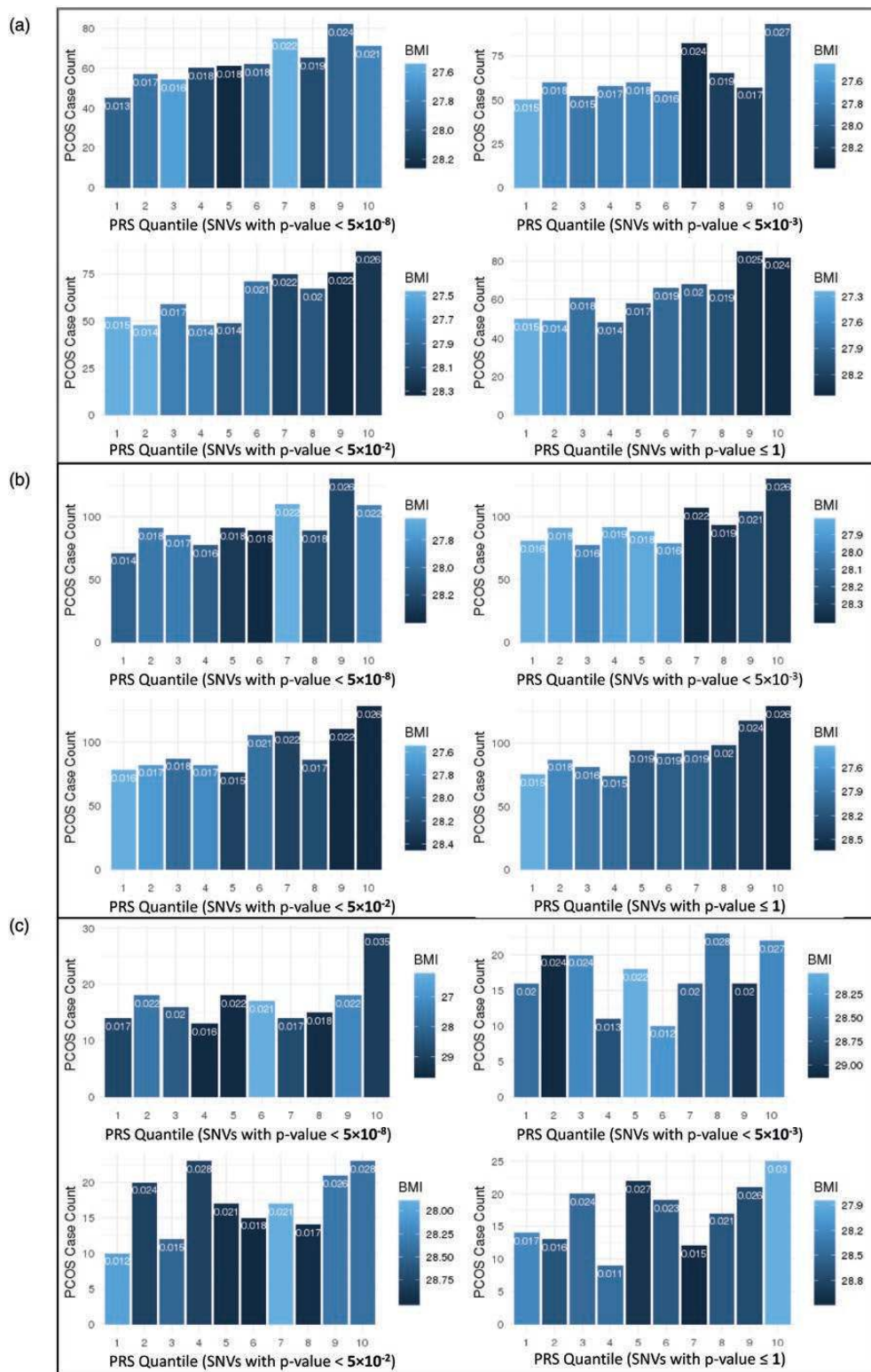
Figure 2. Stratification performance by quantile of PRS models, including PCs 1–4 and site as covariates, in (A) EA, (B) MA, and (C) AA populations. Group 1 includes those with the lowest PRS, and group 10 includes those with the highest. Bar colors indicate the average BMI in the quantile (darker indicates higher BMI), while the proportion of PCOS-diagnosed patients in each group is indicated at the top of each bar.

In the AA cohort, the number of PCOS patients does not always increase with higher PRS quantile, but the observation of an excess of PCOS patients in the highest PRS quantile is generally consistent across the models (Fig. 2c). In the full-inclusive PRS model (SNVs with $P \leq 1$), the rate of PCOS patients increases from 1.7% in the lowest quantile to 3.1% in the highest PRS quantile (see (49)). The observed increase in the rate of PCOS patients is most pronounced in the PRS model with genome-wide significant variants (SNVs with $P < 5 \times 10^{-8}$), as the

Table 4.   Quantile analysis of PCOS PRS in the female European cohort (n = 33 869) (PRS with SNVs $P \leq 5{\times}10^{-8}$ and $P \leq 1$ only)

	Group[a]	PCOS cases	PCOS prop[b](%)	Average BMI(kg/m^2)	Average PRS
PRS with SNVs $P \leq 5{\times}10^{-8}$	1	45	1.3	27.9	−1.750
	2	57	1.7	27.9	−0.950
	3	54	1.6	27.6	−0.813
	4	60	1.8	28.1	−0.239
	5	61	1.8	28.2	−0.068
	6	62	1.8	28.0	0.014
	7	75	2.2	27.6	0.248
	8	65	1.9	28.1	0.810
	9	82	2.4	27.9	0.952
	10	71	2.1	27.8	1.790
PRS with SNVs $P \leq 1$	1	50	1.5	27.3	−1.790
	2	49	1.5	27.5	−1.020
	3	61	1.8	27.8	−0.654
	4	48	1.4	27.9	−0.369
	5	58	1.7	28.0	−0.113
	6	66	2.0	27.8	0.132
	7	68	2.0	28.0	0.386
	8	65	1.9	28.1	0.672
	9	85	2.5	28.4	1.040
	10	82	2.4	28.4	1.720

PCOS-PRS is adjusted with covariates and scaled for standardization.
[a]Higher group number indicates higher PCOS PRS.
[b]Proportion of PCOS case patients in the quantile.
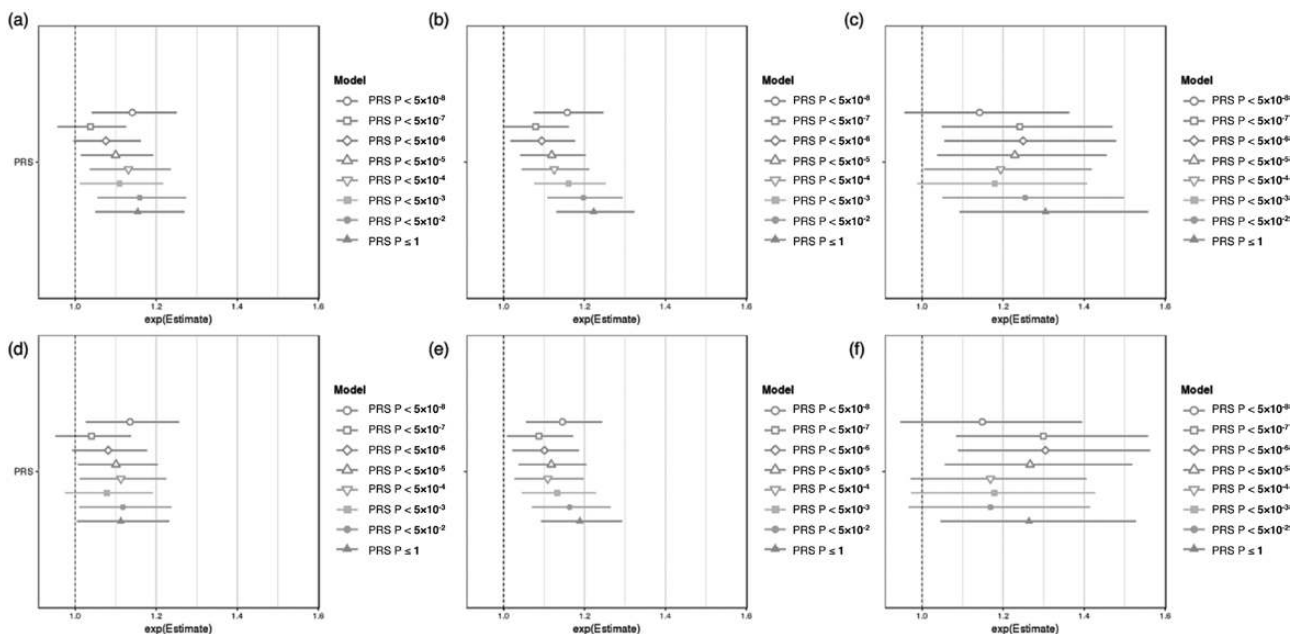
Figure 3. Comparison of odds ratios (ORs) for the PRS and PPRS in (A/D) EA, (B/E) MA, and (C/F) AA cohorts, at different PRS/PPRS inclusion thresholds by GWAS *P*-value. The top row shows OR distributions for the PRS model, which adjusted for basic covariates (PRS model = PCOS ~ PRS + PC1–4 + site). The bottom row shows OR distributions for the PPRS model which adjusted for the same basic covariates as well as PCOS EHR component phenotypes (PPRS model = PCOS ~ PRS + PC1-4 + site + hirsutism + female infertility + irregular menses).

PCOS case rate doubles from 1.7% in the lowest quantile to 3.5% in the highest PRS quantile. We did not identify any notable trends in BMI in AA participants, which is depicted by the quantile color changes in Fig. 2C.

An independent t-test confirms the significant differences of average PRSs between PCOS cases and controls in MA across the models. The PRS difference between PCOS MA cases and controls is 0.165 after scaling with a full-inclusive PRS model, SNVs with $P \leq 1$ (Cohen's d = 0.201, t-test $P = 2.62 \times 10^{-6}$). In AA, only the full-inclusive PRS model shows statistically significant difference between PCOS cases and controls with

a PRS difference of 0.175 (Cohen's d = 0.191, t-test $P = 2.90 \times 10^{-2}$) (see (47)).

Performance improvement by the PRS variable. In the joint ancestry participants, all the prediction models containing the PRS variable provide a better fit over the null model by increasing the average pseudo-R^2 to 0.039, which is an 8.75-fold increase (=0.035/0.004) in explanatory power (Table 2). The subsequent ANOVA analysis confirms the statistical significance of the improved fits over the null model with all $P < 1 \times 10^{-46}$ (48).

In the AA samples, the statistically significant PRS models show the average pseudo-R^2 of 0.034, which has the poorest fit among the ancestries. The models show an average pseudo-R^2 improvement of 7.5-fold increase (=0.030/0.004) from the null model without PRS (Table 2). Even with the lowest average incremental pseudo-R^2 (0.030) among the ancestries, the significant difference between the PRS models and the null model in Africans are confirmed with all ANOVA $P < 5 \times 10^{-3}$ (see (48)).

Development of PPRS prediction algorithms with PRS and PCOS component phenotypes

The addition of PCOS component EHR phenotypes to polygenic risk prediction significantly improved the predictive accuracy (Table 2; model2 and Fig. 3). The average pseudo-R^2 of the PPRS is 0.231 in EA, 0.228 in MA, and 0.215 in AA samples, which indicates an average 14.7% improvement in pseudo-R^2 (19.6% in EA, 13.2% in AA, 11.0% in MA) over the PPRS null model by the inclusion of PCOS component phenotypes. Compared to the basic null model, the PPRS prediction boosts the average predictive performance (pseudo-R^2) by approximately 60 times (61.5-fold in EA, 58.8-fold in AA, 57.0-fold in MA) by the combinational use of PCOS component EHR phenotypes and PRSs. Of note, the PRS variable's P-values in every PPRS model remain consistently valid in the MA samples ($P < 5 \times 10^{-3}$), whereas it was not always significant in AA or even EA samples. The ORs of the PRS and PPRS remain similar across the ancestries and Delong tests (50) confirmed the statistical significance of the difference among the AUC of ROC curves between PRS and PPRS models (Fig. 4).
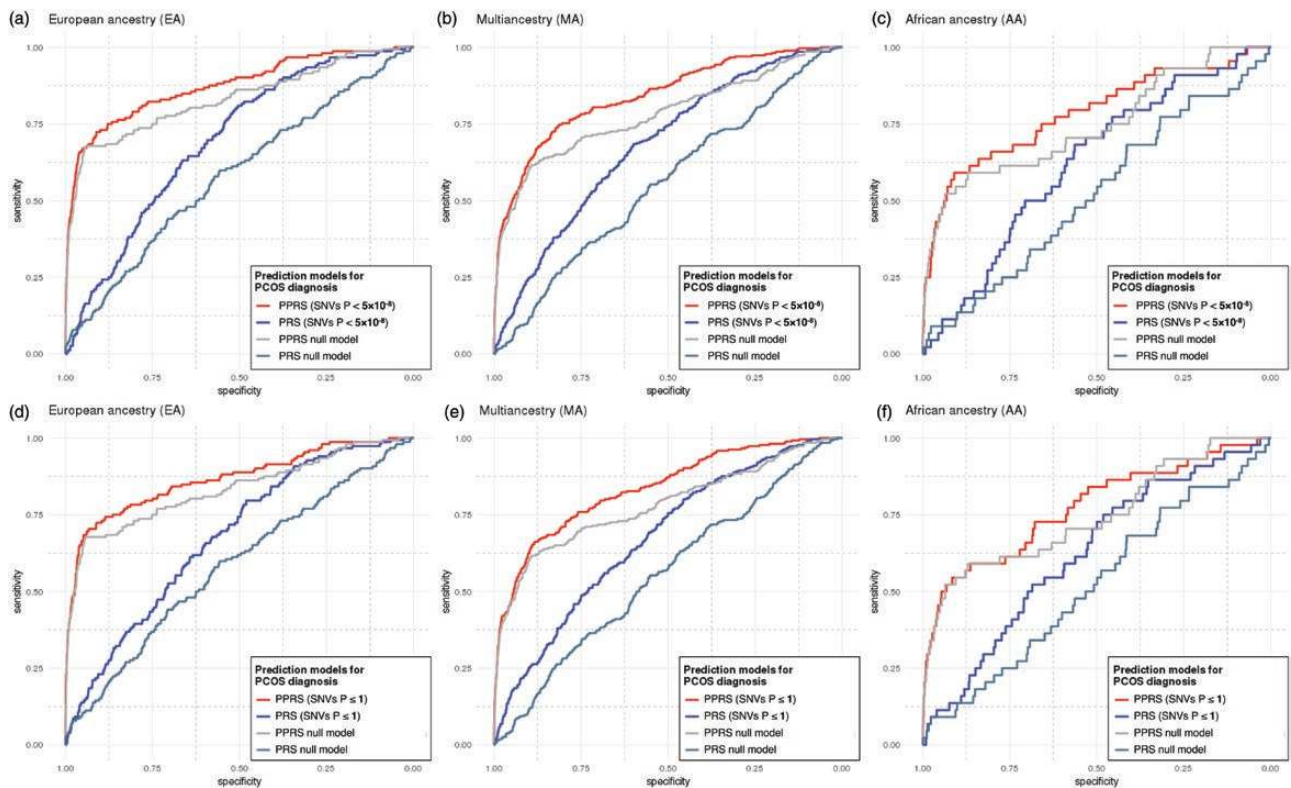
Figure 4. Comparison of receiving operating curves (ROC) of the PPRS and PRS prediction models for PCOS diagnosis. The models with the genome-wide significant SNVs ($P < 5 \times 10^{-8}$) were evaluated in females of (A) EA, (B) MA, and (C) AA cohorts, along with the full-inclusive prediction models ($P \leq 1$) in females of (D) EA, (E) MA, and (F) AA cohorts. The areas under the curve (AUC) are provided in Table 2 and (47). PRS model adjusted for basic covariates (PRS model = PCOS ~ PRS + PC1–4 + site), and PPRS model adjusted for the same basic covariates as well as PCOS EHR component phenotypes (PPRS model = PCOS ~ PRS + PC1–4 + site + hirsutism + female infertility + irregular menses). Null models only included the basic covariates without the PRS variable. Additional Delong test confirmed the statistical significance of the difference among the AUC of ROC curves between PRS and PPRS models: (A) Z = −8.29, $P < 2.20 \times 10^{-16}$, (B) Z = −7.41, $P = 1.27 \times 10^{-13}$, (C) Z = −2.85, $P = 4.31 \times 10^{-3}$, (D) Z = −8.3291, $P < 2.20 \times 10^{-16}$, (E) Z = −8.18, $P = 2.91 \times 10^{-16}$, (F) Z = −3.7523, $P = 1.75 \times 10^{-4}$.

The subsequent ANOVA tested that all the pairs between PPRS and PPRS null models were statistically distinct across the cohorts and every PPRS model show the improved fit over the PPRS null model (see (48)). The average ORs of irregular menstruation (ICD9 code 626.4, ICD10 code N92.6), female infertility (ICD9 code 627, ICD10 code N97.0), and hirsutism (ICD9 code 704.1, ICD10 code L68.0) for PCOS prediction were, as expected, strong across the cohorts: 5.51 (95% CI 4.22–7.18), 10.9 (95% CI 6.44–18.30), and 17.1 (95% CI 12.11–24.19), respectively (see (51)).

Clinical phenome analysis

Associated phenotypes with PRS (PheWAS-1). The general scheme of our PheWAS analyses are depicted in Fig. 5A. Based on the model examination described above, the genome-wide PRS that includes all SNVs with $P \leq 1$ was selected as the best performing PRS model and used for phenome-wide analysis. The phenomes of 49 343 female participants and 41 669 male participants were analyzed separately to test for association with high genetic risk for PCOS.

In the female PheWAS with PRS, 75 EHR phenotypes were identified with phenome-wide significance (Fig. 5B and (52)). "Morbid obesity" (phecode 278.11) and obesity-related endocrine phenotypes, including "overweight, obesity, and other hyperalimentation" (phecode 278), "type 2 diabetes" (phecode 250.2), "essential hypertension" (phecode 401.1), "hypercholesterolemia" (phecode 272.11), "hypertension" (phecode 401), "disorders of lipid metabolism" (phecode 272) are the top-ranked. The phenome-wide significant association of "polycystic ovaries" (phecode 256.4) and PCOS-PRS are observed with one of the largest effect sizes (OR = 1.015) among the result.

As a complex endocrine disorder, the PCOS pathophysiology seems to be tightly linked to the expression of endocrine or circulatory system manifestation. Among the 75 phenome-wide significant traits with PRS, the phenotypes of circulatory system (26.0%) and endocrine/metabolic system (21.0%) appeared the most frequently (Fig. 5D), while the 4 highest associated phenotypes are all endocrine/metabolic features.

Among the remainder of the phenome-wide significant phenotypes, associations of musculoskeletal phenotypes like "osteoarthrosis" (phecode 740 and 740.9) or "calcaneal spur; Exostosis NOS" (phecode 726.4) possibly imply the hormonal changes on the skeletal system impacted by PCOS epidemiology. Multiple symptomatic genitourinary phenotypes of PCOS were also identified: "abnormal mammogram" (phecode 611.1) or "other signs and symptoms in breast" (phecode 613.7).

An obesity-related pulmonary disorder of "sleep apnea" (phecode 327.3) was also observed (classified as neurological phenotype in phecode map) with "obstructive sleep apnea" (phecode 327.32). We could not identify
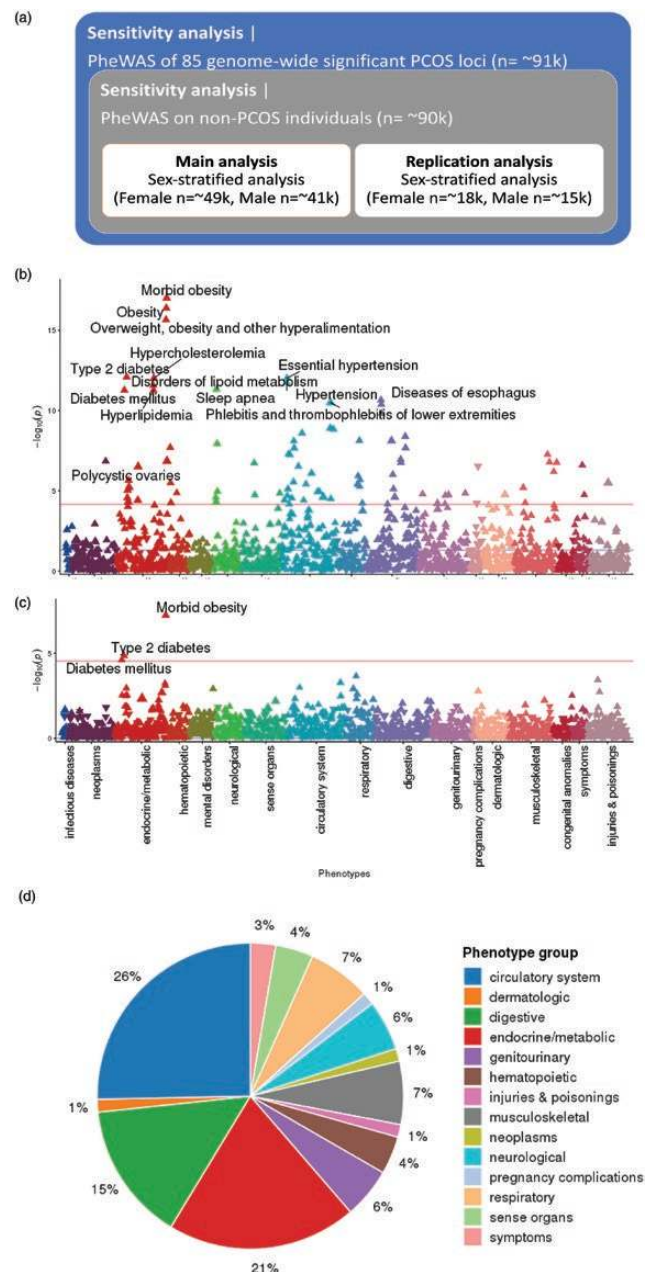

Figure 5. PheWAS scheme and results using PRS. (A) PheWAS scheme and sample sizes; (B) PheWAS Manhattan plot of PRS (SNVs with $P \leq 1$) in the phenomes of 49 343 female participants; (C) PheWAS Manhattan plot of PRS (SNVs with $P \leq 1$) in the phenomes of 41 669 male participants; (D) pie chart summarizing PheWAS groups. In Manhattan plots (B) and (B), the x-axis represents the EHR phenotype categorical group and the y-axis represents the negative log(10) of the PheWAS P-value. Red lines indicate the cut-off for phenome-wide significance. For readability, only the most significant associations are annotated. Full lists of phenome-wide significant results are provided in refs (51,52). The pie chart in (D) shows EHR categories for the 72 phenome-wide significant phenotypes identified through PheWAS of the genome-wide PRS (SNVs with $P \leq 1$).

Table 5.   (A) Sixteen significant phenotypes of PCOS-PRS ($P \leq 1$) female-stratified PheWAS that were phenome-wide significant in the discovery cohort replicated in the independent VU cohort (n=18 096). (B) Three phenome-wide significant results of PCOS-PRS (SNPs with $P \leq 1$) male-stratified PheWAS (n=41 669) and replication cohort (n = 15 612). None of them were replicated in the independent replication analysis.

(A) Discovery analysis

Group	OR	95% CI	P	n_total	n_cases	OR	95% CI
Endocrine/metabolic	1.010	(1.008–1.013)	9.74×10^{-18}	37 108	6790	1.116	(1.054–1.182)
Endocrine/metabolic	1.008	(1.006–1.009)	4.14×10^{-17}	44 267	13 949	1.087	(1.042–1.134)
Endocrine/metabolic	1.007	(1.005–1.009)	2.20×10^{-16}	47 803	17 485	1.077	(1.037–1.120)
Endocrine/metabolic	1.007	(1.005–1.009)	8.18×10^{-13}	42 874	10 800	1.081	(1.036–1.128)
Neurological	1.008	(1.006–1.010)	4.71×10^{-12}	40 673	6503	1.096	(1.036–1.158)
Endocrine/metabolic	1.007	(1.005–1.009)	5.39×10^{-12}	43 325	11 251	1.079	(1.035–1.125)
Digestive	1.008	(1.005–1.011)	4.17×10^{-9}	40 531	4582	1.093	(1.028–1.163)
Digestive	1.012	(1.008–1.016)	7.59×10^{-9}	47 803	2034	1.202	(1.079–1.339)
Neurological	1.007	(1.005–1.010)	1.16×10^{-8}	39 291	5121	1.098	(1.030–1.170)
Digestive	1.008	(1.005–1.011)	2.13×10^{-8}	40 251	4302	1.112	(1.042–1.187)
Musculoskeletal	1.005	(1.003–1.007)	1.71×10^{-7}	43 335	11 354	0.956	(0.915–0.999)
Endocrine/metabolic	1.015	(1.009–1.020)	3.16×10^{-7}	40 696	942	1.174	(1.026–1.343)
Musculoskeletal	1.004	(1.003–1.006)	6.38×10^{-7}	47 803	15 822	0.957	(0.923–0.993)
Endocrine/metabolic	1.008	(1.005–1.011)	2.25×10^{-6}	35 057	2983	1.136	(1.060–1.219)
Endocrine/metabolic	1.010	(1.005–1.014)	9.20×10^{-6}	33 663	1589	1.221	(1.082–1.377)
Genitourinary	1.004	(1.002–1.006)	1.40×10^{-5}	40 468	14 392	0.947	(0.910–0.986)

(B) Discovery analysis

Group	OR	95% CI	P	n_total	n_cases	OR	95% CI
Endocrine/metabolic	1.009	(1.006–1.012)	5.93×10^{-8}	32 456	3489	1.049	(0.978–1.125)
Endocrine/metabolic	1.005	(1.003–1.007)	1.41×10^{-5}	36 835	10 984	1.031	(0.989–1.076)
Endocrine/metabolic	1.005	(1.002–1.007)	2.47×10^{-5}	37 199	11 348	1.029	(0.988–1.073)

any psychological or depression related phenotype that is known to have genetic correlation with the hormonal changes of PCOS.

The overall low range of OR (1.004–1.010) of the PheWAS results should be noted, which is assumedly due to the aggregated effects of the low impact SNVs for PCOS, especially in the full-inclusive PRS with the entire GWAS SNVs. The ORs from the generic PheWAS of individual PCOS SNVs are observed to be higher before merging them into the cumulative PRS, which is described later (see (53)).

In the replication analysis on an independent cohort of 18 096 EA females (BioVU), 16 out of 75 phenome-wide signals from the discovery analysis are replicated including "PCOS" ($P = 1.93 \times 10^{-2}$, phecode 256.4) with the positive OR of 1.174 (Table 5a). Half of the replicated phenotypes (8 out of 16) belong to the endocrine/metabolic category. In particular, the following obesity-related endocrine phenotypes exhibit strong evidence of replication after multiple testing correction ($P < 6.7 \times 10^{-5}$, 0.05/75): "morbid obesity" (phecode 278.11), "obesity" (phecode 278.1), "overweight, obesity and other hyperalimentation" (phecode 278). The well-known comorbidity between "type 2 diabetes" (phecode 250.2) and PCOS is also identified along with other diabetic syndromes like "diabetes mellitus" (phecode 250). Other notable

replicated phenotypes included multiple neurological and digestive manifestations, which have well-known association with obesity, such as "chronic liver disease and cirrhosis" (phecode 571), "bariatric surgery" (phecode 539), and "other chronic nonalcoholic liver disease" (phecode 571.5). An obesity-related pulmonary disorder of "sleep apnea" (phecode 327.3) is also observed (classified as neurological phenotype in phecode map) with "obstructive sleep apnea" (phecode 327.32).

In male-specific PheWAS with PRS (SNVs with $P \leq 1$) model, 3 metabolic phenotypes reached phenome-wide significance in the discovery analysis, "morbid obesity" (phecode 278.11), "type 2 diabetes" (phecode 250.2), and "diabetes mellitus" (phecode 250), which are known risk factors and/or comorbidities for PCOS (Fig. 5B, Table 5B, and (52)). However, none of the associations was replicated in the replication analysis on 15 611 independent males. The replication sample is underpowered and larger sample sizes will be needed to distinguish these results from a true null result. As a result of the smaller sample size for the replication cohort with respect to that of the discovery cohort, several lead to the 95% CI spanning 1 in the replication cohort, and show less statistically significant associations. This can largely be attributed to the winner's curse phenomenon (54).

Sensitivity analysis—Case-excluded analysis (PheWAS-2). After removing 949 PCOS patients in PheWAS investigation, we still identified 68 PRS-phenotype associations that reached phenome-wide significance (see (55)), which is not very different from PheWAS-1. The result might be due to the challenge of current diagnosis practices in identifying PCOS cases, which implies the control groups are not completely excluding PCOS patients and possibly include some mixed signals from the unidentified PCOS cases. Alternatively, it is possible that genetic risk for PCOS remains a robust risk factor for these phenotypes even in the absence of clinical manifestations of PCOS.

The representative signals of diabetes/obesity-related endocrine traits that are identified in PheWAS-1 remained significant: "morbid obesity" (phecode 278.11), "type 2 diabetes" (phecode 250.2), "obesity" (phecode 278.1), "overweight, obesity and other hyperalimentation" (phecode 278), "diabetes mellitus" (phecode 250), "hypercholesterolemia" (phecode 272.11), "disorders of lipid metabolism" (phecode 272) and "hyperlipidemia" (phecode 272.1) etc.

Four phenotypes no longer remained phenome-wide significant in PheWAS-2 compared with PheWAS-1, including "menopausal and postmenopausal disorders" (phecode 627), "iron deficiency anemias, unspecified or not due to blood loss" (phecode 280.1), "sleep disorders" (phecode 327), and "insomnia" (phecode 327.4). A new metabolic phenotype of "disorders of fluid, electrolyte, and acid–base balance" (phecode 276) was phenome-wide significant in PheWAS-2 compared with PheWAS-1, but the association did not remain significant in replication analysis. The phenome-wide significant phenotype with the largest effect size in PheWAS-2 was "localized adiposity" (OR = 1.014, phecode 278.3), same as for PheWAS-1. The range of OR is low in PRS-PheWAS due to the cumulative effect sum of all PCOS susceptibility loci including low-effect variants.

Sensitivity analysis—Associations with individual PCOS susceptibility loci (PheWAS-3). In the individual PheWAS of 85 PCOS genome-wide significant variants, even though no association survives phenome-wide significance, likely due to the multiple testing burden, 11 PCOS variants show notable association to "polycystic ovaries" across the ancestry groups (most significant variant hg19 chr11:30 226 528, OR = 1.36, phecode 256.4), ranked as the second most significant phenotype (see (53)). Out of top 100 associations in PheWAS-3, the largest number of associations were related to the circulatory system for "thrombotic microangiopathy"

(31.0%). Endocrine/metabolic related phenotypes were the second most frequent category (21.0%) composed of either "PCOS" or "ovarian dysfunction", and 12% of the top associations were digestive traits, largely devoted to diverticular diseases. We did not identify any associations related to obesity or diabetes, which were the most significant phenotypic features found in PheWAS-1 and PheWAS-2.

Discussion

A key question in precision medicine is how to identify patients at high risk for a given disease for the goal of targeting preventive care. In this study, we examined the ability of PRS to predict PCOS clinical diagnosis and mine comorbid EHR phenotypes with the ultimate goal of improving diagnostic accuracy for PCOS. We show that a PRS for PCOS can be used (1) to identify patients at elevated risk of PCOS and (2) to determine the comorbid or pleiotropic phenome-wide expression associated with PCOS in a clinical setting.

The primary accomplishment of this study is a systematic enhancement of the polygenic risk prediction by integration of additional disease component phenotypes in the EHR into a PPRS. The onset of hirsutism, menstrual dysfunction, or female infertility are representative symptoms of PCOS and essential in determining clinical hyperandrogenism (15,56,57). They are not required for a diagnosis of PCOS per se, but are useful in suggesting PCOS in a clinical context. The PPRS significantly improves the average explanatory power (pseudo-R^2) of PCOS prediction by 0.221 (59.1-fold increase) compared with the null model without PRS or component phenotypes, and by 0.037 (14.7% increase) over the null model with the component phenotypes alone (Table 2 and Fig. 4). In contrast to the previous studies that attempted to identify PCOS diagnosis with risk score calculation (11,58), our algorithm did not limit risk predictor in a single dimension, using both phenotype and genotype markers with polygenic inheritance, and extensively demonstrated the predictive performance of PPRS with several machine-learning techniques. The findings shown here strengthen the potential clinical utility of PPRS as a disease predictor, particularly when combined with component symptom information available within the EHR.

To date, research has consistently shown that the PRS built from EA GWAS data does not perform as robustly across non-EA samples. In this study, we assessed the performance of a Eurocentrically built PCOS-PRS on the samples of EA, AA, and the joint MA cohorts. Undeniably, validation statistics varied by ancestry

group and the PCOS diagnosis prediction in AA cohort shows the poorest performance. However, it is of note that more than half of the tested models in AA still show statistical significance in terms of regression *P*-value, and those models display a reliable efficiency for PCOS detection in effect size and AUC (Table 3). Interestingly, the ORs for PRS differ across the ancestry cohorts, and somewhat higher in some prediction models in AA (average OR of model1 = 1.25, model2 = 1.28) and MA samples (average OR of model1 = 1.14, model2 = 1.13) than EA samples (average OR of model1 = 1.13, model2 = 1.12). The overall ORs of the PRS variable are fairly stable throughout all polygenic prediction models (OR 1.12–1.28). The observed significance of the PRS variable in the MA cohort, more stable than in the EA or AA participants alone, is likely due to the increased statistical power with larger sample size that counters the sample heterogeneity introduced. In addition, we found that the accumulation of genetic variants did not always increase the predictive capability of the PRS in terms of pseudo-R^2 and OR (Fig. 3 and Table 2). This might be due to the different RAF of PCOS risk variants by different PRS *P*-value cut-offs, and the varying LD structure of the ancestry groups. Previous research has confirmed that the LD pattern varies between EA and Chinese women at the PCOS susceptibility loci encoding LH/choriogonadotropin receptor (*LHCGR*) and follicle-stimulating hormone receptor (*FSHR*) genes, but the reproducible signals of the loci are consistently associated with PCOS regardless of ancestry(10,59). Our sensitivity analysis (PheWAS-3) also suggests the varying phenotypic effect of PCOS loci in different ancestries, but confirms the strong association with PCOS nonetheless. These findings demonstrate the primary role of PCOS-PRS in cumulatively explaining substantial variation of disease susceptibility across ancestries even with differing LD structures, and extend the general utility of PPRS in disease prediction.

Furthermore, our PRS-based phenome-wide analysis revealed several clinical associations that are tightly linked with obesity, confirming the shared metabolic pathways between PCOS and obesity in a phenomic aspect. As obesity is a common finding which can be found in 50% to 65% of PCOS patients (15), and previous Mendelian randomization study revealed the causal relationship of BMI on PCOS etiology (12), many of our findings could be interpreted as phenotypic evidence of comorbid obesity. "Morbid obesity" (phecode 278.11), "hypercholesterolemia" (phecode 272.11), "disorders of lipoid metabolism" (phecode 272), "hyperlipidemia" (phecode 272.1), "hypertension" (phecode 401), or "abnormal glucose" (phecode 250.4) are easily understandable with the context of heightened metabolic risks for obesity. "Sleep

apnea" (phecode 327.3) and "chronic liver disease and cirrhosis" (phecode 571), "GERD" (phecode 530.11), "diseases of esophagus" (phecode 530 and 530.1) are either neurological, pulmonary, or digestive assorted symptoms that are commonly found in the patients with obesity.

It is also noteworthy that there were 75 significant associations identified in women, while in men there were only 3 significantly associated diagnosis (morbid obesity, type 2 diabetes, diabetes mellitus) despite a similar sample size for males and females in the analysis. It is possible that the clinical consequences of high androgens in males are less likely to cause symptoms for which medical treatment is sought, or that these genetic variants only elevate androgen levels in a female "environment" but not a male one. The 3 identified phenotypes in males additionally suggest that if an individual harbors high genetic risk for PCOS, the metabolic manifestations are similar regardless of sex.

Consistent with previous studies (11,12), we identified phenotypic evidence of positive BMI association with genetic risk of PCOS. In the stratification analysis of the PRS, our observation of the increased BMI in individuals with high risk of PCOS are evident in both EA and MA cohorts (Fig. 2). The comorbid phenotypes could be driven by pleiotropy in which PCOS-associated genes also increase BMI, or could be due to underdiagnosis of PCOS itself, in which case the association with obesity phenotypes may be a result of comorbidity with undiagnosed PCOS.

Several limitations to this study need to be acknowledged. First, the sample size of AA participants was relatively small, which increases the likelihood of both false-negative and false-positive findings. Further investigation is needed to fully understand the overlap in PCOS genetic factors across MA participants and the methodological application of Eurocentric PCOS-PRS to other genetic ancestries considering LD structure. Secondly, the phenotypic components we used for polygenic prediction are currently limited to only 3 representative phenotypes: hirsutism, irregular menstruation, and female infertility. Fueled by our PheWAS finding, the work could be extended by incorporating the additional phenotypes that might increase the likelihood of an eventual diagnosis. Also, the phecode of PCOS used for PheWAS was converted from ICD-9-CM 256.4 and ICD-10-CM E28.2, which was used as a proxy for capturing PCOS in the EMR. This phecode may not perfectly capture PCOS as they may or may not capture hyperandrogenemia. The selection bias in our discovery cohort should be acknowledged as well. Two of our participating sites (Geisinger and Marshfield) mainly recruited their patients for the study of obesity and type 2 diabetes, which resulted in a higher proportion of obese

patients into their biobank and therefore may inflate the prevalence of PCOS in these subgroups. Additionally, we also observed 95% confidence intervals for the PheWAS sometimes passed through 1 in our replication PheWAS. This largely can be attributed to the winner's curse phenomenon (54), particularly when there are large differences in sample size between the discovery eMERGE cohort and the replication BioVU cohort. More importantly, the estimated effect sizes between the discovery and replication cohorts are very similar, suggesting validity of the results with differences in statistical significance and 95% CI due to sample size-driven power. Lastly, due to the low diagnosis rate of PCOS patients in current EHR system, it is possible that unidentified PCOS cases could reduce power in each analysis.

Our approach has provided a novel methodological opportunity to stratify patients' genetic risk and to discover the phenomic network associated with PCOS pathogenesis. Integrative analysis of the PRS-PheWAS enables the systematic interrogation of PCOS comorbidity patterns across the phenome, which cannot be readily identified by a single-variant approach. The identified phenomic networks could be used at the stage of first screening, prior to the testing of hormones or imaging of ovaries, or to help the patient and physician decide whether more extensive testing would be useful for PCOS diagnosis. As genomics-based precision medicine becomes more widely adopted as part of routine care, this approach should improve cost-effectiveness for PCOS screening/diagnosis by saving unnecessary screening tests or physician involvement in identifying possible PCOS cases. Further, it also permits detection of PCOS patients prior to diagnosis by a physician, allowing earlier interventions thereby reducing costs from long-term complications. Finally, from a precision medicine perspective, such an approach may provide a greater understanding of a patient's clinical presentation and suspected diagnosis based on specific phenotypic or genetic variations.

Acknowledgments

Additional Information

References

1. Davies N. PCOS: polycystic ovarian syndrome. *Diabetes Self Manag.* 2016;**33**(1):44–47.
2. 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):4650–4658.

3. The Endocrine Society. Endocrine Facts and Figures: Reproduction and Development. 2017. http://endocrinefacts.org/wp-content/uploads/Endocrine_Facts_Figures_Reproduction_and_Development.pdf.

4. Yawn V. Polycystic ovarian syndrome. *Adv NPs PAs*. 2012;**3**(12):11–4; quiz 15.

5. Vink JM, Sadrzadeh S, Lambalk CB, Boomsma DI. Heritability of polycystic ovary syndrome in a Dutch twin-family study. *J Clin Endocrinol Metab*. 2006;**91**(6):2100–2104.

6. Jahanfar S, Eden JA, Nguyen T, Wang XL, Wilcken DE. A twin study of polycystic ovary syndrome and lipids. *Gynecol Endocrinol*. 1997;**11**(2):111–117.

7. Jahanfar S, Eden JA, Warren P, Seppälä M, Nguyen TV. A twin study of polycystic ovary syndrome. *Fertil Steril*. 1995;**63**(3):478–486.

8. Hayes MG, Urbanek M, Ehrmann DA, et al.; Reproductive Medicine Network. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations. *Nat Commun*. 2015;**6**:7502.

9. Shi Y, Zhao H, Shi Y, et al. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet*. 2012;**44**(9):1020–1025.

10. Chen ZJ, Zhao H, He L, et al. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet*. 2011;**43**(1):55–59.

11. Day F, Karaderi T, Jones MR, et al.; 23andMe Research Team. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *Plos Genet*. 2018;**14**(12):e1007813.

12. Day FR, Hinds DA, Tung JY, et al. Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat Commun*. 2015;**6**:8464.

13. Broekmans FJ, Knauff EA, Valkenburg O, Laven JS, Eijkemans MJ, Fauser BC. PCOS according to the Rotterdam consensus criteria: change in prevalence among WHO-II anovulation and association with metabolic factors. *BJOG*. 2006;**113**(10):1210–1217.

14. Li L, Baek KH. Molecular genetics of polycystic ovary syndrome: an update. *Curr Mol Med*. 2015;**15**(4):331–342.

15. Futterweit W. Polycystic ovary syndrome: clinical perspectives and management. *Obstet Gynecol Surv*. 1999;**54**(6):403–413.

16. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. *Int J Env Res Pub He*. 2018;**15**(11): 2589-2601.

17. Carmina E. Diagnosis of polycystic ovary syndrome: from NIH criteria to ESHRE-ASRM guidelines. *Minerva Ginecol*. 2004;**56**(1):1–6.

18. Dewailly D. Diagnostic criteria for PCOS: is there a need for a rethink? *Best Pract Res Clin Obstet Gynaecol*. 2016;**37**:5–11.

19. Agapova SE, Cameo T, Sopher AB, Oberfield SE. Diagnosis and challenges of polycystic ovary syndrome in adolescence. *Semin Reprod Med*. 2014;**32**(3):194–201.

20. Fritsche LG, Gruber SB, Wu Z, et al. Association of polygenic risk scores for multiple cancers in a phenome-wide study: results from the Michigan genomics initiative. *Am J Hum Genet*. 2018;**102**(6):1048–1061.

21. Purcell SM, Wray NR, Stone JL, et al.; International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;**460**(7256):748–752.

22. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;**50**(9):1219–1224.

23. Zheutlin AB, Dennis J, Karlsson Linnér R, et al. Penetrance and pleiotropy of polygenic risk scores for schizophrenia in 106,160 patients across four health care systems. *Am J Psychiatry*. 2019;**176**(10):846–55.

24. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics*. 2014;**30**(16):2375–2376.

25. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*. 2010;**26**(9):1205–1210.

26. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;**51**(4):584–591.

27. Rosenberg NA, Edge MD, Pritchard JK, Feldman MW. Interpreting polygenic scores, polygenic adaptation, and human phenotypic differences. *Evol Med Public Health*. 2019;**2019**(1):26–34.

28. Duncan L, Shen H, Gelaye B, et al. Analysis of polygenic score usage and performance across diverse human populations. *bioRxiv*. 2018:398396.

29. Martin AR, Kanai M, Kamatani Y, et al. Hidden 'risk' in polygenic scores: clinical use today could exacerbate health disparities. *bioRxiv*. 2018:441261.

30. Martin AR, Gignoux CR, Walters RK, et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet*. 2017;**100**(4):635–649.

31. Curtis D. Polygenic risk score for schizophrenia is more strongly associated with ancestry than with schizophrenia. *Psychiatr Genet*. 2018;**28**(5):85–89.

32. McCarty CA, Chisholm RL, Chute CG, et al.; eMERGE Team. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics*. 2011;**4**:13.

33. Stanaway IB, Hall TO, Rosenthal EA, et al.; eMERGE Network. The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with the herpes zoster medical record phenotype. *Genet Epidemiol*. 2019;**43**(1):63–81.

34. Kho AN, Pacheco JA, Peissig PL, et al. Electronic medical records for genetic research: results of the eMERGE consortium. *Sci Transl Med*. 2011;**3**(79):79re1.

35. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable1.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-v3yd-n323 https://digitalhub.northwestern.edu/files/e8d0a9e6-bfbc-4c83-8223-38b6d8596a5e.

36. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;**81**(3):559–575.

37. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics*. 2015;**31**(9):1466–1468.

38. Denny JC, Bastarache L, Ritchie MD, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol*. 2013;**31**(12):1102–1110.

39. Wu P, Gifford A, Meng X, et al. Mapping ICD-10 and ICD-10-CM codes to phecodes: workflow development and initial evaluation. *JMIR Med Inform*. 2019;**7**(4):e14325.

40. Denny JC, Bastarache L, Roden DM. Phenome-wide association studies as a tool to advance precision medicine. *Annu Rev Genomics Hum Genet*. 2016;**17**:353–373.

41. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;**48**(10):1284–1287.

42. McCarthy S, Das S, Kretzschmar W, et al.; Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;**48**(10):1279–1283.

43. McFadden D. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*. 1973:105–142. https://eml.berkeley.edu/reprints/mcfadden/zarembka.pdf

44. Plomin R, Haworth CM, Davis OS. Common disorders are quantitative traits. *Nat Rev Genet*. 2009;**10**(12):872–878.

45. Krapohl E, Euesden J, Zabaneh D, et al. Phenome-wide analysis of genome-wide polygenic scores. *Mol Psychiatry*. 2016;**21**(9):1188–1193.

46. Joo YY, Actkins K, Pachec JA, et al. Data from: SupplementaryFigure1.tiff. DigitalHub. Galter Health Sciences Library & Learning

Center. Deposited October 3, 2019. 10.18131/g3-xv60-ms55, https://digitalhub.northwestern.edu/files/c565c8d2-4f0b-4fb7-ad0b-79b06b8c4482.

47. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable2.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-r7m1-wq23, https://digitalhub.northwestern.edu/files/fdce3659-6dd4-4b01-8f10-4a6a0b12a7fe.

48. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable3.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-tdm6-hh94 https://digitalhub.northwestern.edu/files/959bcdb0-560f-410d-add5-401a8064027a.

49. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable4.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-81aj-nt23 https://digitalhub.northwestern.edu/files/50a90ed9-c941-4404-9929-9e7268c2c0f8.

50. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;**44**(3):837–845.

51. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable5.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-4s8p-sc21 https://digitalhub.northwestern.edu/files/ac6e2904-2241-490e-a9ee-cfd75c9c20f9.

52. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable6.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-zyxw-wj46 https://digitalhub.northwestern.edu/files/6f730204-0d8d-4558-818d-e71379ef2e65.

53. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable7.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-jrm9-q777 https://digitalhub.northwestern.edu/files/b6109a74-725e-4b46-a38d-efadbaca444f.

54. Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. *Am J Hum Genet*. 2007;**80**(4):605–615.

55. Joo YY, Actkins K, Pachec JA, et al. Data from: PCOSPheWAS_E3ver2_Manuscript_SuppTable8.xlsx. DigitalHub. Galter Health Sciences Library & Learning Center. Deposited October 3, 2019. 10.18131/g3-t800-xz46 https://digitalhub.northwestern.edu/files/b90321d5-4989-4088-a11f-26b397f65384.

56. Goodman NF, Cobin RH, Futterweit W, Glueck JS, Legro RS, Carmina E; American Association of Clinical Endocrinologists (AACE); American College of Endocrinology (ACE); Androgen Excess and PCOS Society (AES). American Association of Clinical Endocrinologists, American College of Endocrinology, and Androgen Excess and PCOS Society Disease State Clinical Review: guide to the best practices in the evaluation and treatment of polycystic ovary syndrome–part 1. *Endocr Pract*. 2015;**21**(11):1291–1300.

57. Rosenfield RL, Lucky AW. Acne, hirsutism, and alopecia in adolescent girls. Clinical expressions of androgen excess. *Endocrinol Metab Clin North Am*. 1993;**22**(3):507–532.

58. Deshmukh H, Papageorgiou M, Kilpatrick ES, Atkin SL, Sathyapalan T. Development of a novel risk prediction and risk stratification score for polycystic ovary syndrome. *Clin Endocrinol (Oxf)*. 2019;**90**(1):162–169.

59. Mutharasan P, Galdones E, Peñalver Bernabé B, et al. Evidence for chromosome 2p16.3 polycystic ovary syndrome susceptibility locus in affected women of European ancestry. *J Clin Endocrinol Metab*. 2013;**98**(1):E185–E190.