



45 University Advanced Computing Center for Research and Education. This resource is funded by  
46 institutional, private, and federal grants, which include NIH funded Shared Instrumentation Grant  
47 S10OD017985 and S10RR025141 and CTSA grants UL1TR002243, UL1TR000445, and  
48 UL1RR024975.

49

50 **Author contributions:** LKD and KVA designed the study. KVA performed the statistical  
51 analyses. All authors contributed to the interpretation of the results, the revision of the  
52 manuscript, and approved the final manuscript.

53

54 **Competing interests:** The authors declare they have no competing interests.

55

56 **Data availability statement:** All GWAS summary statistics used in this study are publicly  
57 available and can be obtained from the following consortium groups and websites: International  
58 PCOS Consortium (<https://doi.org/10.17863/CAM.27720>), UK Biobank  
59 (<http://www.nealelab.is/uk-biobank>), HERMES (<http://www.broadcvdi.org/>), and  
60 CARDIoGRAMplusC4D (<http://www.cardiogramplusc4d.org/>). The summary statistics used from  
61 the Million Veteran Project can be accessed through dbGaP under accession number  
62 phs001672.v6.p1. The electronic health record data that support the findings of this study are  
63 available from Vanderbilt University Medical Center, but restrictions apply to the availability of  
64 these data, which were used under license for the current study, and so are not publicly  
65 available. Data are, however, available from the authors upon reasonable request and with  
66 permission of Vanderbilt University Medical Center.

67 **Abstract**

68 Females with polycystic ovary syndrome (PCOS), the most common endocrine disorder in  
69 women, have an increased risk of developing metabolic disorders such as insulin resistance,  
70 obesity, and type 2 diabetes (T2D). Furthermore, while only diagnosable in females, males with  
71 a family history of PCOS can also exhibit a poor cardiometabolic profile. Therefore, we aimed to  
72 elucidate the role of sex in the relationship between PCOS and its comorbidities by conducting  
73 bidirectional genetic risk score analyses in both sexes. We conducted a phenome-wide  
74 association study (PheWAS) using PCOS polygenic risk scores (PCOS<sub>PRS</sub>) to understand the  
75 pleiotropic effects of PCOS genetic liability across 1,380 medical conditions in females and  
76 males recorded in the Vanderbilt University Medical Center electronic health record (EHR)  
77 database. After adjusting for age and genetic ancestry, we found that European descent males  
78 with higher PCOS<sub>PRS</sub> were significantly more likely to develop cardiovascular diseases than  
79 females at the same level of genetic risk, while females had a higher odds of developing T2D.  
80 Based on observed significant associations, we tested the relationship between PRS for  
81 comorbid conditions (e.g., T2D, body mass index, hypertension, etc.) and found that only PRS  
82 generated for BMI and T2D were associated with a PCOS diagnosis. We then further  
83 decomposed the T2D<sub>PRS</sub> association with PCOS by adjusting the model for measured BMI and  
84 BMI<sub>residual</sub> (enriched for the environmental contribution to BMI). Results demonstrated that  
85 genetically regulated BMI primarily accounted for the relationship between T2D<sub>PRS</sub> and PCOS.  
86 This was further supported in a mediation analysis, which only revealed clinical BMI  
87 measurements, but not BMI<sub>residual</sub>, as a strong mediator for both sexes. Overall, our findings  
88 show that the genetic architecture of PCOS has distinct metabolic sex differences, but these  
89 associations are only apparent when PCOS<sub>PRS</sub> is explicitly modeled. It is possible that these  
90 pathways are less explained by the direct genetic risk of metabolic traits than they are by the  
91 risk factors shared between them, which can be influenced by biological variables such as sex.  
92

## 93 Introduction

94 Polycystic ovary syndrome (PCOS) is a highly heritable endocrine disorder that affects  
95 5%-21% of females of reproductive age who are typically diagnosed by having two or more of  
96 the following features under the Rotterdam criteria: polycystic ovaries, oligo- and anovulation, or  
97 hyperandrogenism [1–3]. Although Rotterdam is the most common PCOS criteria, other criteria  
98 can be used for diagnosis, including the National Institutes of Health criteria, the Androgen  
99 Excess and PCOS Criteria, or the 2018 International Evidence Based PCOS guidelines [2].  
100 Each criterion slightly differs in requirements in an effort to cover the range of PCOS symptoms  
101 that are exhibited in patients. However, as a result of the differing diagnostic criteria, the  
102 heterogenous presentation of symptoms, and the prevalent comorbidities that reside outside of  
103 diagnostic requirements patients often spend years seeking a diagnosis or worse, may be one  
104 of the 75% of females estimated to be undiagnosed [4,5].

105 Many clinicians select criteria based on their perception of the most defining PCOS  
106 feature [6]. In some cases, as with the Androgen Excess and PCOS Society criteria [7], this will  
107 mean hyperandrogenism, a symptom that typically manifests as acne, hirsutism, or alopecia [8].  
108 Androgen excess is also hypothesized to underlie many of the comorbid metabolic dysfunctions  
109 experienced by patients such as insulin resistance, obesity, metabolic syndrome, type 2  
110 diabetes, and cardiovascular diseases (CVD). However, as previous studies have shown, the  
111 genetic risk factors present in patients with metabolic manifestations of PCOS could differ from  
112 others who have primary presentations of reproductive dysfunction [9,10].

113 PCOS is multifactorial and twin studies estimate heritability at 70% [11–14]. With an  
114 underlying polygenic architecture, multiple variants are hypothesized to be involved in the  
115 development of PCOS [15,16]. Furthermore, the variation of clinical features can be partially  
116 explained by ancestry informative markers, indicating potential population specific effects  
117 [17,18]. Rare variants in genes, such as *DENND1A*, have also been identified in family studies  
118 for PCOS alongside many other genetic variants identified from GWAS [19]. Despite the small

119 effect size of individual common variants, aggregation of common risk variants together as a  
120 polygenic (or genetic) risk score (PRS) reflects the overall additive genetic liability to PCOS in  
121 individuals. This marker of disease risk is associated with PCOS diagnosis in multiple ancestries  
122 and offers many advantages to parsing out the genetic etiology of PCOS that is entangled with  
123 its comorbid presentations [11–13]. Furthermore, there is increasing evidence that a spectrum  
124 of clinical PCOS manifestations for PCOS is also correlated with the PRS for PCOS [11].

125 Therefore, in this study, we aimed to determine if the PRS for PCOS demonstrated  
126 pleiotropic associations with other health conditions in a hospital biobank population through a  
127 phenome-wide association study (PheWAS). By using a genetic model, we were able to  
128 determine sex-differentiated effects associated with PCOS genetic risk, revealing the impact of  
129 PCOS-associated inherited genetic variation in males, despite the fact that PCOS is only  
130 diagnosed in females. We further implemented several analyses to determine the mediating role  
131 of body mass index (BMI) on cardiometabolic comorbidities, and further identified the  
132 directionality of their effects through causal mediation analyses.

## 133 **Methods**

### 134 *VUMC EHR-linked Biorepository*

135  
136  
137  
138 Vanderbilt University Medical Center (VUMC) is a tertiary care hospital in Nashville,  
139 Tennessee, with several outpatient clinics throughout Tennessee and the surrounding states  
140 offering primary and secondary care. Medical records have been electronically documented at  
141 VUMC since the early 1990s, resulting in a clinical research database of over 3 million EHRs  
142 referred to as the Synthetic Derivative [20]. EHRs include demographic information, health  
143 information documented through International Classification of Disease, Ninth Revision (ICD9)  
144 and Tenth Revision (ICD10) codes, procedural codes (CPT), clinical notes, medications, and  
145 laboratory values. This information is linked with a DNA biorepository known as BioVU. Use of

146 EHRs and genetic data for this study was approved by the Vanderbilt University Institutional  
147 Review Board (IRB #160279).

148  
149 *Genetic data*

150  
151 BioVU contains 94,474 individuals genotyped on the MEGA<sup>EX</sup> platform [21]. The QC  
152 pipeline removed SNPs with low genotyping call rate ( $< 0.98$ ) and individual subjects who were  
153 related ( $\pi\text{-hat} > 0.2$ ), had low call rates ( $< 0.98$ ), sex discrepancies, and excessive  
154 heterozygosity ( $F_{\text{het}} > 0.2$ ). A principal component (PC) analysis (PCA) was performed on  
155 remaining individuals to determine genetic ancestry using FlashPCA2 [22]. BioVU genotyped  
156 samples were stratified by ancestral origin based on PCs herein referred to as the European  
157 (EUR) descent or African (AFR) descent dataset. Extended details for the quality control of  
158 these datasets have been described previously [23].

159  
160 *Publicly available summary statistics*

161 Genome-wide association study (GWAS) summary statistics were acquired for PCOS,  
162 BMI, diastolic blood pressure, systolic blood pressure, pulse, type 2 diabetes, heart failure, and  
163 coronary artery disease (i.e., all of the significant phenotypic associations observed in the  
164 PheWAS models) to further establish the strength of associations between PCOS and its  
165 comorbidities through subsequent analyses. Each GWAS was selected based on public  
166 availability, sample size, and sample diversity.

167 The Genetic Investigation of ANthropometric Traits (GIANT) consortium body mass  
168 index (BMI) summary statistics included 339,224 individuals of European and non-European  
169 descent from over 125 studies. These BMI summary statistics were used as a proxy for obesity  
170 summary statistics [24]. Blood pressure traits (diastolic, systolic, and pulse) and type 2 diabetes  
171 (T2D) summary statistics were obtained from the Million Veterans Program (MVP), a large  
172 biobank consortium effort that houses biobank data from various sites in the Department of

173 Veterans Affairs health system [25]. Blood pressure traits were generated from a trans-ethnic  
174 sample of over 750,000 individuals from MVP [26]. T2D summary statistics were generated from  
175 a meta-analysis using data from 1.4 million participants in various biobanks and consortia  
176 groups [27]. Heart failure summary statistics were collected from 47,309 cases and 930,014  
177 controls of European ancestry across nine studies in the Heart Failure Molecular Epidemiology  
178 for Therapeutic Targets (HERMES) consortium as a proxy for heart disease [28]. Finally,  
179 coronary artery disease (CAD) datasets generated from the Coronary Artery Disease Genome-  
180 wide Replication and Meta-analysis plus The Coronary Artery Disease (CARDIoGRAMplusC4D)  
181 consortium were used as the genetic measurement for the coronary atherosclerosis phenotype  
182 [29]. This meta-analysis assembled 60,801 cases and 123,504 controls of multiple ancestries  
183 across forty-eight study sites.

184  
185 *Statistical analysis*

186  
187 *Generation of polygenic risk scores (PRS)*

188  
189 PCOS<sub>PRS</sub> were calculated with PRS-CS software using the weighted sums of the risk  
190 allele effects as reported in the summary statistics from the Day et al. GWAS of PCOS and  
191 applying a Bayesian continuous shrinkage parameter select SNP features and to model linkage  
192 disequilibrium [16,30]. The details of these methods have been previously described elsewhere  
193 [11]. We calculated PCOS<sub>PRS</sub> for both EUR and AFR BioVU genotyped ancestry samples which  
194 were previously shown to be associated with a PCOS diagnosis defined by a coded strict PCOS  
195 definition in our previously published EHR-based algorithm [11].

196  
197 *Phenome-wide association study (PheWAS)*

198 Next, we were interested in identifying the pleiotropic effects of the genetic susceptibility  
199 to PCOS on the medical phenome. Therefore, in a PheWAS framework we analyzed the effects  
200 of PCOS<sub>PRS</sub> across 1,380 medical conditions. This analysis was first performed in a female

201 sample for our EUR and AFR ancestry datasets to validate the PRS and identify potential  
202 pleiotropic relationships (**Supplementary Figures 1-5**). Although males cannot be diagnosed  
203 with PCOS, they still harbor genetic risk for PCOS. Thus, we extended the analysis to males to  
204 examine sex differences. Finally, we performed a sex-combined PheWAS to increase statistical  
205 power. In the sex-combined logistic regression model, covariates included median age of  
206 individuals medical record, sex, and the first ten principal components. In the sex-stratified  
207 models, covariates included median age and the first ten PCs.

208

#### 209 *Interaction analysis*

210 For each phenotype with evidence of a significant main effect of PCOS<sub>PRS</sub> in either sex,  
211 we tested for two-way interactions (sex \* PCOS<sub>PRS</sub>) to determine which of the significant sex  
212 stratified PheWAS associations were influenced by biological effects of sex. Selected  
213 phenotypes of interest that met the phenome-wide false discovery rate ( $q < 0.05$ ) were tested  
214 for interactions ( $N = 10$ ). For these interaction analyses, sex-combined models included the  
215 main effects of PCOS<sub>PRS</sub>, median age of individuals medical record, sex, and the top ten PCs.

216

#### 217 *Sensitivity analyses*

218 Several sensitivity analyses were performed to assess the robustness of the significant  
219 phenome-wide findings. First, BMI is strongly correlated with both PCOS and its comorbidities,  
220 and thus, can influence the strength of the results [31]. Therefore, we adjusted each model for  
221 BMI (median measurement across an individual's EHR) to observe whether the significant  
222 phenotypes associated with PCOS<sub>PRS</sub> were independent of obesity-related effects. In addition to  
223 adjusting for BMI, the model was adjusted for median age, sex (only in the sex-combined  
224 sample), and the top ten PCs.

225 Next, we evaluated which phenotypes were dependent on a PCOS diagnosis in females  
226 by accounting for PCOS diagnosis [11]. This analysis allowed us to identify true pleiotropic



227 associations and provided insight into which phenotypes were exclusively correlated with  
228 genetic risk, even in the absence of PCOS.

229 Finally, given that BMI has strong contributions from both genetic and environmental  
230 sources of variance, we revisited the previous PheWAS analyses to specifically account for the  
231 environmental contribution of BMI. To do this, we calculated the residuals of median BMI  
232 adjusted for BMI<sub>PRS</sub> (residuals(medianBMI ~ BMI<sub>PRS</sub>)), herein referred to as BMI<sub>residual</sub>. BMI<sub>residual</sub>  
233 were then used in a subsequent sensitivity analysis of the previously described PheWAS.

234

### 235 *Genetic correlation*

236 Linkage Disequilibrium Score Regression (LDSC) was used to calculate genetic  
237 correlation, an estimate of genetic similarity, between traits [32]. LDSC only utilizes GWAS  
238 summary statistics and is not sensitive to sample overlap, which may be present across the  
239 publicly available GWAS datasets used in this study. This method utilizes the effect estimate of  
240 each SNP and accounts for the effects of SNPs in linkage disequilibrium based on the GWAS  
241 reference population. The European reference panel was used for all analyses based on the  
242 demographic majority of each of the GWAS samples.

243

### 244 *Logistic regression models*

245 Public GWAS datasets were used to generate PRS for phenome-wide significant  
246 phenotypes identified in the PheWAS. PRS were then used as the independent variable in a  
247 logistic regression model against PCOS diagnosis as the dependent variable. BioVU genotyped  
248 datasets were filtered to females and contained 361 PCOS cases and 29,035 controls in the  
249 EUR dataset and 189 PCOS cases and 9,229 controls in the AFR dataset. Selection of PCOS  
250 cases and controls have been described elsewhere [11]. In brief, cases required PCOS billing  
251 codes and no exclusion codes. Controls excluded individuals who had any inclusion or  
252 exclusion codes [11]. All models were adjusted for median age of the individuals medical record,

253 and the top ten PCs for each ancestry. Median BMI was included as a covariate in the  
254 sensitivity analysis.

255 Lastly, we adjusted for BMI<sub>residual</sub> in the logistic regression models in place of BMI. Next,  
256 we performed a sensitivity analysis to test the effect of BMI<sub>residual</sub> and BMI<sub>PRS</sub> separately on  
257 PCOS diagnosis. A multiple testing correction of  $p < 7.35e-04$  (0.05/68) was implemented to  
258 account for all statistical tests. All statistical analyses were done using R 3.6.0.

259

### 260 *Mediation analysis*

261 To test whether BMI or BMI<sub>residual</sub> mediates the pleiotropic relationships between PCOS  
262 and associated cardiometabolic conditions, a mediation analysis was employed using the  
263 *mediation* R package for both AFR and EUR genotyped samples [33]. For the first set of  
264 mediation analyses, we modeled BMI<sub>PRS</sub> and BMI<sub>residual</sub> as exposures, T2D diagnosis as the  
265 mediator, and PCOS case status as the outcome. Another reciprocal model was tested with the  
266 same exposures, but with PCOS diagnosis as the mediator and T2D as the outcome. This  
267 analysis tested the hypothesis that the risk conferred by BMI on one phenotype (e.g., PCOS) in  
268 turn increases risk for another other phenotype (e.g., T2D). We restricted this analysis to the  
269 female population. All models were further adjusted for median age and the top ten PCs.

270 In our second set of mediation analyses, we tested whether genetic risk for PCOS or  
271 T2D also influenced BMI<sub>residual</sub> through the manifestation of either clinical condition. For example,  
272 we modeled PCOS<sub>PRS</sub> as the exposure variable, T2D diagnosis as the mediator and BMI<sub>residual</sub>  
273 as the outcome. Finally, in a female only analysis, T2D<sub>PRS</sub> was modeled the exposure variable,  
274 PCOS diagnosis was included as the mediator, and BMI<sub>residual</sub> again was modeled as the  
275 outcome.

276 Finally, for our last set of mediation analyses, we examined the mediating effect of both  
277 measured BMI and BMI<sub>residual</sub> on the relationship between PCOS<sub>PRS</sub> (exposure) and the  
278 diagnosis of hypertension and hypertensive heart disease (i.e., outcomes that demonstrated

279 evidence of significant interactions with sex). T2D was also tested because it was significantly  
280 associated with PCOS<sub>PRS</sub> in both sexes. PCOS<sub>PRS</sub> was modeled as the exposure variable and  
281 T2D (phecode = 270.2) hypertension (phecode = 411), and hypertensive heart disease  
282 (phecode = 401.21) were modeled as the outcome variables. Models were considered  
283 significant when the average direct effect (ADE) and average causal mediation effect (ACME)  
284 both passed  $p < 7.35e-04$ . All mediation analyses were performed with 1,000 bootstrap  
285 simulations to estimate the confidence intervals and determine an empirical p-value.

286

## 287 **Results**

288

### 289 *PCOS<sub>PRS</sub> PheWAS results*

290 The PCOS<sub>PRS</sub> PheWAS was applied to 66,903 EUR individuals from BioVU. When  
291 restricted to females (N = 37,240), two phenotypes were significantly associated with PCOS<sub>PRS</sub>  
292 (Bonferroni corrected p-value = 4.56e-05) (**Figure 1a**): T2D (odds ratio [OR] = 1.10, 95%  
293 confidence interval [CI] = 1.06-1.14,  $p = 5.54e-07$ ) and diabetes mellitus (OR = 1.09, 95% CI =  
294 1.05-1.13,  $p = 2.20e-06$ ). When the analysis was restricted to males (N = 29,663; **Figure 1b**),  
295 hypertensive heart disease was significant (OR = 1.15, 95% CI = 1.08-1.23,  $p = 2.07e-05$ )  
296 alongside a cluster of nominally ( $p < 0.05$ ) associated cardiovascular phenotypes with similar  
297 effect sizes (OR = 1.06 to 1.08). This included hypertension (OR = 1.06, 95% CI = 1.03-1.10,  
298  $p = 1.30e-04$ ), essential hypertension (OR = 1.06, 95% CI = 1.03-1.10,  $p = 2.49e-04$ ), and  
299 coronary atherosclerosis (OR = 1.06, 95% CI = 1.02-1.10,  $p = 1.40e-03$ ). T2D (OR = 1.07, 95% CI  
300 = 1.03-1.11,  $p = 7.04e-04$ ) and several other endocrine phenotypes were also modestly  
301 associated with PCOS<sub>PRS</sub> in males.

302 In the EUR sex-combined model, four phenotypes were significantly associated with  
303 PCOS<sub>PRS</sub> (**Figure 2**). This included T2D and diabetes mellitus as the top associations with the  
304 same OR of 1.08 (T2D 95% CI = 1.06-1.11,  $p = 3.15e-09$ ; diabetes mellitus 95% CI = 1.05-1.11,

305  $p = 1.03e-08$ ). Following was obesity (OR = 1.07, 95% CI = 1.04-1.11,  $p = 9.68e-06$ ) and  
306 hypertensive heart disease (OR = 1.11, 95% CI = 1.06-1.16,  $p = 3.57e-05$ ). Three  
307 cardiovascular diseases were also nominally associated with PCOS<sub>PRS</sub>, presumably due to the  
308 addition of males. PCOS<sub>PRS</sub> yielded an OR of 1.21 (95% CI = 1.08-1.36,  $p = 7.82e-04$ ) for an  
309 association with polycystic ovaries, falling just short of phenome-wide significance.

310 No associations passed Bonferroni correction in the AFR ancestry analysis, which  
311 included a total of 12,383 individuals (**Supplementary Figure 1**).

312

### 313 *Interaction analysis*

314 The sex interaction analysis demonstrated that males with a high PCOS<sub>PRS</sub> were  
315 (interaction  $p = 7.24e-03$ ) more likely to be diagnosed with hypertension (OR<sub>Males</sub> = 1.06, 95% CI  
316 = 1.03-1.10,  $p = 1.30e-04$ ; OR<sub>Females</sub> = 1.03, 95% CI = 1.00-1.06,  $p = 0.07$ ), essential  
317 hypertension (interaction  $p = 7.71e-03$ , OR<sub>Males</sub> = 1.06, 95% CI = 1.03-1.10,  $p = 2.49e-04$ ;  
318 OR<sub>Females</sub> = 1.03, 95% CI = 1.00-1.06,  $p = 0.08$ ), and hypertensive heart disease (interaction  
319  $p = 0.01$ , OR<sub>Males</sub> = 1.15, 95% CI = 1.08-1.23,  $p = 2.07e-05$ ; OR<sub>Females</sub> = 1.05, 95% CI = 0.98-1.13,  
320  $p = 0.16$ ) than females with the same PCOS<sub>PRS</sub> (**Table 1**). These sex differences were also  
321 observed when calculating the prevalence for each trait by decile of PCOS<sub>PRS</sub>, which only  
322 showed an upwards trend across deciles for traits with a significant sex effect (**Figure 3**).  
323 Together, these results showed that sex is an important modifier of PCOS genetic risk.

324

### 325 *Sensitivity analyses adjusting for BMI and PCOS case status*

326 In our first set of PheWAS sensitivity analyses for the EUR population, we adjusted for  
327 median BMI in all of the PRS-PheWAS models. There were no surviving associations in the  
328 sex-combined model or stratified analyses, suggesting BMI may mediate the pleiotropic effects  
329 of PCOS<sub>PRS</sub> (**Supplementary Figures 2a and 3**). In a separate sensitivity analysis (female only)  
330 we found that after adjusting for PCOS diagnosis, females with a high PCOS<sub>PRS</sub> still

331 demonstrated a significant positive phenome-wide association with T2D and diabetes mellitus  
332 (**Supplementary Figure 2b**).

333

#### 334 *Genetic correlation results*

335 We found that BMI ( $r_g = 43.32\%$ ,  $p = 8.77e-19$ ) and T2D ( $r_g = 29.26\%$ ,  $p = 3.25e-09$ ) had  
336 the strongest genetic correlation with PCOS (**Table 2**). Heart failure ( $r_g = 26.51\%$ ,  $p = 0.0025$ )  
337 and pulse pressure ( $r_g = 13.46\%$ ,  $p = 0.011$ ) were also modestly significantly with PCOS. CAD  
338 bordered the significance threshold, but systolic and diastolic blood pressure were not  
339 significantly genetically correlated with PCOS.

340

#### 341 *Cardiometabolic PRS Analysis of PCOS diagnosis*

342 Outside of T2D and BMI, none of the PRS built for CAD, heart failure, or blood pressure  
343 were significantly associated with a PCOS diagnosis in females of either EUR or AFR descent  
344 (**Figure 4**).

345 BMI<sub>PRS</sub> was positively associated with PCOS diagnosis for both EUR and AFR ancestry  
346 populations ( $p_{EUR} = 3.71e-08$ ,  $p_{AFR} = 0.02$ ), but not independently of clinically measured BMI  
347 values. T2D<sub>PRS</sub> in the EUR dataset yielded similar results with PCOS diagnosis, again losing  
348 significance upon addition of a BMI covariate in the model ( $OR_{unadjusted} = 1.16$ , 95% CI = 1.04-  
349 1.30,  $p = 0.007$ ;  $OR_{adjusted} = 1.09$ , 95% CI = 0.97-1.22,  $p = 0.16$ ). To determine whether this  
350 reduction in effect size was due to the genetic correlation between T2D and BMI, we also tested  
351 a model in which T2D<sub>PRS</sub> was adjusted for BMI<sub>residual</sub> (i.e., variance remaining in BMI after  
352 removing variance due to BMI<sub>PRS</sub>) instead of BMI. This model indeed recovered the original  
353 association between T2D<sub>PRS</sub> ( $OR_{EUR} = 1.18$ , 95% CI = 1.05-1.32,  $p_{EUR} = 0.005$ ) and PCOS  
354 diagnosis for the EUR population, suggesting that *genetically predicted* BMI mediates the  
355 association between T2D<sub>PRS</sub> and PCOS diagnosis (**Figure 5**).

356

357 *Mediation analysis in females with BMI as an exposure variable*

358 To further quantify the degree to which BMI influenced the shared genetic pathways  
359 between PCOS and T2D, a mediation analysis was performed with BMI<sub>PRS</sub> and BMI<sub>residual</sub> as the  
360 exposure variables, PCOS diagnosis as the outcome variable, and T2D as the mediator. The  
361 reciprocal model was then tested with T2D diagnosis as the outcome and PCOS as the  
362 mediator (**Table 3**). We found that both phenotypes acted as significant mediators when BMI<sub>PRS</sub>  
363 and BMI<sub>residual</sub> were exposure variables. PCOS diagnosis had a stronger mediating effect on T2D  
364 when BMI<sub>residual</sub> was the exposure variable compared to BMI<sub>PRS</sub> (13.6% vs 7.1%). However, the  
365 opposite was true for T2D, which mediated more of risk conferred by BMI<sub>PRS</sub> than BMI<sub>residual</sub> on  
366 the PCOS outcome (9% vs 2.1%).

367

368 *PheWAS sensitivity analysis with BMI<sub>residual</sub>*

369 We revisited the PheWAS analysis in light of the findings for BMI<sub>residual</sub>. By covarying for  
370 BMI<sub>residual</sub>, (e.g., environmental estimate of BMI), we observed almost no difference from the  
371 original results in our female sample (**Supplementary Figure 4a**). T2D remained significant in  
372 females (OR =1.10, 95% CI = 1.06-1.15, p = 1.69e-06), as did diabetes mellitus (OR =1.09,  
373 95% CI = 1.05-1.13, p = 9.91e-06). These associations also appeared in the combined dataset  
374 alongside obesity (**Supplementary Figure 5**). Although none of the phenotypes passed  
375 Bonferroni correction for males, many of the associations improved in estimate effects  
376 (**Supplementary Figure 4b**).

377

378 *Mediation analysis in the sex-combined and sex-stratified samples*

379 Upon finding BMI<sub>residual</sub> as a significant risk exposure for T2D and PCOS in the first  
380 mediation analysis, but not a strong confounder in the PCOS<sub>PRS</sub> PheWAS analysis, we tested  
381 whether the genetic risk conferred by one condition (e.g., T2D<sub>PRS</sub>) was mediated by the clinical  
382 manifestation of the other (e.g., PCOS diagnosis) when examining BMI<sub>residual</sub> as the outcome.

383 Here we observed a modest association that showed PCOS diagnosis mediated 33% of the risk  
384 conferred by T2D<sub>PRS</sub> on BMI<sub>residual</sub> whereas T2D diagnosis did not explain any of the variance in  
385 BMI<sub>residual</sub> due to PCOS<sub>PRS</sub> (**Table 4**).

386 Lastly, the effects of BMI on PCOS comorbidities T2D, hypertension, and hypertensive  
387 heart disease were explored (**Table 5**). We found that BMI<sub>residual</sub> was not a significant mediator  
388 for any of the tested models, even though clinical BMI was a significant mediator for T2D in  
389 females and a significant mediator for hypertensive heart disease in males. This again  
390 implicated *genetically regulated* BMI as the main source of the mediating effect for males since  
391 median BMI significantly mediated 14.8% of the variance in hypertensive heart disease ( $p =$   
392  $<2e-16$ ) and mediated 23.7% of the variance in hypertension ( $p = 0.002$ ) due to PCOS<sub>PRS</sub>. The  
393 lack of an association with BMI<sub>residual</sub> in females also suggests *genetically regulated* BMI as the  
394 primary mediator for T2D, where it mediated 31.5% of the variance explained by the PCOS<sub>PRS</sub>.

## 395 396 **Discussion**

397  
398 Through a comprehensive analysis of PCOS genetic risk across multiple phenotypes,  
399 we identified sex differences in the cardiometabolic traits associated with PCOS genetic risk.  
400 Among these, the most notable difference was that males with high PCOS<sub>PRS</sub> were at greater  
401 risk of cardiovascular conditions than females at the same level of genetic risk. Genetically  
402 predicted BMI was revealed as a primary mediator of this risk in males, whereas both  
403 environmental and genetic BMI variance components mediated risk of T2D conferred by  
404 PCOS<sub>PRS</sub> in females. Furthermore, only T2D<sub>PRS</sub> and BMI<sub>PRS</sub> were associated with PCOS  
405 diagnosis, indicating that many of the associations observed in the PCOS<sub>PRS</sub> PheWAS were  
406 primarily driven by PCOS genetic risk and not the genetic effects of the identified comorbidities.

407 There is growing interest in studying PCOS related effects in males whether through  
408 their relationship to first-degree family members with PCOS or by an equivalent phenotype [34–  
409 36]. Generally, males with mothers or sisters with PCOS tend to exhibit a poorer

410 cardiometabolic profile which can be observed as early as infancy [35]. A previous study  
411 suggests that males who exhibit a high genetic risk for PCOS are more likely to present with  
412 morbid obesity, T2D, and diabetes mellitus, two findings which were confirmed in this study [12].  
413 Although we also found associations with CVD phenotypes in males, these were largely  
414 influenced by genetically predicted BMI which can significantly increase the lifetime risk for CVD  
415 and mortality rates of high-risk individuals [37]. Genetic susceptibility for PCOS could be an  
416 additional catalyst for these events in males, making individuals already predisposed to adverse  
417 metabolic outcomes more vulnerable.

418 In an effort to determine whether genetic risk for phenotypes comorbid with PCOS could  
419 also increase risk for PCOS, we conducted separate multivariable logistic regressions with PRS  
420 for BMI, diastolic blood pressure, systolic blood pressure, pulse pressure, T2D, heart failure,  
421 and CAD on the PCOS diagnosis outcome and found no significant associations outside of  
422 T2D<sub>PRS</sub> and BMI<sub>PRS</sub>. This was unsurprising as the relationship between PCOS and CVD are still  
423 poorly elucidated and often debated [38]. It is possible that CVD is more prevalent in genetically  
424 high-risk males than females. The sex-difference in CVD prevalence may also be reflected in  
425 the PRS which may fail to fully capture CVD genetic risk for PCOS when stratified by sex due to  
426 fewer females in the discovery sample.

427 This study provides novel insight into PCOS genetic etiology and continues to  
428 underscore the importance of BMI in PCOS risk. As with T2D, we found that genetically  
429 predicted BMI is primarily responsible for the phenotypic association between T2D<sub>PRS</sub> and  
430 PCOS. However, PCOS<sub>PRS</sub> genetic risk led to more comorbid traits in the PheWAS in which  
431 both environmental and genetically predicted BMI influenced the metabolic associations. This  
432 may mean that PCOS patients with a family history of T2D can have an increased risk for  
433 morbid outcomes, as females with PCOS are already more likely to develop T2D at an earlier  
434 age [39]. This genetic susceptibility could also explain the high prevalence of insulin resistance  
435 in PCOS patients, which can be as high as 70% across all BMI strata [40]. These effects should



436 be investigated further, as the genetic and biological pathways could differ in lean PCOS  
437 patients who also experience a high rate of insulin resistance [41].

438         This study offers many strengths. Firstly, we showed that cardiometabolic associations  
439 vary with sex and that the metabolic outcomes related to PCOS genetic architecture can be  
440 further understood by studying both males and females. Secondly, we decomposed BMI into  
441 genetically predicted (i.e.,  $BMI_{PRS}$ ) and environmentally enriched (i.e.,  $BMI_{residual}$ ) and evaluated  
442 their respective roles in mediating the cardiometabolic profiles associated with  $PCOS_{PRS}$ .  
443 However, limitations include low power to detect any significant associations in our African  
444 descent sample. To date, there is no PCOS GWAS of African descent individuals, limiting all  
445 current similar studies to building PRS using European-based genetic variants, which do not  
446 perform as well in non-European populations [11,12]. Second, we only examined one  
447 environmental risk factor. Although many effects such as lifestyle and diet can be captured  
448 through BMI, it is not an exhaustive measurement nor does it accurately account for the full  
449 wellness of an individual [42,43]. Although other anthropometric features like hip-to-waist ratio  
450 (WHR) may be better indicators of health for some phenotypes [44], this information is not  
451 routinely collected in clinical settings or reported in EHRs. Furthermore, evidence does suggest  
452 that clinically ascertained BMI may be more informative for PCOS than WHR [45]. Finally,  
453 despite using the largest PCOS GWAS to date for this analysis, our  $PCOS_{PRS}$  still only explains  
454 a small portion of PCOS genetic variance. As these analyses expand, so too will our ability to  
455 detect the full genetic spectrum of PCOS and its subphenotypes.

456         PCOS is a multifaceted disorder with genetic architecture that is reflective of its  
457 heterogeneous outcomes. This polygenic structure captures a spectrum of metabolic  
458 comorbidities that is even more apparent when compared between sexes. Our findings show  
459 that males with high PCOS liability are indeed a high-risk group and those with a family history  
460 of PCOS should be closely monitored for hypertension and CVD. This is also true for females  
461 with PCOS and a family history of T2D, whose genetic risk could induce more severe comorbid

462 outcomes. As such, management and screening strategies should be updated to reflect  
463 advances in PCOS etiology. This call to action is paramount and requires both widespread  
464 dissemination of risk factor information to relevant stakeholders and increases in PCOS  
465 research priorities and funding. This becomes even more crucial as PCOS comorbidities are  
466 often under-recognized in clinical settings and metabolic conditions are underutilized in PCOS  
467 screening methods [46–49].

468 **References**

469

- 470 1. Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003  
471 consensus on diagnostic criteria and long-term health risks related to polycystic ovary  
472 syndrome (PCOS). *Hum Reprod.* 2004;19: 41–47. doi:10.1093/humrep/deh098
  
- 473 2. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al. Recommendations  
474 from the international evidence-based guideline for the assessment and management of  
475 polycystic ovary syndrome. *Hum Reprod.* 2018;33: 1602–1618.  
476 doi:10.1093/humrep/dey256
  
- 477 3. Shorakae S, Boyle J, Teede H. Polycystic ovary syndrome: a common hormonal condition  
478 with major metabolic sequelae that physicians should know about. *Intern Med J.* 2014;44:  
479 720–726. doi:10.1111/imj.12495
  
- 480 4. Gibson-Helm M, Teede H, Dunaif A, Dokras A. Delayed Diagnosis and a Lack of  
481 Information Associated With Dissatisfaction in Women With Polycystic Ovary Syndrome. *J*  
482 *Clin Endocrinol Metab.* 2017;102: 604–612. doi:10.1210/jc.2016-2963
  
- 483 5. March WA, Moore VM, Willson KJ, Phillips DIW, Norman RJ, Davies MJ. The prevalence  
484 of polycystic ovary syndrome in a community sample assessed under contrasting  
485 diagnostic criteria. *Human Reproduction.* 2010;25: 544–551. doi:10.1093/humrep/dep399
  
- 486 6. Cussons AJ, Stuckey BGA, Walsh JP, Burke V, Norman RJ. Polycystic ovarian syndrome:  
487 marked differences between endocrinologists and gynaecologists in diagnosis and  
488 management. *Clin Endocrinol.* 2005;62: 289–295. doi:10.1111/j.1365-2265.2004.02208.x
  
- 489 7. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit  
490 W, et al. Criteria for Defining Polycystic Ovary Syndrome as a Predominantly  
491 Hyperandrogenic Syndrome: An Androgen Excess Society Guideline. *The Journal of*  
492 *Clinical Endocrinology & Metabolism.* 2006;91: 4237–4245. doi:10.1210/jc.2006-0178
  
- 493 8. Sanchez-Garrido MA, Tena-Sempere M. Metabolic dysfunction in polycystic ovary  
494 syndrome: Pathogenic role of androgen excess and potential therapeutic strategies. *Mol*  
495 *Metab.* 2020;35: 100937. doi:10.1016/j.molmet.2020.01.001
  
- 496 9. Dapas M, Lin FTJ, Nadkarni GN, Sisk R, Legro RS, Urbanek M, et al. Distinct subtypes of  
497 polycystic ovary syndrome with novel genetic associations: An unsupervised, phenotypic  
498 clustering analysis. *PLoS Med.* 2020;17: e1003132. doi:10.1371/journal.pmed.1003132
  
- 499 10. Zhang Y, Movva VC, Williams MS, Lee MTM. Polycystic Ovary Syndrome Susceptibility  
500 Loci Inform Disease Etiological Heterogeneity. *JCM.* 2021;10: 2688.  
501 doi:10.3390/jcm10122688
  
- 502 11. Actkins KV, Singh K, Hucks D, Velez Edwards DR, Aldrich M, Cha J, et al. Characterizing  
503 the clinical and genetic spectrum of polycystic ovary syndrome in electronic health records.  
504 *J Clin Endocrinol Metab.* 2020. doi:10.1210/clinem/dgaa675
  
- 505 12. Joo YY, Actkins K, Pacheco JA, Basile AO, Carroll R, Crosslin DR, et al. A Polygenic and  
506 Phenotypic Risk Prediction for Polycystic Ovary Syndrome Evaluated by Phenome-Wide

- 507 Association Studies. *J Clin Endocrinol Metab.* 2020;105: 1918–1936.  
508 doi:10.1210/clinem/dgz326
- 509 13. Lee H, Oh J-Y, Sung Y-A, Chung HW. A genetic risk score is associated with polycystic  
510 ovary syndrome-related traits. *Hum Reprod.* 2016;31: 209–215.  
511 doi:10.1093/humrep/dev282
- 512 14. Vink JM, Sadrzadeh S, Lambalk CB, Boomsma DI. Heritability of polycystic ovary  
513 syndrome in a Dutch twin-family study. *J Clin Endocrinol Metab.* 2006;91: 2100–2104.  
514 doi:10.1210/jc.2005-1494
- 515 15. Day FR, Hinds DA, Tung JY, Stolk L, Styrkarsdottir U, Saxena R, et al. Causal  
516 mechanisms and balancing selection inferred from genetic associations with polycystic  
517 ovary syndrome. *Nat Commun.* 2015;6: 8464. doi:10.1038/ncomms9464
- 518 16. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, et al. Large-scale genome-wide  
519 meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for  
520 different diagnosis criteria. *PLoS Genet.* 2018;14: e1007813.  
521 doi:10.1371/journal.pgen.1007813
- 522 17. Louwers YV, Lao O, Fauser BCJM, Kayser M, Laven JSE. The impact of self-reported  
523 ethnicity versus genetic ancestry on phenotypic characteristics of polycystic ovary  
524 syndrome (PCOS). *J Clin Endocrinol Metab.* 2014;99: E2107-16. doi:10.1210/jc.2014-1084
- 525 18. Bjornes AC, Saxena R, Welt CK. Relationship between polycystic ovary syndrome and  
526 ancestry in European Americans. *Fertil Steril.* 2016;106: 1772–1777.  
527 doi:10.1016/j.fertnstert.2016.08.033
- 528 19. Dapas M, Sisk R, Legro RS, Urbanek M, Dunaif A, Hayes MG. Family-based quantitative  
529 trait meta-analysis implicates rare noncoding variants in DENND1A in polycystic ovary  
530 syndrome. *J Clin Endocrinol Metab.* 2019. doi:10.1210/jc.2018-02496
- 531 20. Robinson JR, Wei W-Q, Roden DM, Denny JC. Defining Phenotypes from Clinical Data to  
532 Drive Genomic Research. *Annu Rev Biomed Data Sci.* 2018;1: 69–92.  
533 doi:10.1146/annurev-biodatasci-080917-013335
- 534 21. Bien SA, Wojcik GL, Zubair N, Gignoux CR, Martin AR, Kocarnik JM, et al. Strategies for  
535 Enriching Variant Coverage in Candidate Disease Loci on a Multiethnic Genotyping Array.  
536 *PLoS One.* 2016;11: e0167758. doi:10.1371/journal.pone.0167758
- 537 22. Abraham G, Qiu Y, Inouye M. FlashPCA2: principal component analysis of Biobank-scale  
538 genotype datasets. *Bioinformatics.* 2017;33: 2776–2778. doi:10.1093/bioinformatics/btx299
- 539 23. Dennis JK, Sealock JM, Straub P, Lee YH, Hucks D, Actkins K, et al. Clinical laboratory  
540 test-wide association scan of polygenic scores identifies biomarkers of complex disease.  
541 *Genome Med.* 2021;13: 6. doi:10.1186/s13073-020-00820-8
- 542 24. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body  
543 mass index yield new insights for obesity biology. *Nature.* 2015;518: 197–206.  
544 doi:10.1038/nature14177

- 545 25. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, et al. Million Veteran  
546 Program: A mega-biobank to study genetic influences on health and disease. *J Clin*  
547 *Epidemiol.* 2016;70: 214–223. doi:10.1016/j.jclinepi.2015.09.016
- 548 26. Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association  
549 study of blood pressure determinants in over 750,000 individuals. *Nat Genet.* 2019;51: 51–  
550 62. doi:10.1038/s41588-018-0303-9
- 551 27. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, et al. Discovery of  
552 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million  
553 participants in a multi-ancestry meta-analysis. *Nat Genet.* 2020;52: 680–691.  
554 doi:10.1038/s41588-020-0637-y
- 555 28. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide  
556 association and Mendelian randomisation analysis provide insights into the pathogenesis  
557 of heart failure. *Nat Commun.* 2020;11: 163. doi:10.1038/s41467-019-13690-5
- 558 29. Munz M, Richter GM, Loos BG, Jepsen S, Divaris K, Offenbacher S, et al. Genome-wide  
559 association meta-analysis of coronary artery disease and periodontitis reveals a novel  
560 shared risk locus. *Sci Rep.* 2018;8: 13678. doi:10.1038/s41598-018-31980-8
- 561 30. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian  
562 regression and continuous shrinkage priors. *Nat Commun.* 2019;10: 1776.  
563 doi:10.1038/s41467-019-09718-5
- 564 31. Barber TM, Hanson P, Weickert MO, Franks S. Obesity and Polycystic Ovary Syndrome:  
565 Implications for Pathogenesis and Novel Management Strategies. *Clin Med Insights*  
566 *Reprod Health.* 2019;13: 1179558119874042. doi:10.1177/1179558119874042
- 567 32. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. An atlas of  
568 genetic correlations across human diseases and traits. *Nat Genet.* 2015;47: 1236–1241.  
569 doi:10.1038/ng.3406
- 570 33. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal  
571 mediation analysis. *J Stat Softw.* 2014. Available:  
572 <https://oar.princeton.edu/jspui/handle/88435/pr1gj2f>
- 573 34. Di Guardo F, Ciotta L, Monteleone M, Palumbo M. Male Equivalent Polycystic Ovarian  
574 Syndrome: Hormonal, Metabolic, and Clinical Aspects. *Int J Fertil Steril.* 2020;14: 79–83.  
575 doi:10.22074/ijfs.2020.6092
- 576 35. Recabarren SE, Smith R, Rios R, Maliqueo M, Echiburú B, Codner E, et al. Metabolic  
577 profile in sons of women with polycystic ovary syndrome. *J Clin Endocrinol Metab.*  
578 2008;93: 1820–1826. doi:10.1210/jc.2007-2256
- 579 36. Subramaniam K, Tripathi A, Dabadghao P. Familial clustering of metabolic phenotype in  
580 brothers of women with polycystic ovary syndrome. *Gynecol Endocrinol.* 2019;35: 601–  
581 603. doi:10.1080/09513590.2019.1566451

- 582 37. Khan SS, Ning H, Wilkins JT, Allen N, Carnethon M, Berry JD, et al. Association of Body  
583 Mass Index With Lifetime Risk of Cardiovascular Disease and Compression of Morbidity.  
584 *JAMA Cardiol.* 2018;3: 280–287. doi:10.1001/jamacardio.2018.0022
- 585 38. Papadakis G, Kandaraki E, Papalou O, Vryonidou A, Diamanti-Kandarakis E. Is  
586 cardiovascular risk in women with PCOS a real risk? Current insights. *Minerva Endocrinol.*  
587 2017;42: 340–355. doi:10.23736/S0391-1977.17.02609-8
- 588 39. Rubin KH, Glintborg D, Nybo M, Abrahamsen B, Andersen M. Development and Risk  
589 Factors of Type 2 Diabetes in a Nationwide Population of Women With Polycystic Ovary  
590 Syndrome. *J Clin Endocrinol Metab.* 2017;102: 3848–3857. doi:10.1210/jc.2017-01354
- 591 40. Diamanti-Kandarakis E, Dunaif A. Insulin resistance and the polycystic ovary syndrome  
592 revisited: an update on mechanisms and implications. *Endocr Rev.* 2012;33: 981–1030.  
593 doi:10.1210/er.2011-1034
- 594 41. Escobar-Morreale HF. Polycystic ovary syndrome: definition, aetiology, diagnosis and  
595 treatment. *Nat Rev Endocrinol.* 2018;14: 270–284. doi:10.1038/nrendo.2018.24
- 596 42. Rothman KJ. BMI-related errors in the measurement of obesity. *Int J Obes (Lond).* 2008;32  
597 Suppl 3: S56-59. doi:10.1038/ijo.2008.87
- 598 43. Chrysant SG, Chrysant GS. The single use of body mass index for the obesity paradox is  
599 misleading and should be used in conjunction with other obesity indices. *Postgrad Med.*  
600 2019;131: 96–102. doi:10.1080/00325481.2019.1568019
- 601 44. Li C, Engström G, Hedblad B, Calling S, Berglund G, Janzon L. Sex differences in the  
602 relationships between BMI, WHR and incidence of cardiovascular disease: a population-  
603 based cohort study. *Int J Obes (Lond).* 2006;30: 1775–1781. doi:10.1038/sj.ijo.0803339
- 604 45. Venkatesh SS, Ferreira T, Benonisdottir S, Rahmioglu N, Becker CM, Granne I, et al. The  
605 role of obesity in female reproductive conditions: A Mendelian Randomisation study.  
606 *Genetic and Genomic Medicine*; 2021 Jun. doi:10.1101/2021.06.01.21257781
- 607 46. Rodgers RJ, Avery JC, Moore VM, Davies MJ, Azziz R, Stener-Victorin E, et al. Complex  
608 diseases and co-morbidities: polycystic ovary syndrome and type 2 diabetes mellitus.  
609 *Endocr Connect.* 2019;8: R71–R75. doi:10.1530/EC-18-0502
- 610 47. Piltonen TT, Ruokojärvi M, Karro H, Kujanpää L, Morin-Papunen L, Tapanainen JS, et al.  
611 Awareness of polycystic ovary syndrome among obstetrician-gynecologists and  
612 endocrinologists in Northern Europe. *PLOS ONE.* 2019;14: e0226074.  
613 doi:10.1371/journal.pone.0226074
- 614 48. Dokras A, Saini S, Gibson-Helm M, Schulkin J, Cooney L, Teede H. Gaps in knowledge  
615 among physicians regarding diagnostic criteria and management of polycystic ovary  
616 syndrome. *Fertil Steril.* 2017;107: 1380-1386.e1. doi:10.1016/j.fertnstert.2017.04.011
- 617 49. Mott MM, Kitos NR, Coviello AD. Practice Patterns in Screening for Metabolic Disease in  
618 Women with PCOS of Diverse Race-Ethnic Backgrounds. *Endocr Pract.* 2014;20: 855–  
619 863. doi:10.4158/EP13414.OR

620

621 **Table 1. Hypertension is Significantly Modified by Sex for PCOS<sub>PRS</sub>.**

622

Phenotype	Estimate	SE	Z-Value	P	OR	LCI	UCI	Cases	Controls
Hypertension	-0.05	0.02	-2.69	7.24E-03	0.95	0.91	0.99	24,574	35,352
Essential hypertension	-0.05	0.02	-2.66	7.71E-03	0.95	0.91	0.99	24,159	35,352
Hypertensive heart disease	-0.11	0.04	-2.51	0.01	0.89	0.82	0.98	2,696	35,352
Coronary atherosclerosis	-0.04	0.03	-1.46	0.14	0.96	0.91	1.01	9,123	49,700
Ischemic Heart Disease	-0.03	0.03	-1.28	0.20	0.97	0.92	1.02	10,015	49,700
Loss of teeth or edentulism	-0.17	0.14	-1.21	0.23	0.84	0.64	1.11	184	61,061
Obesity	0.03	0.03	0.92	0.36	1.03	0.97	1.09	5,835	51,152
Type 2 diabetes	0.02	0.02	0.90	0.37	1.02	0.97	1.07	8,990	51,286
Diabetes mellitus	0.02	0.02	0.65	0.52	1.02	0.97	1.06	9,424	51,286
Overweight, obesity and other hyperalimentation	0.01	0.03	0.23	0.82	1.01	0.95	1.06	6,841	51,152

623

624 We used a logistic regression model to test the interaction of effect between PCOS<sub>PRS</sub> and sex.

625 Models were adjusted for age, sex, genetic ancestry, and the interaction term. Estimate effects

626 are in the direction of females.

627

628 **Table 2. PCOS Shares Genetic Architecture with Cardiometabolic Traits**

629

Traits	Genetic Correlation	SE	Z	P
Diastolic Blood Pressure	-3.60%	0.07	-0.54	0.59
Systolic Blood Pressure	9.33%	0.05	1.79	0.07
Pulse Pressure	13.46%	0.05	2.54	0.01
Type 2 Diabetes	29.26%	0.05	5.92	3.25e-09
Heart Failure	26.51%	0.09	3.02	0.003
Coronary Artery Diseases	17.35%	0.09	1.95	0.05
Body Mass Index	43.32%	0.05	8.85	8.77e-19

630

631



632 **Table 3. BMI Genetic Variance is a Causal Risk Factor for PCOS and T2D**

633

Exposure	Mediator	Outcome	ADE $\beta$ (95% CI) <i>P</i> = <2e-16	ACME $\beta$ (95% CI) <i>P</i> = <2e-16	Proportion Mediated (95% CI) <i>P</i> = <2e-16
BMI <sub>PRS</sub>	PCOS	T2D	0.025 (0.014-0.037) <i>P</i> = <2e-16	0.002 (6.92E-04-0.004) <i>P</i> = <2e-16	0.071 (0.025-0.144) <i>P</i> = <2e-16
	T2D	PCOS	0.025 (0.013-0.037) <i>P</i> = <2e-16	0.002 (0.001-0.004) <i>P</i> = <2e-16	0.090 (0.039-0.191) <i>P</i> = <2e-16
BMI <sub>residual</sub>	PCOS	T2D	0.037 (0.023-0.052) <i>P</i> = <2e-16	0.006 (0.002-0.010) <i>P</i> = <2e-16	0.136 (0.049-0.254) <i>P</i> = <2e-16
	T2D	PCOS	0.127 (0.104-0.150) <i>P</i> = <2e-16	0.003 (8.12-04- 0.005) <i>P</i> = 0.004	0.021 (0.006-0.042) <i>P</i> = 0.004

634

635 ADE = average direct effect, ACME = average causal mediation effect, CI = confidence interval

636

637

638 **Table 4. Genetic effects of T2D Contributes to Environmental BMI.**

639

Sex	Exposure	Mediator	Outcome	ADE $\beta$ (95% CI) <i>P</i>	ACME $\beta$ (95% CI) <i>P</i>	Proportion Mediated (95% CI) <i>P</i>
Sex- Combined	PCOS <sub>PRS</sub>	T2D	BMI <sub>residual</sub>	0.041 (-0.055-0.147) <i>P</i> = 0.44	0.059 (0.033-0.088) <i>P</i> = <2e-16	0.589 (-1.112-3.248) <i>P</i> = 0.06
Females				0.024 (-0.102-0.172) <i>P</i> = 0.78	0.059 (0.041-0.121) <i>P</i> = <2e-16	0.652 (-4.638-6.128) <i>P</i> = 0.18
Males				0.040 (-0.101-0.176) <i>P</i> = 0.58	0.038 (0.004-0.074) <i>P</i> = 0.02	0.342 (-2.354-4.320) <i>P</i> = 0.31
Females	T2D <sub>PRS</sub>	PCOS	BMI <sub>residual</sub>	0.268 (0.018-0.519) <i>P</i> = 0.04	0.139 (0.046-0.242) <i>P</i> = 0.002	0.338 (0.122-0.863) <i>P</i> = 0.004

640

641

642  
643  
644  
645

**Table 5. Genetically Regulated BMI is a Stronger Mediator in PCOS<sub>PRS</sub> Shared Metabolic Pathways.**

Sex	Mediator	Outcome	ADE $\beta$ (95% CI) <i>P</i>	ACME $\beta$ (95% CI) <i>P</i>	Proportion Mediated (95% CI) <i>P</i>
Sex- Combined	BMI	T2D	0.009 (0.004-0.014) <i>P</i> = <2e-16	0.005 (0.004-0.007) <i>P</i> = <2e-16	0.367 (0.250 - 0.585) <i>P</i> = <2e-16
	BMI Residuals		0.0134 (0.008- 0.019) <i>P</i> = <2e-16	0.001 (-1.76E-04- 0.002) <i>P</i> = 0.09	0.072 (-0.011- 0.155) <i>P</i> = 0.09
	BMI	Hypertension	0.007 (-2.13e-04-0.013) <i>P</i> = 0.062	0.005 (0.004-0.006) <i>P</i> = <2e-16	0.435 (0.269-1.043) <i>P</i> = <2e-16
	BMI Residuals		0.012 (0.005-0.0178) <i>P</i> = 0.002	8.89e-04 (-3.13E-04-0.002) <i>P</i> = 0.13	0.072 (-0.036-0.207) <i>P</i> = 0.13
	BMI	Hypertensive Heart Disease	0.008 (0.003-0.013) <i>P</i> = 0.006	0.002 (0.001-0.003) <i>P</i> = <2e-16	0.237 (0.144-0.457) <i>P</i> = 0.002
	BMI Residuals		0.009 (0.004-0.014) <i>P</i> = 0.002	0.0006 (-0.0001-0.001) <i>P</i> = 0.09	0.062 (-0.011-0.152) <i>P</i> = 0.09
Females	BMI	T2D	0.012 (0.005-0.019) <i>P</i> = <2e-16	0.005 (0.004-0.007) <i>P</i> = <2e-16	0.315 (0.208-0.535) <i>P</i> = <2e-16
	BMI Residuals		0.016 (0.009-0.023) <i>P</i> = <2e-16	9.64e-04 (-6.00e-04-0.003) <i>P</i> = 0.22	0.058 (-0.037-0.154) <i>P</i> = 0.22
	BMI	Hypertension	0.002 (-0.007-0.010) <i>P</i> = 0.73	0.006 (0.004-0.007) <i>P</i> = <2e-16	0.742 (-2.165-5.363) <i>P</i> = 0.08

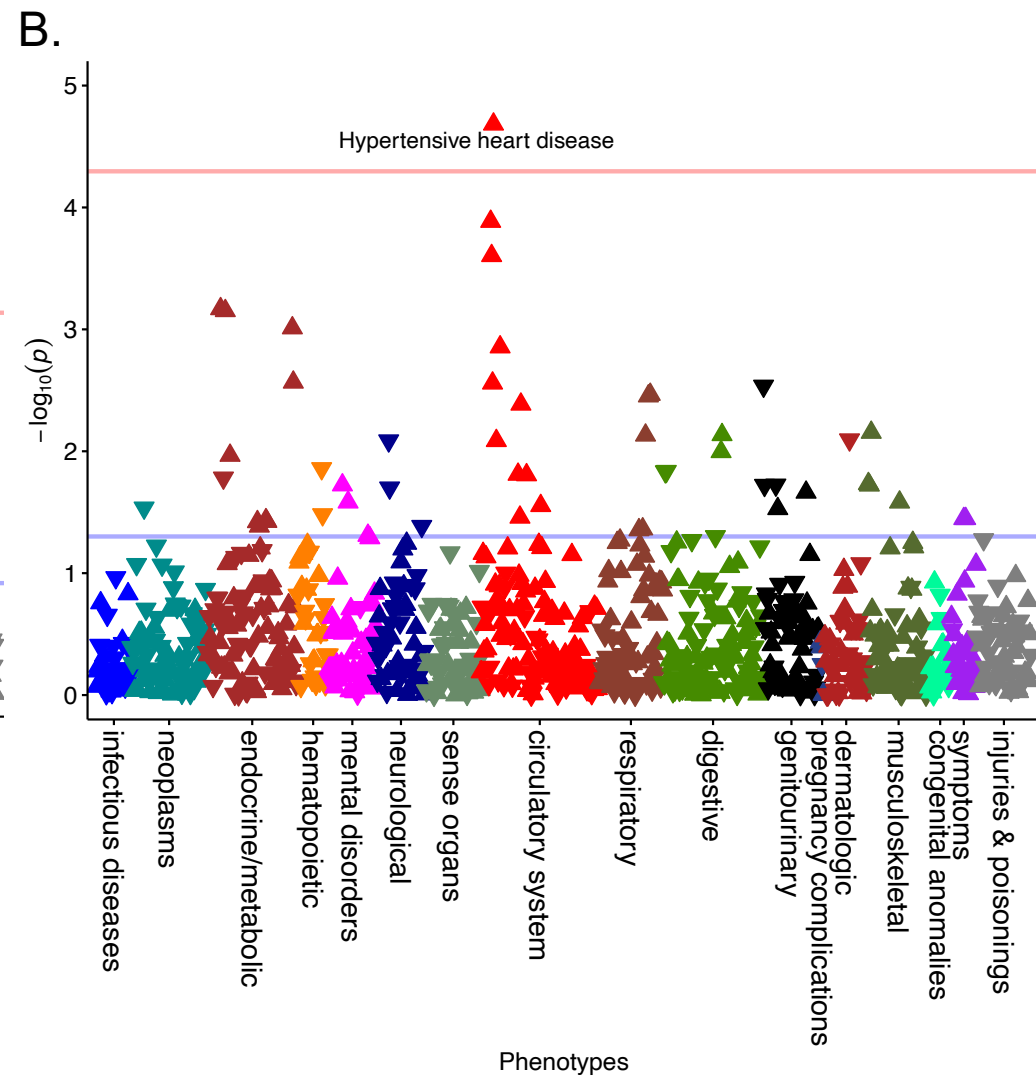
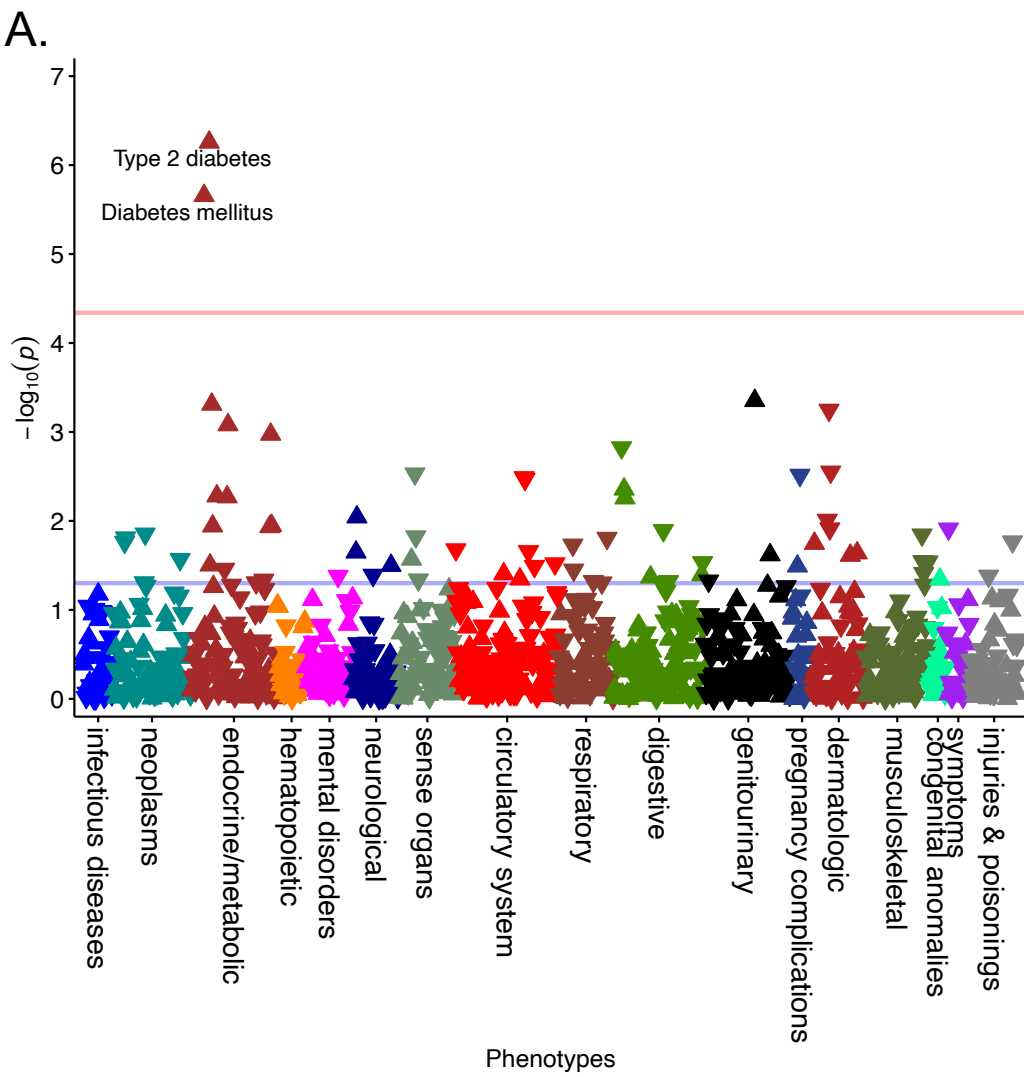
	BMI Residuals		0.007 (-0.002-0.015) <i>P</i> = 0.11	9.01e-04 (-7.86E-04-0.002) <i>P</i> = 0.26	0.112 (-0.426-0.806) <i>P</i> = 0.30
	BMI	Hypertensive Heart Disease	0.003 (-0.002- 0.009) <i>P</i> = 0.31	0.002 (0.001- 0.003) <i>P</i> = <2e-16	0.378 (-1.602-3.278) <i>P</i> = 0.082
	BMI Residuals		0.005 (-9.79e-04-0.011) <i>P</i> = 0.11	5.50e-04 (-0.0002-0.001) <i>P</i> = 0.14	0.097 (-0.367-0.781) <i>P</i> = 0.20
Males	BMI	T2D	0.007 (-0.002-0.016) <i>P</i> = 0.13	0.005 (0.003-0.007) <i>P</i> = <2e-16	0.430 (0.186-1.377) <i>P</i> = 0.02
	BMI Residuals		0.011 (0.001-0.020) <i>P</i> = 0.03	0.001 (-8.26e-04-0.003) <i>P</i> = 0.27	0.086 (-0.115-0.421) <i>P</i> = 0.28
	BMI	Hypertension	0.014 (0.004-0.024) <i>P</i> = 0.006	0.004 (0.003-0.006) <i>P</i> = <2e-16	0.237 (0.127-0.530) <i>P</i> = 0.002
	BMI Residuals		0.018 (0.009-0.027) <i>P</i> = <2e-16	8.65e-04 (-8.84e-04-0.003) <i>P</i> = 0.30	0.046 (-0.049 -0.156) <i>P</i> = 0.30
	BMI	Hypertensive Heart Disease	0.015 (0.006- 0.025) <i>P</i> = <2e-16	0.003 (0.001- 0.004) <i>P</i> = <2e-16	0.148 (0.059-0.349) <i>P</i> = <2e-16
	BMI Residuals		0.017 (0.008- 0.027) <i>P</i> = <2e-16	0.0006 (-8.82e-04- 0.002) <i>P</i> = 0.43	0.034 (-0.063-0.139) <i>P</i> = 0.43

646

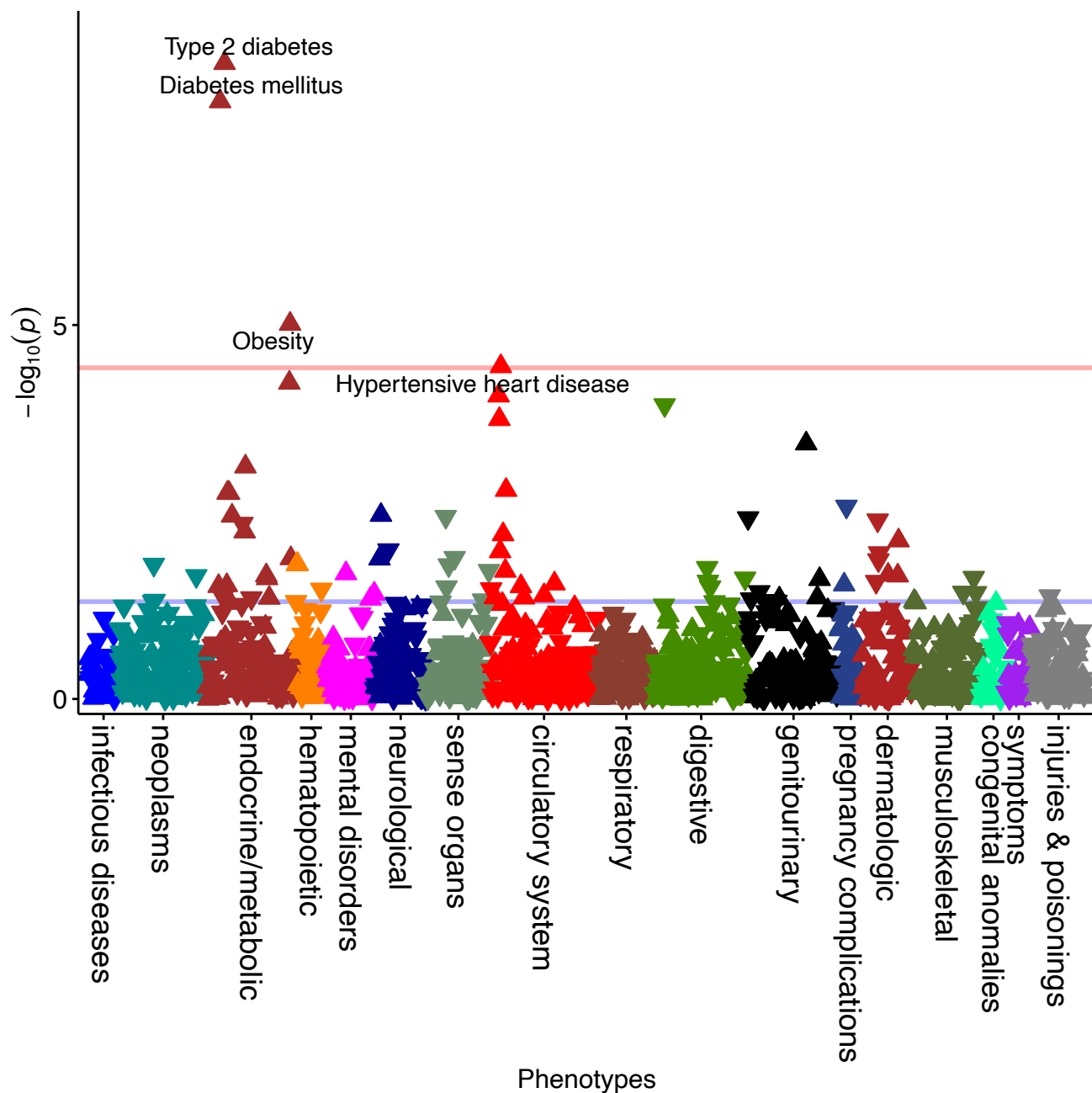
647 PCOS<sub>PRS</sub> was analyzed as the exposure variable. *P* < 0.05 was considered statistically significant. ADE = average direct effect,

648 ACME = average causal mediation effect, CI = confidence interval

649



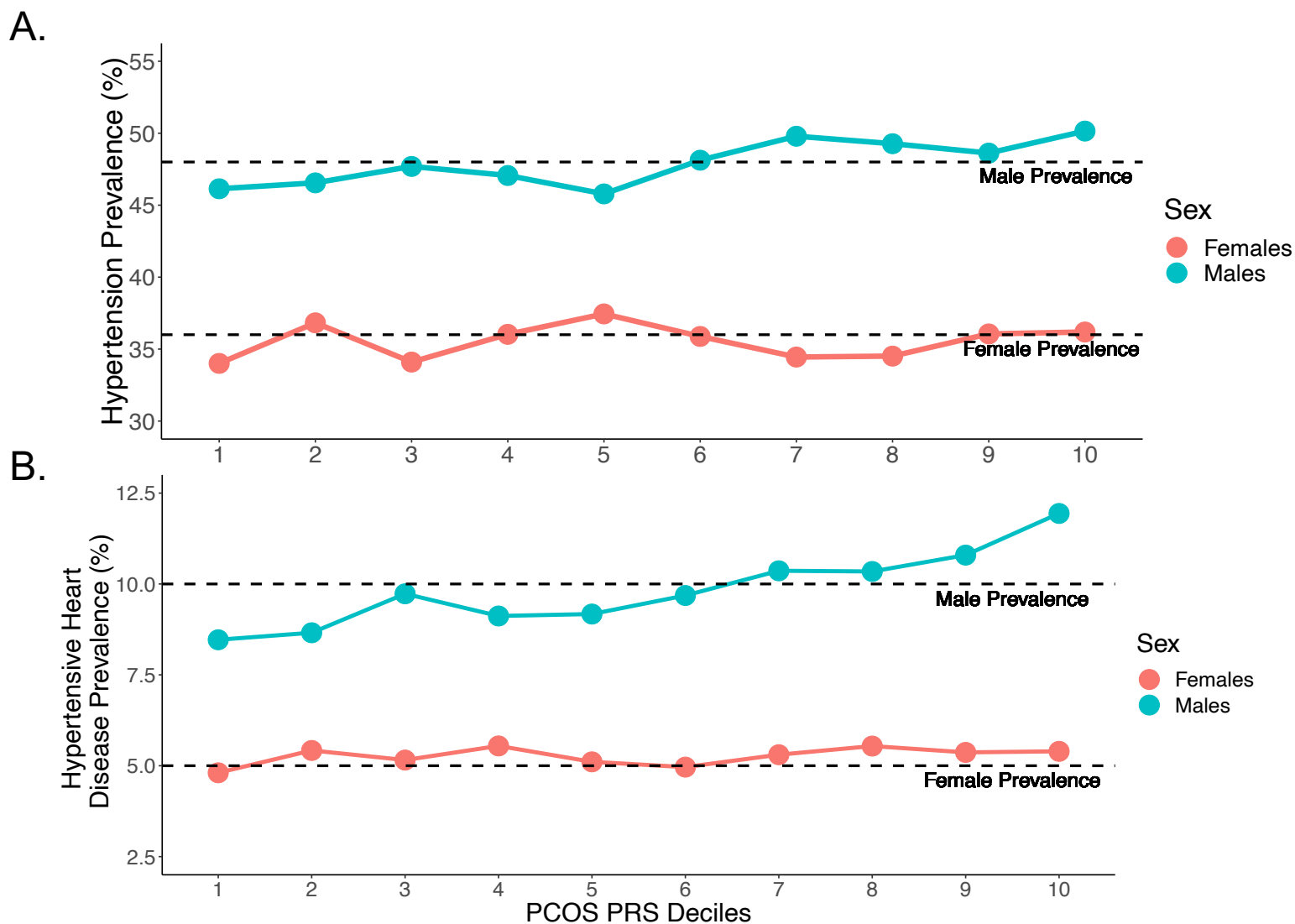
650 **Figure 1. Metabolic Associations Vary by Sex for Genetically High-Risk Individuals.** PCOS<sub>PRS</sub> phenome-wide association study  
 651 results are displayed in the Manhattan plots for (A) females and (B) males. The red line represents the Bonferroni correction of  $P =$   
 652  $4.56e-05$  and  $P = 5.07e-05$ , respectively. Arrows in the upward direction represent increased risk.  
 653



654  
655 **Figure 2. Sex Combined PCOS<sub>PRS</sub> PheWAS Analysis Increases Detection of Associations.**

656 The phenome-wide association study (PheWAS) was performed on the full dataset (females  
657 and males) to increase power. Significant results that passed Bonferroni correction ( $P = 3.73e-$   
658  $05$ ) are annotated in the Manhattan plot.

659  
660  
661



662

663 **Figure 3. The Prevalence of Hypertensive Metabolic Traits Differs Across the PCOS<sub>PRS</sub>**

664 **Strata.** (A) Top Figure: Hypertension, (B) Bottom Figure: Hypertensive Heart Disease. The

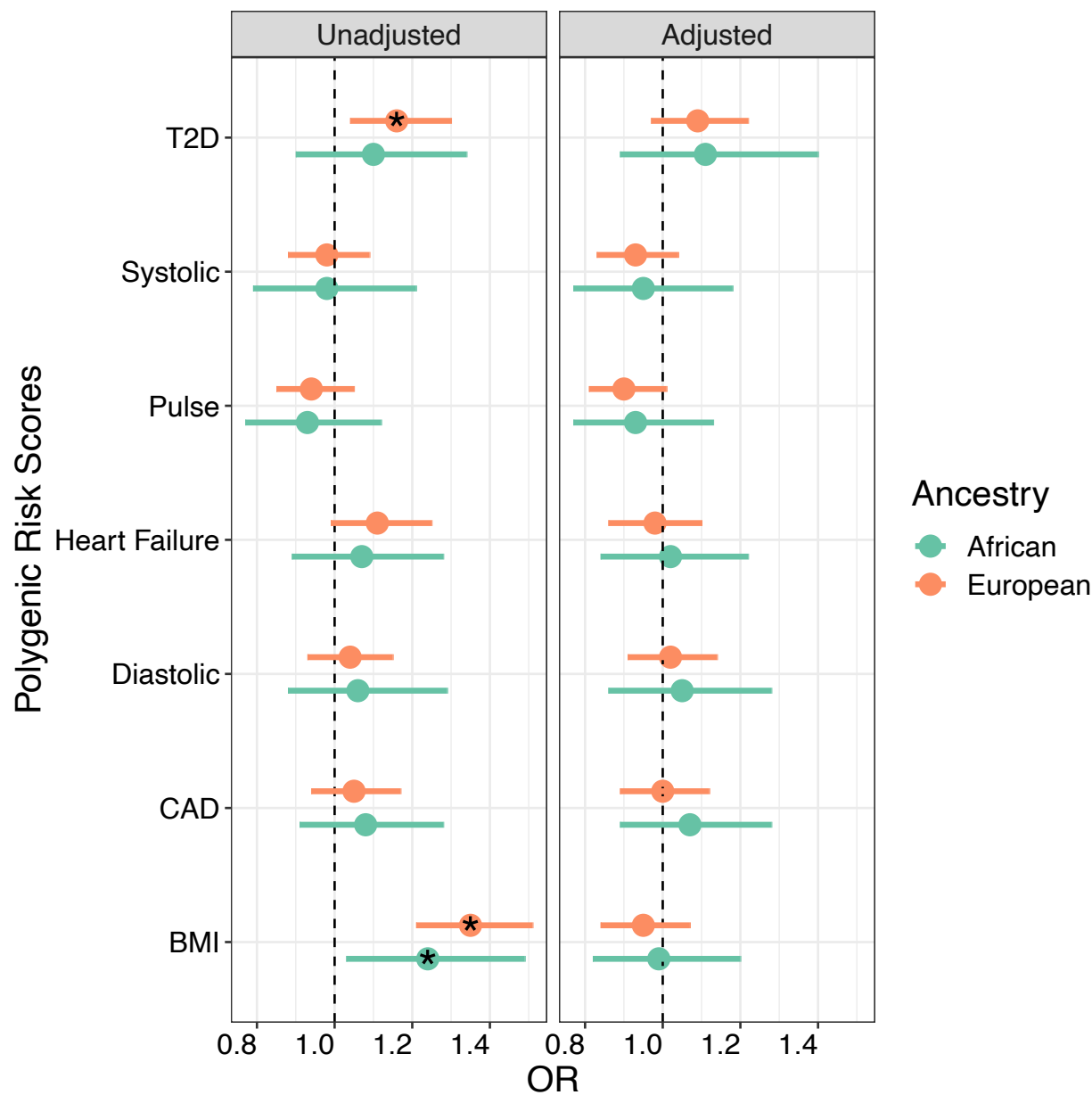
665 prevalence of the significant phenotypes from the sex interaction analysis are shown. The

666 prevalence for hypertension in our dataset was 48% for males and 36% for females. The

667 prevalence for hypertensive heart disease was 10% for males and 5% for females.

668

669

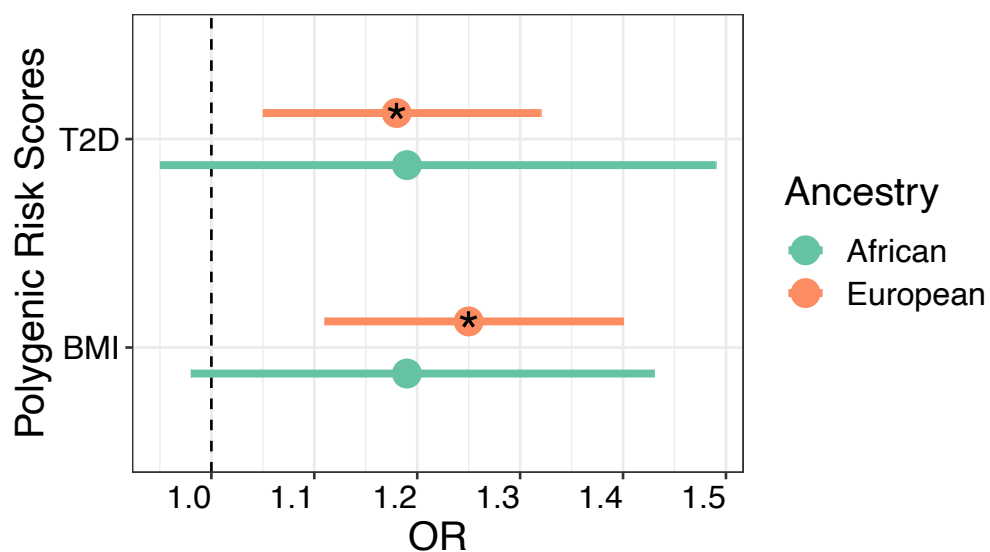


670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680

**Figure 4. T2D<sub>PRS</sub> and BMI<sub>PRS</sub> are Associated with PCOS Case Status.** The results from the logistic regression analysis between PCOS diagnosis and polygenic risk scores (PRS) generated for significant metabolic traits from the PheWAS analysis are displayed in the forest plot. One model was unadjusted for body mass index (unadjusted) and the other model was adjusted for body mass index (adjusted). \* indicates P < 0.05



681



682

683 **Figure 5. T2D<sub>PRS</sub> is Associated with PCOS After Accounting for BMI<sub>residual</sub>.** The logistic  
684 regression model between T2D<sub>PRS</sub> and PCOS diagnosis was covaried for age, the top ten  
685 principal components for ancestry, and BMI<sub>residual</sub>. BMI<sub>residual</sub> represents BMI after the removal of  
686 BMI genetic variance from the variable. \* indicates P < 0.05