

# 1 **The role of obesity in female reproductive conditions:**

## 2 **A Mendelian Randomisation study**

3 Samvida S. Venkatesh<sup>1,2\*</sup>, Teresa Ferreira<sup>1,3</sup>, Stefania Benonisdottir<sup>1</sup>, Nilufer Rahmioglu<sup>2,3</sup>,  
4 Christian M. Becker<sup>3</sup>, Ingrid Granne<sup>3</sup>, Krina T. Zondervan<sup>2,3</sup>, Michael V. Holmes<sup>4,5</sup>, Cecilia M.  
5 Lindgren<sup>1,2,3,4,6</sup>, Laura B. L. Wittemans<sup>1,3\*</sup>

6  
7 <sup>1</sup>Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of  
8 Oxford, Oxford, United Kingdom

9 <sup>2</sup>Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford,  
10 Oxford, United Kingdom

11 <sup>3</sup>Nuffield Department of Women's and Reproductive Health, Medical Sciences Division,  
12 University of Oxford, Oxford, United Kingdom

13 <sup>4</sup>Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

14 <sup>5</sup>Medical Research Council Population Health Research Unit, University of Oxford, Oxford,  
15 United Kingdom

16 <sup>6</sup>Broad Institute of Harvard and MIT, Cambridge, MA, United States of America

17  
18 \* Corresponding Author

19 E-mail: [samvida@well.ox.ac.uk](mailto:samvida@well.ox.ac.uk) (S.S.V.); [laura.wittemans@wrh.ox.ac.uk](mailto:laura.wittemans@wrh.ox.ac.uk) (L.B.L.W.)

## 20 **Abstract**

21 **Background** - Obesity is observationally associated with altered risk of many female reproductive  
22 conditions. These include polycystic ovary syndrome (PCOS), abnormal uterine bleeding,  
23 endometriosis, infertility, and pregnancy-related disorders. However, the roles and mechanisms of  
24 obesity in the aetiology of reproductive disorders remain unclear.

25 **Methods and Findings** - We estimated observational and genetically predicted causal associations  
26 between obesity, metabolic hormones, and female reproductive conditions using logistic  
27 regression, generalised additive models, and Mendelian randomisation (two-sample, non-linear,  
28 and multivariable) applied to data from UK Biobank and publicly available genome-wide  
29 association studies (GWAS).

30 Body mass index (BMI), waist-hip ratio (WHR), and WHR adjusted for BMI (WHRadjBMI) were  
31 observationally (odds ratios (ORs) = 1.02 – 1.87 per 1 S.D. obesity trait) and  
32 causally (ORs = 1.06 – 2.09) associated with uterine fibroids (UF), PCOS, heavy menstrual  
33 bleeding (HMB), and pre-eclampsia. Causal effect estimates of WHR and WHRadjBMI, but not  
34 BMI, were attenuated compared to their observational counterparts. Genetically predicted visceral  
35 adipose tissue mass was causal for the development of HMB, PCOS, and pre-eclampsia (ORs =  
36 1.01 - 3.38). Increased waist circumference also posed a higher causal risk (ORs = 1.16 – 1.93) for  
37 the development of these disorders and UF than did increased hip circumference (ORs = 1.06 –  
38 1.10). Leptin, fasting insulin, and insulin resistance each mediated between 20% -50% of the total  
39 causal effect of obesity on pre-eclampsia. Reproductive conditions clustered based on shared  
40 genetic components of their aetiological relationships with obesity.

41 **Conclusions** - In this first systematic, large-scale, genetics-based analysis of the aetiological  
42 relationships between obesity and female reproductive conditions, we found that common indices  
43 of overall and central obesity increased risk of reproductive disorders to heterogenous extents,  
44 mediated by metabolic hormones. Our results suggest exploring the mechanisms mediating the  
45 causal effects of overweight and obesity on gynaecological health to identify targets for disease  
46 prevention and treatment.

## 47 **Introduction**

48 Obesity is commonly understood as the excess accumulation of body fat which leads to increased  
49 health risks. In women, body mass index (BMI) is associated with increased prevalence of  
50 gynecological conditions, including excessive and abnormal menstrual bleeding (1, 2),  
51 endometriosis and uterine fibroids (UF) (3, 4), polycystic ovary syndrome (PCOS) (5, 6),  
52 complications of pregnancy such as pre-eclampsia and eclampsia (7), miscarriage (8, 9), and  
53 infertility (10, 11). These are often non-linear and heterogeneous relationships. While the risk of  
54 anovulatory infertility and recurrent miscarriages are highest in obese women, underweight  
55 women also have increased risk of infertility (9, 12). The association of BMI with endometriosis  
56 varies by disease severity, as women with advanced-stage endometriosis have lower BMI than  
57 those with minimal disease, and the inverse BMI-endometriosis association is stronger in women  
58 with infertility (13, 14). Finally, although the severity of PCOS and menstrual disorders increases  
59 with overall obesity, women presenting with these conditions are more likely to store fat in the  
60 abdominal region, regardless of their BMI (2, 5).

61 Observational epidemiological studies are limited in assessing causality, due to confounding and  
62 reverse causation. The Mendelian randomisation (MR) framework is a genetics-based instrumental

63 variable approach that relies on the random and fixed assignment of genetic variants at conception  
64 to estimate the causal effect size of genetically predicted exposures on an outcome. MR has  
65 previously indicated causal effects of genetically predicted BMI on the development of some  
66 subtypes of ovarian cancer (odds ratio (OR) = 1.29 per 5 units of BMI) (15), endometrial cancer  
67 (OR = 2.06 per 5 units of BMI) (16), and PCOS (OR = 4.89 per 1 S.D. higher BMI) (17). However,  
68 the aetiological role of obesity and body fat distribution on many other female reproductive  
69 diseases has not been reported. It is especially relevant to investigate the effects of fat distribution,  
70 as there are intricate metabolic and endocrine links between adipose tissue and female reproductive  
71 organs. Yet, causal investigations of such relationships are lacking.

72 Leptin, which is a hormone secreted by adipocytes, and elevated in individuals with obesity, is  
73 increased in women with endometriosis, UF, and adverse pregnancy outcomes, even when  
74 adjusted for BMI (7, 18-21). Obesity-induced insulin resistance additionally increases the risk and  
75 severity of PCOS and pre-eclampsia by dysregulating steroid hormone and metabolic pathways  
76 (5, 22, 23). The dysregulation of sex hormones, including oestrogen and testosterone, is likely to  
77 play a role in the obesity-driven development of female reproductive disorders due to its close  
78 associations with body fat (22, 24). Yet, the causal impact of these factors in mediating the  
79 relationships between obesity and gynaecological diseases has not been detailed.

80 Here, we apply logistic regression, generalised additive models, two-sample, non-linear, and  
81 multivariable MR to dissect the relationships of overall obesity and body fat distribution with a  
82 range of female reproductive disorders, and investigate the mediating role of metabolic factors  
83 including leptin and insulin.

84

## 85 **Methods**

### 86 **Observational associations in UK Biobank**

87 UK Biobank (UKBB) is a prospective UK-based cohort study with approximately 500,000  
88 participants aged 40-69 at recruitment on whom a range of medical, environmental, and genetic  
89 information is collected (25). We included 257,193 individuals self-identifying as females of white  
90 ancestry in UKBB in our analyses. Baseline measurements of BMI (total body weight (kg) /  
91 standing height<sup>2</sup> (m<sup>2</sup>)) and waist-to-hip ratio (WHR) (waist circumference (cm) / hip  
92 circumference (cm)), and WHR adjusted for BMI (WHRadjBMI) were used to estimate general  
93 and central obesity, respectively. Cases of reproductive conditions were identified based on ICD9  
94 and ICD10 primary and secondary diagnoses from hospital inpatient data, self-reported illness  
95 codes, and primary care records (**Table A in S1 Table**). We fitted logistic regression models to  
96 estimate the associations of BMI, WHR, and WHRadjBMI with prevalence of endometriosis  
97 (7,703 cases, 249,490 controls), heavy menstrual bleeding (17,229 cases, 239,964 controls),  
98 infertility (2,194 cases, 254,999 controls), self-reported stillbirth, spontaneous miscarriage or  
99 termination (81,102 cases, 176,091 controls), PCOS (746 cases, 256,447 controls), pre-eclampsia  
100 (2,242 cases, 254,951 controls), and uterine fibroids (19,192 cases, 238,001 controls). Case  
101 definitions for pre-eclampsia included eclampsia cases to capture cases in which the former may  
102 have developed into the latter. For each disease, individuals not included in the case group were  
103 used as controls. BMI, WHR and WHRadjBMI were adjusted for age, age-squared, assessment  
104 centre, and smoking status. The residuals were rank-based inverse normally transformed. Multiple  
105 testing correction was applied using the false discovery rate (FDR) to evaluate statistical  
106 significance while minimising false negatives (26).

107 We also tested associations without adjustment for smoking status, as it has previously been  
108 suggested that higher BMI increases risk of smoking (27) and adjustment for both could therefore  
109 induce collider bias. Adjustment for menopause status was not performed as up to 42% of women  
110 with reproductive disorders in UKBB report being unsure of their menopause status as compared  
111 to 16% of women who do not have a recorded history or presence of a reproductive condition  
112 **(Table A in S1 Table).**

113 To evaluate the presence of non-linear observational associations between obesity and each  
114 reproductive trait, fractional polynomial regression following the closed test procedure was  
115 performed using the mfp v1.5.2 R package (28). This algorithm tests for the presence of an overall  
116 association, the likelihood of non-linearity, and selects the best-fitting fractional polynomial  
117 function. We also fitted generalised additive models (GAM) to the same data, allowing for  
118 smoothing of the obesity trait with splines, using the mgcv 1.8-31 R package (29). All models were  
119 adjusted for age, age-squared, assessment centre, and smoking status. Model fits were compared  
120 with Akaike's Information Criterion (AIC) (30).

## 121 **Two-sample Mendelian Randomisation**

122 Genetic instruments for BMI, WHR, and WHRadjBMI were selected based on the sentinel variants  
123 at genome-wide significant loci ( $P < 5E-9$ ) reported in the largest publicly available European  
124 ancestry GWAS of Genetic Investigation of ANthropometric Traits (GIANT) and UKBB (max N  
125 individuals = 806,801) (31). Similarly, genetic instruments for predicted visceral adipose tissue  
126 (VAT) mass (N individuals = 325,153) (32), waist circumference (N individuals = 462,166) and  
127 hip circumference (N individuals = 462,117) (33), and waist-specific and hip-specific WHR (N  
128 individuals = 18,330) (34) were selected based on the largest publicly available GWAS.

129 Three instrument weighting strategies were considered where sex-stratified GWAS results were  
130 available: (i) SNPs from combined-sexes GWAS with combined-sexes weights (effect sizes), (ii)  
131 combined-sexes SNPs with female-specific weights, or (iii) female-specific SNPs with female-  
132 specific weights. The method of female-specific SNPs with female-specific weights produced the  
133 strongest instruments as evaluated by F-statistics and was thus chosen for analysis (**Table B in S1**  
134 **Table**). Additionally, due to concerns of ascertainment bias in UKBB (35, 36), sensitivity analyses  
135 with combined-sexes instruments (combined-sexes SNPs with combined-sexes weights) were also  
136 performed.

137 Associations of the genetic instruments for obesity traits with female reproductive diseases were  
138 obtained by performing a fixed-effect inverse-variance weighted meta-analysis of publicly  
139 available GWAS summary statistics from two large biobank projects - FinnGen and UKBB (37).  
140 The meta-analysis was performed using METAL (38) by matching the relevant ICD codes (**Table**  
141 **C in S1 Table**) for the following traits: infertility (4,996 cases, 421,223 controls), pre-eclampsia  
142 (2,711 cases, 480,373 controls), and uterine fibroids (21,835 cases, 456,551 controls). For  
143 endometriosis, summary statistics were obtained by request from a recent European ancestry  
144 GWAS (39) and meta-analysed as above with publicly available FinnGen and UKBB summary  
145 statistics (12,210 cases, 450,183 controls). For heavy menstrual bleeding (HMB) (9,813 cases,  
146 210,946 controls), sporadic miscarriage, i.e. 1-2 miscarriages (50,060 cases, 174,109 controls),  
147 and multiple consecutive miscarriage, i.e.  $\geq 3$  consecutive miscarriages (750 cases, 150,215  
148 controls), publicly available summary statistics were obtained from recent European ancestry  
149 GWAS that include UKBB individuals (3, 40). For PCOS, estimates were based on a fixed-effect  
150 inverse variance-weighted meta-analysis of published GWAS summary statistics (38), publicly  
151 available GWAS results by FinnGen, and a European-ancestry GWAS run in UKBB using SAIGE

152 (11,186 cases, 273,812 controls). As a sensitivity analysis, all MR tests were performed using  
153 disease association estimates based on FinnGen only, where available, to alleviate bias due to  
154 sample overlap between the exposure and outcome GWAS sources.

155 Power to detect MR associations was calculated using two methods, one designed for general two-  
156 sample MR (41) and the other for MR performed on binary outcomes (42). Briefly, these methods  
157 calculate power by accounting for GWAS sample size, proportion of cases in case-control GWAS,  
158 and variance explained by genetic instruments for the exposure. The power to detect a true odds  
159 ratio (OR) association of 1.1 or more extreme at an unadjusted significance level of 0.05 was  
160 estimated.

161 Instrument SNPs were extracted from the outcome GWAS results, harmonised for consistency in  
162 the alleles, and MR was performed using the TwoSample MR v0.5.4 R package (43). Three  
163 methods for MR - inverse-variance weighted (IVW), MR-Egger, and weighted median - were  
164 evaluated, and the best method was selected via Rucker's framework (44). Briefly, this framework  
165 advises to choose the MR method with least heterogeneity as assessed by Cochran's Q-statistic,  
166 while accounting for the trade-off between power and pleiotropy (45). Inverse-variance weighted  
167 results, which were the best method chosen by Rucker's framework for all tested associations, are  
168 reported in the main text, but results from all methods are calculated for robustness and displayed  
169 in the supplementary information. Multiple-hypothesis testing correction was applied with the  
170 FDR method and significance established at  $FDR < 0.05$ . MR-Egger intercept tests were performed  
171 to detect horizontal pleiotropy, and single-SNP and leave-one-out analyses were used to identify  
172 outlier SNPs driving relationships (43).

173 Reverse MR for obesity traits regressed on female reproductive conditions was performed as  
174 detailed above. Genetic instruments for endometriosis (14,926 cases, 189,715 controls) (46),  
175 PCOS (10,174 cases, 103,164 controls) (47), and uterine fibroids (20,406 cases, 223,918 controls)  
176 (3) were constructed from index variants identified by the largest European ancestry GWAS for  
177 each trait. Instrument strength was assessed by F-statistics (endometriosis, 16 SNPs,  $F = 5.13$ ;  
178 PCOS, 14 SNPs,  $F = 41.6$ ; UF, 29 SNPs,  $F = 11.1$ ). Associations of genetic instruments for these  
179 reproductive conditions with BMI, WHR, and WHRadjBMI were obtained from female-specific  
180 summary statistics from the above-mentioned GIANT-UKBB meta-analysis (31).

### 181 **Non-linear Mendelian Randomisation**

182 For non-linear MR analyses, we selected female UK Biobank participants of white British ancestry  
183 with no second-degree or closer relatives in the study, as identified by the UKBB team (48), to  
184 avoid violation of the MR assumption of random assignment of genetic variants; 207,705 women  
185 were retained following this selection. Genetic instruments for BMI were constructed for each  
186 individual using female-specific index variants from Pulit et al.'s GIANT-UKBB meta-analysis  
187 (31). The instruments for BMI explained 4.15% of trait variance after adjustment for age, age-  
188 squared, smoking status, assessment centre, genotyping array, and the first ten genetic principal  
189 components to account for population stratification. Binomial non-linear MR, a method designed  
190 to assess causal relationships in different exposure strata while avoiding collider bias, was  
191 performed using the fractional polynomial method with 100 quantiles and the piecewise linear  
192 method with 10 quantiles (49). Outcomes were restricted to female reproductive disorders with  
193 prevalence  $> 5\%$  in UKBB (i.e. HMB, miscarriage, and UF) to maintain sufficient sample sizes in  
194 each quantile to estimate localised average causal effects. We assessed non-linearity with the  
195 fractional polynomial non-linearity and Cochran's Q tests, and tested heterogeneity of the

196 instrumental variable (IV) with the Cochran's Q and trend tests. All analyses were performed with  
197 the nlmr v2.0 R package (49).

## 198 **MR with Mediation Analysis**

199 To investigate the extent to which obesity affects female reproductive disorders via hormone-  
200 related mediators, two-step MR by the product of coefficients method was performed using GWAS  
201 summary statistics. This method was chosen as female reproductive disease phenotypes are binary  
202 outcomes with disease prevalence < 10% in UK Biobank, for which two-step MR provides the  
203 least biased estimates of mediation (50). Summary statistics for leptin (N = 33,987) (51), fasting  
204 insulin (N = 51,750) (52), and insulin sensitivity (N = 16,753) (53) were obtained from publicly  
205 available European-ancestry GWAS sources that do not include samples from UKBB to minimise  
206 bias from sample overlap (**Table B in S1 Table**).

207 In the first step of two-step MR, the mediators were regressed on obesity-related exposures using  
208 summary statistics MR methods described above. The direction of causality for all relationships  
209 was confirmed with the MR-Steiger directionality test (54) and reciprocal MR with mediator  
210 instruments and obesity-related exposures as outcomes were performed to ensure correct direction  
211 of causality. In the second step, multivariable MR (MVMR) was performed using combined  
212 genetic instruments for each obesity trait and hormone to estimate the independent effect of the  
213 mediator on each outcome after adjusting for the value of the exposure; and to estimate the  
214 independent effect of the exposure on outcome when adjusted for the value of each mediator. This  
215 was only done for traits where the total unadjusted effect of the exposure on the outcome was  
216 significant (FDR < 0.05). Odds ratios (ORs) for binary outcomes were converted to log ORs to  
217 calculate mediated effect by the product of coefficients method. The proportion of effect mediated

218 was calculated by dividing indirect effect over total effect. Standard errors were estimated with the  
219 delta method (55).

## 220 **Disease and SNP Clustering**

221 To assess similarities in the aetiological relationships of different reproductive conditions with  
222 obesity traits, we projected single SNP causal estimates for BMI, WHR and WHRadjBMI on the  
223 reproductive traits, estimated using the Wald ratio, in a two-dimensional space using UMAP. SNPs  
224 were annotated to their nearest gene with SNPsnap (56).

225 To identify the genetic instruments driving the causal obesity-reproductive trait association, and  
226 identify clusters of SNPs with distinct causal effect sizes, we clustered SNPs by the magnitude of  
227 their causal effect using mixture model clustering in the MR-Clust v0.1.0 R package (56). For each  
228 obesity trait-reproductive disease pair, the algorithm distinguishes the genetic instruments for the  
229 obesity traits that do not have an effect on the disease (“null cluster”), from those which have a  
230 similar scaled effect on the disease (the “substantial clusters”), and those that have a scaled effect  
231 that cannot be grouped with other variants (“junk cluster”).

## 232 **Code availability**

233 All scripts used in analyses are deposited at

234 [https://github.com/lindgrengroup/obesity\\_femrepr\\_MR](https://github.com/lindgrengroup/obesity_femrepr_MR)

235

236

237

## 238 **Results**

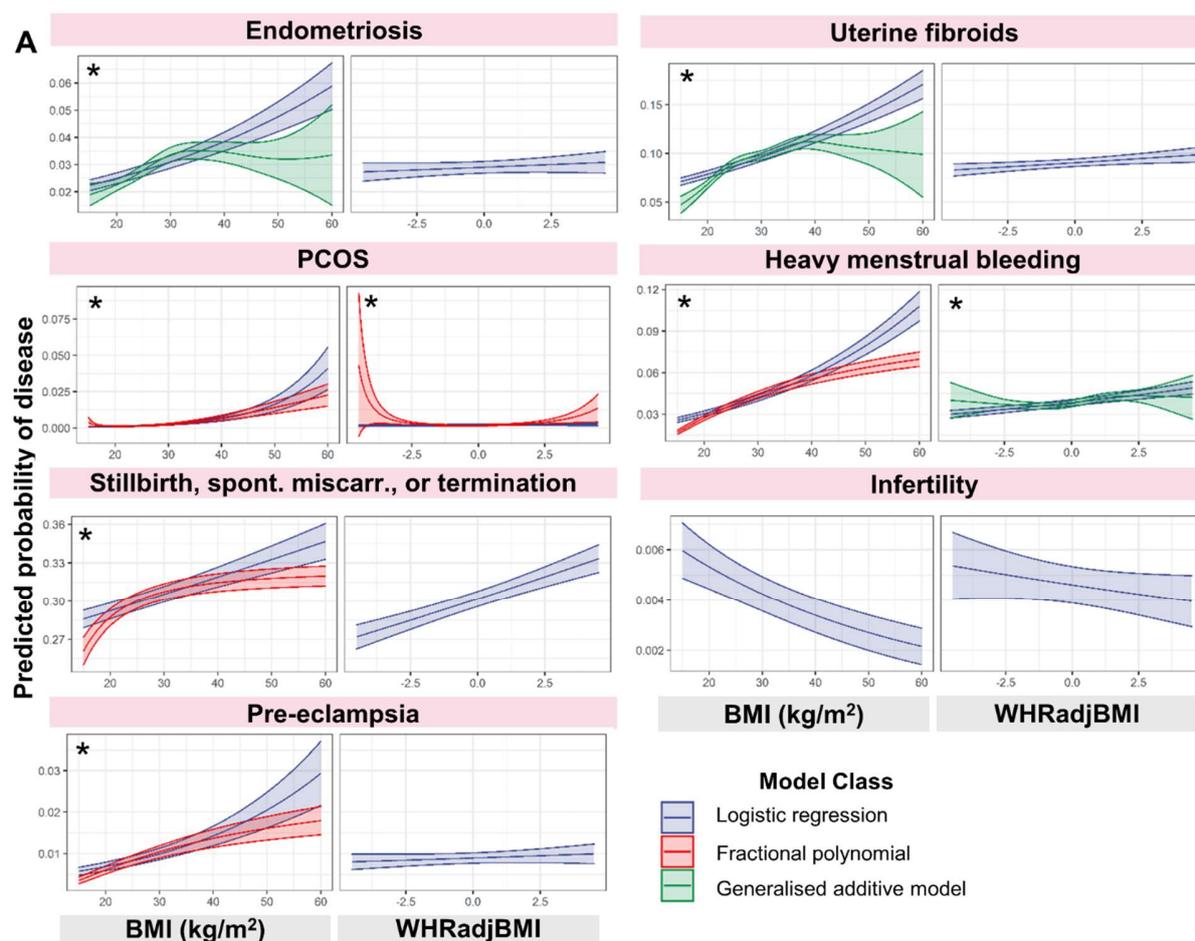
### 239 **Obesity traits are observationally associated with female reproductive diseases in UK**

#### 240 **Biobank**

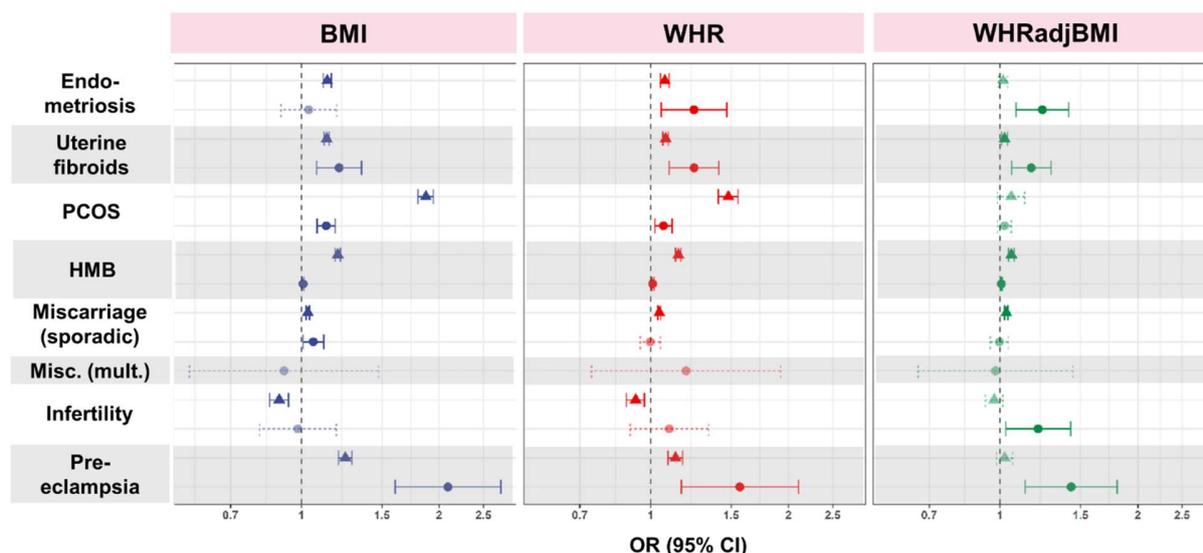
241 BMI at baseline assessment (age 40-69) was positively associated with the prevalence of most  
242 female reproductive disorders in UKBB, with the strongest association observed between BMI and  
243 pre-eclampsia (odds ratio (OR) per 1 S.D. higher BMI = 1.87,  $P = 1.90E-64$ ). Associations with  
244 WHRadjBMI were null or lower than those for WHR (ORs for WHRadjBMI v. WHR for PCOS  
245 = 1.06 v. 1.48, pre-eclampsia = 1.02 v. 1.13, endometriosis = 1.02 v. 1.08, HMB = 1.06 v. 1.14,  
246 and UF = 1.02 v. 1.08), indicating that BMI may be driving many of the associations between  
247 WHR and female reproductive diseases (**Figures 1 & 2, Table 1**). Infertility was the only disorder  
248 for which BMI (OR = 0.894,  $P = 2.16E-07$ ) and WHR (OR = 0.927,  $P = 4.08E-04$ ) were inversely  
249 associated with disease.

250 Non-linear models explained the associations of BMI with many reproductive disorders better than  
251 linear models. We observed inverted-U and plateau relationships with endometriosis (linear AIC  
252 = 67091, generalised additive model (GAM) AIC = 67051), uterine fibroids (linear AIC = 134160,  
253 GAM AIC = 134094), HMB (linear AIC = 116687, GAM AIC = 116636), miscarriage (linear AIC  
254 = 314828, GAM AIC = 314819), and pre-eclampsia (linear AIC = 24826, GAM AIC = 24814)  
255 (**Figure 1, Table D in S1 Table**). All three obesity traits displayed U-shaped relationships with  
256 PCOS.

257 Observational estimates between all obesity traits and female reproductive disorders did not differ  
 258 with or without adjustment for smoking status (Table E in S1 Table). Statistical significance after  
 259 multiple-testing correction was established at  $FDR < 0.05$ , unadjusted  $P < 0.04$ .



**Figure 1: (A) Predicted probability of developing female reproductive disorders as a function of obesity-related traits.** A series of logistic regression, fractional polynomial, and generalized additive models were fitted to estimate the probability of developing female reproductive disorders as a function of obesity-related traits in UKBB. Predicted fits for logistic and best-fitting non-linear models where better than the logistic regression (as evaluated with AIC) and 95% confidence intervals about the mean are displayed. \* indicates that non-linear models fit the data better than linear models. Spont. miscar. = spontaneous miscarriage.



**Figure 2: Comparison of observational and causal relationships between obesity-related traits and female reproductive disorders.** Odds ratios (ORs) and 95% confidence intervals per 1 S.D. higher obesity trait displayed. Significant relationships (FDR adjusted p-value < 0.05) are in solid lines while non-significant (n.s.) ones are shown with dotted lines. For observational results, BMI, WHR, or WHRadjBMI adjusted for age, age-squared, region (assessment centre), and smoking status are used as predictors in a logistic regression model. Causal relationships between genetically predicted obesity-related traits and female reproductive disorders are assessed by two-sample Mendelian randomisation. The displayed method (inverse-variance weighted) is determined via Rucker's model selection framework to minimise heterogeneity of the estimate. HMB = heavy menstrual bleeding, Misc. (mult.) = multiple consecutive miscarriage. Miscarriage (sporadic) = self-reported stillbirth, spontaneous miscarriage, or termination for observational results and sporadic miscarriage for MR results.

**Significance**  
 —●— FDR < 0.05  
 -●- n.s.  
**Association**  
 ▲ observational  
 ● causal (MR)

**Table 1: Observational and genetic associations between obesity traits and female reproductive disorders.**

Diagnosis	Obesity trait	Logistic regression		Mendelian randomisation		
		OR (95% CI)	Unadj. p-value	# SNPs	OR (95% CI) per 1 S.D. higher obesity trait	Unadj. p-value
Endometriosis	BMI	1.14 (1.12 - 1.16)	4.45E-29	264	1.04 (0.902 - 1.19)	0.606
	WHR	1.07 (1.05 - 1.10)	1.00E-09	190	1.24 (1.05 - 1.47)	1.00E-02
	WHRadjBMI	1.02 (0.99 - 1.04)	0.188	250	1.24 (1.08 - 1.41)	1.67E-03
Heavy menstrual bleeding	BMI	1.20 (1.18 - 1.22)	7.78E-117	268	1.01 (1.004 - 1.013)	3.62E-04
	WHR	1.15 (1.13 - 1.16)	8.35E-69	191	1.01 (1.005 - 1.02)	1.42E-04
	WHRadjBMI	1.06 (1.04 - 1.07)	5.30E-13	251	1.01 (1.002 - 1.011)	5.20E-03
Infertility	BMI	0.894 (0.852 - 0.936)	2.16E-07	267	0.982 (0.810 - 1.19)	0.856
	WHR	0.927 (0.884 - 0.969)	4.08E-04	191	1.10 (0.901 - 1.34)	0.355
	WHRadjBMI	0.971 (0.929 - 1.01)	0.176	251	1.21 (1.03 - 1.43)	0.0214
Miscarriage	BMI	1.03 (1.02 - 1.04)	4.28E-14	265	1.06 (1.01 - 1.12)	0.0238

<b>(sporadic)</b>	WHR	1.04 (1.04 - 1.05)	3.82E-24	190	0.998 (0.947 - 1.05)	0.933
	WHRadjBMI	1.03 (1.02 - 1.04)	1.87E-13	250	0.996 (0.953 - 1.04)	0.878
<b>Miscarriage (multiple consecutive)</b>	BMI			254	0.917 (0.570 - 1.48)	0.720
	WHR			184	1.20 (0.743 - 1.92)	0.462
	WHRadjBMI			240	0.978 (0.662 - 1.44)	0.911
<b>PCOS</b>	BMI	1.87 (1.80 - 1.94)	1.90E-64	268	1.13 (1.08 - 1.19)	7.60E-08
	WHR	1.48 (1.41 - 1.55)	3.32E-26	191	1.07 (1.02 - 1.11)	4.30E-03
	WHRadjBMI	1.06 (0.986 - 1.13)	0.124	251	1.02 (0.990 - 1.06)	0.222
<b>Pre-eclampsia</b>	BMI	1.25 (1.21 - 1.29)	3.85E-25	266	2.09 (1.60 - 2.73)	5.16E-08
	WHR	1.13 (1.09 - 1.17)	4.97E-09	191	1.57 (1.16 - 2.10)	2.92E-03
	WHRadjBMI	1.02 (0.982 - 1.07)	0.272	250	1.43 (1.13 - 1.80)	2.46E-03
<b>Uterine fibroids</b>	BMI	1.14 (1.12 - 1.15)	2.43E-63	268	1.21 (1.08 - 1.35)	9.93E-04
	WHR	1.08 (1.06 - 1.09)	2.75E-23	191	1.24 (1.10 - 1.41)	6.20E-04
	WHRadjBMI	1.02 (1.01 - 1.04)	2.94E-03	251	1.17 (1.06 - 1.29)	1.95E-03

## 260 **Body fat distribution is causally related to risk of female reproductive diseases**

261 Two-sample MR indicated that higher genetically predicted WHR and/or WHRadjBMI are causal  
 262 for higher risk of pre-eclampsia (OR per 1 S.D. higher WHR = 1.57,  $P = 2.92E-03$ ; WHRadjBMI  
 263 = 1.43,  $P = 2.46E-03$ ), endometriosis (WHR = 1.24,  $P = 1.00E-02$ ; WHRadjBMI = 1.24,  $P =$   
 264 1.67E-03), uterine fibroids (WHR = 1.24,  $P = 6.20E-04$ ; WHRadjBMI = 1.17,  $P = 1.95E-03$ ),  
 265 infertility (WHRadjBMI = 1.21,  $P = 2.14E-02$ ), and PCOS (WHR = 1.07,  $P = 4.30E-03$ ) (**Table**  
 266 **1, Figure 2**). The causal estimates of WHR and WHRadjBMI on most reproductive disorders were  
 267 higher than their observational counterparts. While genetically predicted BMI also increased risk  
 268 of most female reproductive disorders (ORs per 1 S.D. higher BMI = 1.01 for HMB to 2.09 for  
 269 pre-eclampsia), MR estimates of associations between BMI and HMB (OR = 1.01,  $P = 3.62E-04$ ),  
 270 endometriosis (OR = 1.04,  $P = 0.606$ ), and PCOS (OR = 1.13,  $P = 7.60E-08$ ) were much attenuated  
 271 compared to observational results.

272 Genetically predicted visceral adipose tissue (VAT) mass was causal for the development of pre-  
 273 eclampsia (OR per 1 kg increase in predicted VAT mass = 3.08,  $P = 6.65E-07$ ), PCOS (OR = 1.15,

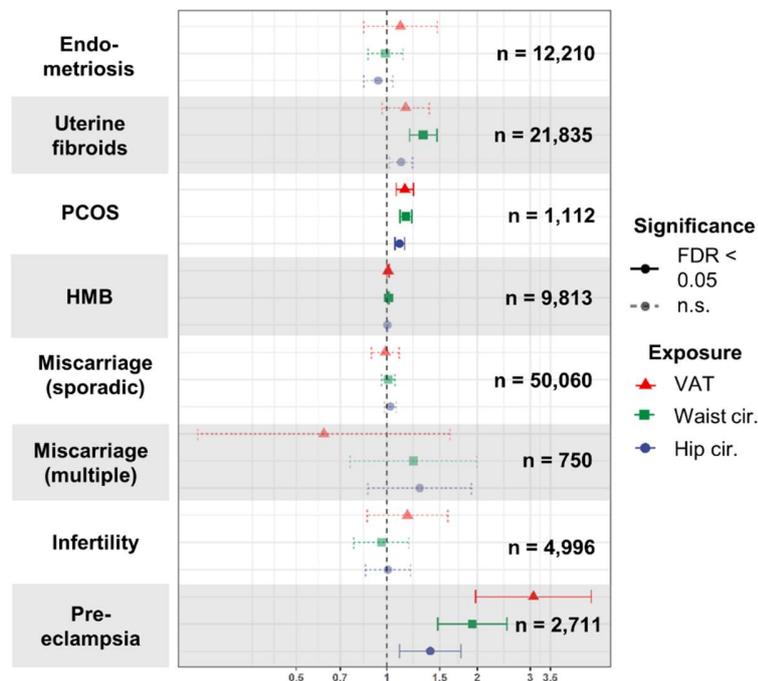
274  $P = 3.24E-05$ ), and HMB (OR = 1.01,  $P = 0.0125$ ) (**Figure 3** and **Table F in S1 Table**). The  
275 differential effect of body fat distribution on female reproductive traits was further reflected in the  
276 heterogeneous causal effects of waist circumference (WC) and hip circumference (HC) on disease  
277 development. Increased WC posed a higher risk than did increased HC for pre-eclampsia (ORs per  
278 1 S.D. higher WC = 1.93 v. HC = 1.40, heterogeneity  $P\text{-het} = 0.0373$ ), uterine fibroids (WC = 1.32  
279 v. HC = 1.12,  $P\text{-het} = 7.70E-03$ ), and PCOS (WC = 1.16 v. HC = 1.10,  $P\text{-het} = 0.0325$ ). We did  
280 not see this heterogeneity in observational associations (all  $P\text{-het} > 0.164$ ) (**Figure 3** and **Table F**  
281 **in S1 Table**).

282 No significant causal effects were found when restricting MR analyses to genetic instruments with  
283 a specific effect of waist but not hip circumference, or on hip but not waist circumference (**Table**  
284 **H in S1 Table** and **S1 Figure**), but the power based on these instruments to detect odds ratios  
285 more extreme than 1.1 was limited to 5% - 20% (**Table I in S1 Table**). No non-linear MR models  
286 explained the causal effects of BMI on any reproductive disorder better than linear MR models  
287 (**S3 Figure**). However, the power to detect non-linear effects was severely limited by the lower  
288 number of cases in each quantile of the BMI distribution in which analyses were run. Statistical  
289 significance after multiple-testing correction was established at  $FDR < 0.05$ , unadjusted  $P < 0.03$ .

290 SNPs identified in female-only GWAS and with female-specific weights for BMI, WHR, and  
291 WHRadjBMI (29) were found to be the strongest instruments, with F-statistics  $> 60$ ; instrument  
292 strength for waist- and hip-circumference was  $> 45$  (**Table B in S1 Table**). We found MR  
293 estimates to be consistent between the different MR methods (heterogeneity  $P > 0.321$ ), when  
294 based only on FinnGen summary statistics (heterogeneity  $P > 0.163$ ), or with combined-sex  
295 instruments (heterogeneity  $P > 0.999$ ), suggesting that the findings were not dependent on the  
296 adopted MR method, or substantially biased due to sample overlap between exposure and outcome

297 GWAS sources or ascertainment bias in UKBB (35, 36) (**Tables F, J, K in S1 Table & S1 & S2**  
298 **Figure**).

299 We did not find evidence for reverse causal effects of endometriosis, PCOS, or uterine fibroids on  
300 BMI, WHR, and WHRadjBMI (**Table L in S1 Table**). However, these estimates may be biased  
301 by weak genetic instruments for endometriosis (F-statistic = 5.13) and UF (F-statistic = 11.1), and  
302 high heterogeneity for all associations (Cochran's  $Q P < 4.71E-06$ ). We were limited in assessing  
303 reverse causality of other female reproductive conditions on obesity traits by the lack of large-  
304 scale publicly available GWAS summary statistics.



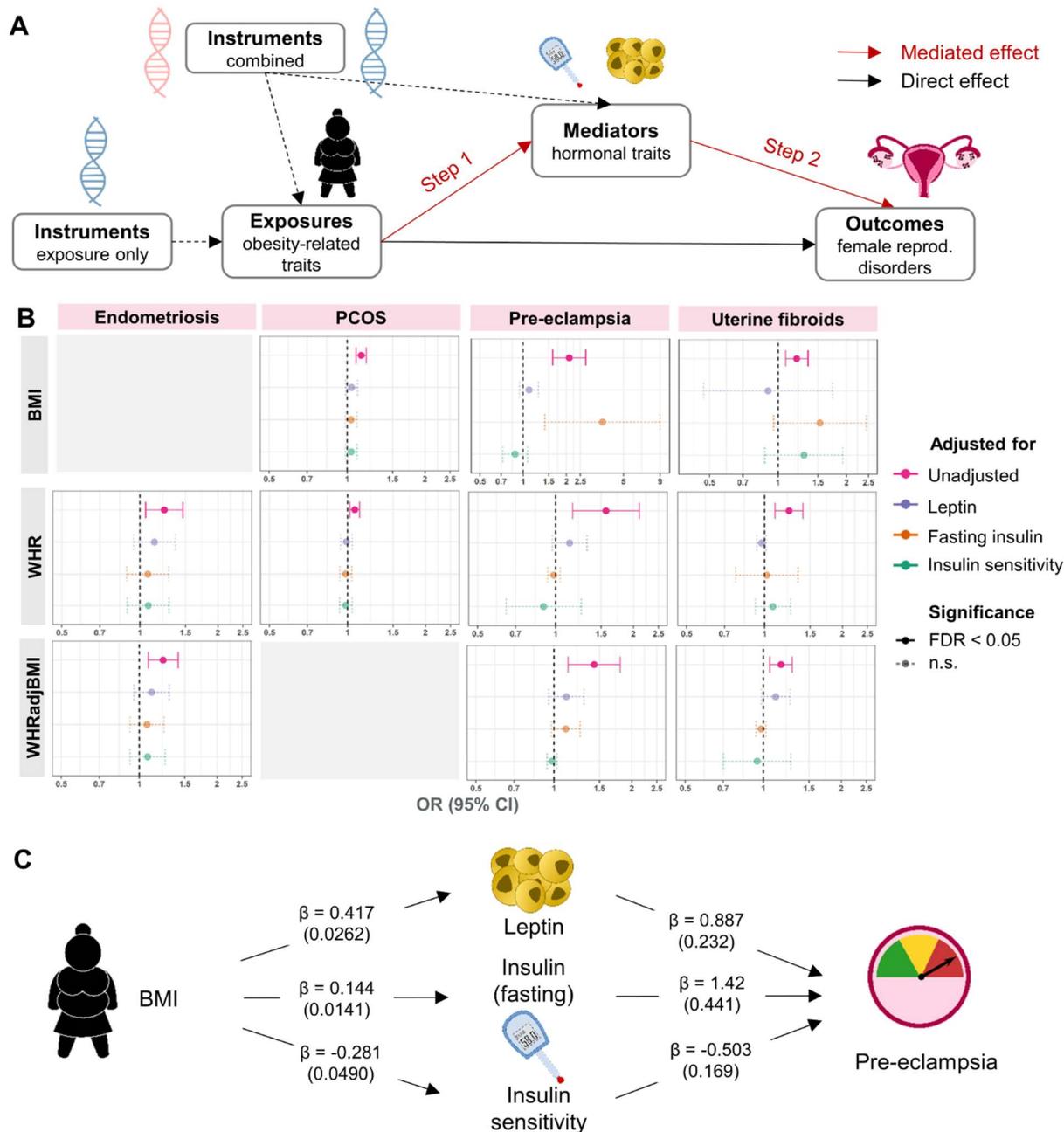
**Figure 3: Causal relationships between genetically predicted visceral adipose tissue mass, waist circumference, hip circumference and female reproductive disorders.** Odds ratios (ORs) per 1 kg increase in predicted visceral adipose tissue (VAT) mass or per 1 S.D. increase in waist circumference or hip circumference and 95% confidence intervals are estimated by two-sample MR. The displayed method (inverse-variance weighted) is determined via Rucker's model selection framework to minimise heterogeneity of the estimate. Number of cases for each disease is indicated. Significant relationships (FDR adjusted p-value < 0.05) are in solid lines while non-significant ones are shown with dotted lines. Cir. = circumference, VAT = predicted visceral adipose tissue mass

### 305 **Leptin and insulin mediate the causal effects of obesity on female reproductive disorders**

306 We applied a series of MR-based mediation analyses (58, 59) to study the role of hormonal factors  
 307 - leptin and insulin resistance - in mediating the causal relationships between obesity and female  
 308 reproductive health (**Figure 4A**). The effects of BMI, WHR, and WHRadjBMI on endometriosis,  
 309 PCOS, pre-eclampsia, and UF were attenuated (95% CIs of ORs all contain 1) when adjusted for  
 310 leptin, fasting insulin, or insulin sensitivity as measured by the modified Stumvoll Insulin

311 Sensitivity Index (ISI) (**Figure 4B, Table M in S1 Table**). Furthermore, these hormones influence  
312 risk of pre-eclampsia independently of obesity. After adjustment for BMI, leptin ( $\beta = 0.887$ ,  $P =$   
313  $1.28E-04$ ), fasting insulin ( $\beta = 1.42$ ,  $P = 1.29E-03$ ), and ISI ( $\beta = -0.503$ ,  $P = 2.99E-03$ ) were all  
314 associated with risk of pre-eclampsia (**Figure 4C, Table 2**). Similarly, fasting insulin ( $\beta = 1.27$ ,  $P$   
315  $= 7.15E-03$ ) and ISI ( $\beta = -0.793$ ,  $P = 1.46E-05$ ) had causal effects on pre-eclampsia upon  
316 adjustment for WHR. Leptin, fasting insulin, and ISI did not have significant causal effects on  
317 endometriosis, PCOS, or UF after adjustment for obesity traits. Statistical significance after  
318 multiple-testing correction was established at  $FDR < 0.05$ , unadjusted  $P < 0.01$ .

319 We calculated the proportion of total obesity effect mediated by the above hormones for disorders  
320 where the effects of obesity traits and mediators were significant at unadjusted  $P < 0.05$ . We found  
321 that leptin (50.2% of effect of BMI on pre-eclampsia), fasting insulin (between 27.7% - 36.6% of  
322 different effects), and ISI (between 19.1% - 50.1% of different effects) each mediated the total  
323 causal effect of obesity traits on female reproductive disorders (**Table 3**).



**Figure 4: (A) Method outline for two-step MR to estimate hormonally mediated effects between obesity and female reproductive disorders.** In step 1, the effect of exposures on mediators is estimated using instruments for the exposure alone, while in step 2 the independent effect of the mediators on outcomes is estimated using multivariable MR (MVMR) adjusted for exposures. All SNP-phenotype effect estimates come from different GWAS sources. **(B) MVMR estimates for obesity traits on female reproductive disorders adjusted for mediators.** MVMR was performed with combined genetic instruments for each exposure-mediator combination, displayed here for relationships where unadjusted exposure-outcome effect was significant (FDR < 0.05). **(C) Example of a mediated relationship.** Shown here for BMI effect on pre-eclampsia. Effect size estimates for exposure-mediator (betas and standard errors) and for mediator-outcome (log odds ratios and standard errors) are shown.

**Table 2:** Multivariable MR estimates of female reproductive disorders regressed on metabolic hormones, adjusted for obesity traits.

Outcome	Exposure	Adjusted for	$\beta$ +/- S.E. per 1 S.D. higher exposure	Unadjusted P-value
Endometriosis	Leptin	Unadjusted	-0.0745 +/- 0.188	0.692
		BMI	-0.0457 +/- 0.102	0.654
		WHR	-0.131 +/- 0.126	0.297
		WHRadjBMI	-0.0964 +/- 0.128	0.453
Uterine fibroids	Leptin	Unadjusted	-0.132 +/- 0.323	0.683
		BMI	0.0881 +/- 0.102	0.389
		WHR	-4.12E-03 +/- 0.138	0.976
		WHRadjBMI	-0.0834 +/- 0.14	0.550
PCOS	Leptin	Unadjusted	-0.134 +/- 0.0701	0.0567
		BMI	0.0491 +/- 0.0409	0.231
		WHR	8.44E-03 +/- 0.0476	0.859
		WHRadjBMI	-0.0310 +/- 0.0473	0.511
Pre-eclampsia	Leptin	Unadjusted	0.181 +/- 0.501	0.718
		BMI	0.887 +/- 0.232	1.28E-04
		WHR	0.468 +/- 0.309	0.129
		WHRadjBMI	-0.102 +/- 0.333	0.760
Endometriosis	Fasting insulin	Unadjusted	0.150 +/- 0.247	0.544
		BMI	0.0996 +/- 0.209	0.634
		WHR	0.186 +/- 0.254	0.465
		WHRadjBMI	0.423 +/- 0.239	0.0770
Uterine fibroids	Fasting insulin	Unadjusted	0.0263 +/- 0.342	0.939
		BMI	0.311 +/- 0.183	0.090
		WHR	0.313 +/- 0.209	0.135
		WHRadjBMI	0.262 +/- 0.202	0.195
PCOS	Fasting insulin	Unadjusted	-0.0334 +/- 0.158	0.833
		BMI	0.0283 +/- 0.0772	0.714
		WHR	0.0932 +/- 0.0788	0.237
		WHRadjBMI	0.0289 +/- 0.0726	0.690
Pre-eclampsia	Fasting insulin	Unadjusted	-0.0379 +/- 0.873	0.965
		BMI	1.42 +/- 0.441	1.29E-03
		WHR	1.27 +/- 0.473	7.15E-03
		WHRadjBMI	0.476 +/- 0.467	0.308
Endometriosis	Insulin sensitivity	Unadjusted	-0.0940 +/- 0.133	0.478
		BMI	-0.101 +/- 0.0864	0.244
		WHR	-0.131 +/- 0.100	0.192
		WHRadjBMI	-0.102 +/- 0.106	0.336
Uterine fibroids	Insulin sensitivity	Unadjusted	-0.223 +/- 0.137	0.103
		BMI	-0.0951 +/- 0.0716	0.184
		WHR	-0.202 +/- 0.0830	0.0151

		WHRadjBMI	-0.120 +/- 0.0855	0.165
PCOS	Insulin sensitivity	Unadjusted	-0.0239 +/- 0.0616	0.698
		BMI	-0.0257 +/- 0.0290	0.376
		WHR	-0.0355 +/- 0.0293	0.226
		WHRadjBMI	-0.0443 +/- 0.0294	0.132
Pre-eclampsia	Insulin sensitivity	Unadjusted	-0.354 +/- 0.295	0.229
		BMI	-0.503 +/- 0.169	2.99E-03
		WHR	-0.793 +/- 0.183	1.461E-05
		WHRadjBMI	-0.519 +/- 0.205	0.0110

**Table 3:** Proportion of effect mediated for exposure - mediator - outcome relationships.

Exposure	Mediator	Outcome	Log OR (S.E.) per 1 S.D. higher exposure <i>Unadj. pval</i>			Proportion of Effect Mediated (95% CI)
			Exposure - Outcome	Mediator - Outcome	Exposure - Mediator	
BMI	Leptin	Pre-eclampsia	0.737 (0.135) <i>P = 5.16E-08</i>	0.887 (0.232) <i>P = 1.28E-04</i>	0.417 (0.0262) <i>P = 7.94E-57</i>	50.2% (18.2% - 82.2%)
	Fasting insulin			1.42 (0.441) <i>P = 1.29E-03</i>	0.144 (0.0141) <i>P = 1.61E-24</i>	27.7% (7.40% - 48.0%)
	Insulin sensitivity			-0.503 (0.169) <i>P = 2.99E-03</i>	-0.281 (0.0490) <i>P = 9.83E-09</i>	19.1% (3.33% - 35.0%)
WHR	Fasting insulin	Pre-eclampsia	0.449 (0.151) <i>P = 2.92E-03</i>	1.27 (0.473) <i>P = 7.15E-03</i>	0.129 (0.0159) <i>P = 5.04E-16</i>	36.6% (0% - 73.7%)
	Insulin sensitivity			-0.793 (0.183) <i>P = 1.46E-05</i>	-0.283 (0.0601) <i>P = 2.44E-06</i>	50.1% (4.98% - 95.3%)
	Insulin sensitivity	Uterine fibroids	0.218 (0.0637) <i>P = 6.20E-04</i>	-0.202 (0.083) <i>P = 1.51E-02</i>	-0.283 (0.0601) <i>P = 2.44E-06</i>	26.2% (0% - 54.4%)
WHRadjBMI	Insulin sensitivity	Pre-eclampsia	0.358 (0.118) <i>P = 2.46E-03</i>	-0.519 (0.205) <i>P = 1.14E-02</i>	-0.168 (0.0427) <i>P = 8.21E-05</i>	24.4% (0% - 51.8%)

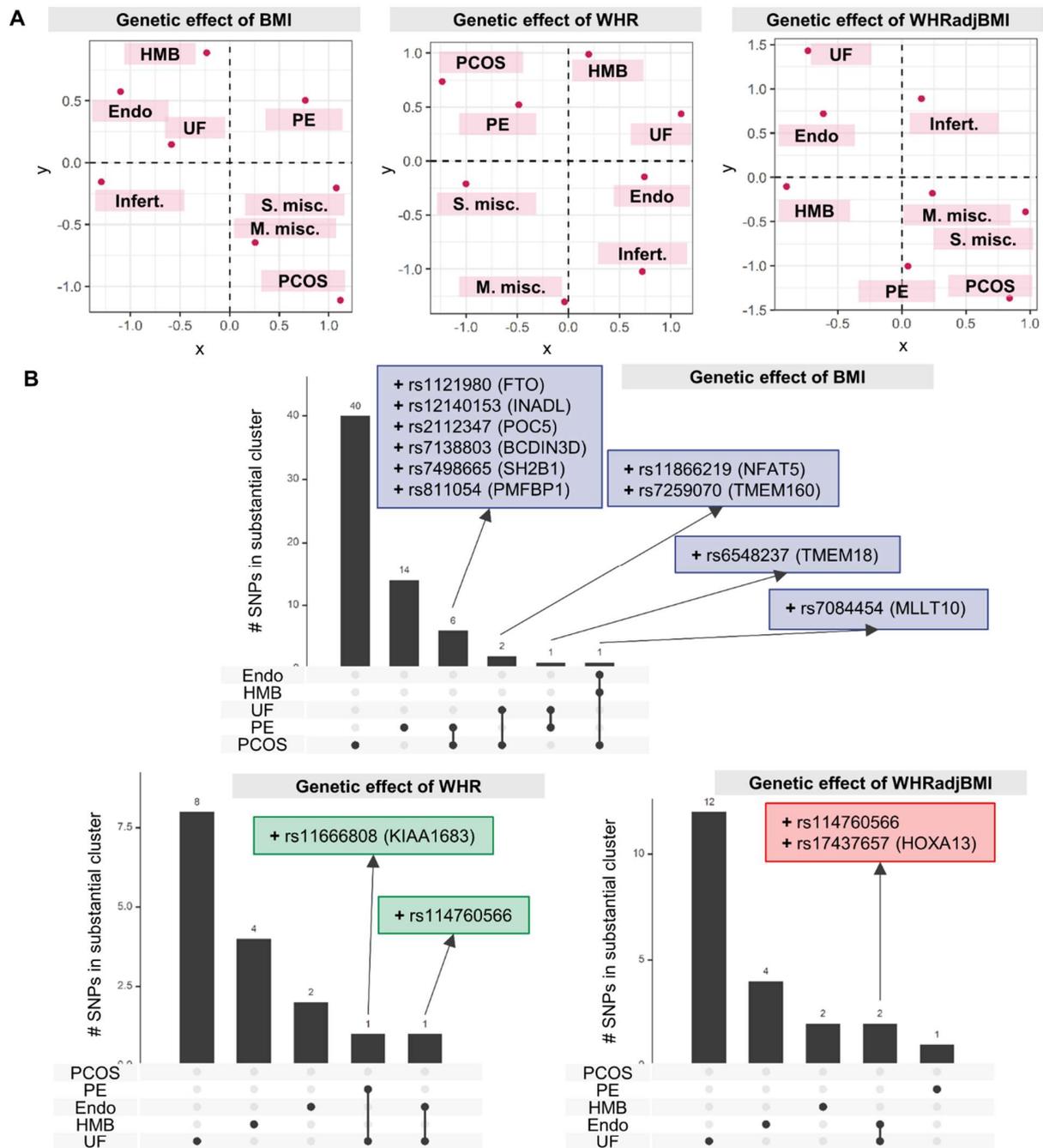
324 **Other metabolic and hormone pathways may drive the aetiological relationships of obesity**  
325 **with female reproductive diseases**

326 We assessed the similarities in the aetiological relationships of different reproductive conditions  
327 with obesity, by projecting the single SNP causal estimates for BMI, WHR and WHRadjBMI on  
328 the reproductive traits in a two-dimensional space using UMAP (**Figure 5A**). The UMAP  
329 projections based on all obesity traits clustered endometriosis and UF together with infertility and  
330 HMB, which were further separated from miscarriage (sporadic and multiple consecutive). While  
331 PCOS and pre-eclampsia were grouped closely in UMAP plots of the effect of WHR and  
332 WHRadjBMI variants, they were separated by BMI-associated variants. This reflects a shared  
333 genetic component of the aetiological role of general and central obesity in the three groups of  
334 reproductive conditions.

335 We further examined if different aspects of obesity play an aetiological role in different  
336 reproductive conditions. For each obesity trait-reproductive disease pair, we grouped the genetic  
337 instruments for the obesity traits by those that do not have an effect on the disease (“null cluster”),  
338 those which have a similar scaled effect on the disease (the “substantial clusters”), and those that  
339 have a scaled effect that cannot be grouped with other variants (“junk cluster”) using MRClust  
340 (57). One substantial cluster was identified for each pair of obesity traits and reproductive  
341 conditions. The only exception to this was with WHRadjBMI and UF, for which two substantial  
342 clusters were identified, one with positive causal effect and the other with negative effect (**S4**  
343 **Figure and Table N in S1 Table**). Of the 4 SNPs in the negative effect cluster, rs2277339  
344 (missense variant in *PRIMI* and upstream of *HSD17B6*, involved in steroid biosynthesis) is  
345 associated with primary ovarian insufficiency, early menopause, and PCOS (60, 61), and  
346 rs11694173 is intronic to *THADA*, which is also associated with PCOS (47). On the other hand, 4

347 of 10 SNPs in the positive effect cluster are associated with metabolic traits - rs12328675 and  
348 rs2459732 with circulating leptin (62), rs6905288 with type 2 diabetes and thyroid stimulating  
349 hormone (63, 64), and rs4686696 is intronic to insulin-like growth factor *IGF2BP2*.

350 SNPs with high probability of belonging to the substantial cluster ( $\geq 80\%$  probability) were  
351 generally unique to each obesity-disease relationship, with no more than 2 variants shared between  
352 any two clusters (**Figure 5B**). However, 6 BMI index SNPs had positive causal effect estimates  
353 for both PCOS and pre-eclampsia, including rs1121980 in the adipose-associated gene *FTO* and  
354 rs7498665 in *SH2B1*, linked to insulin resistance in obesity. The BMI-associated variant  
355 rs7084454 (intronic to *MLLT10*) was shared by substantial clusters for PCOS, endometriosis, and  
356 UF, while rs114760566 (mapped to *HMGAI*, associated with type 2 diabetes and multiple  
357 lipomatosis) was shared by endometriosis and UF. We evaluated the biological effect of the top  
358 SNPs in each substantial cluster with the DEPICT algorithms for pathway enrichment and gene  
359 prioritisation. We recapitulated the known associations of GEMIN5 subnetwork enrichment in  
360 SNPs causal for BMI-PCOS, which has previously been implicated in the aetiology of PCOS (65).  
361 Gene prioritisation for WHR-endometriosis causal SNPs highlighted *TBX15*, an important  
362 mesodermal transcription factor with roles in endometrial and ovarian cancer (66, 67).



**Figure 5: (A) Female reproductive disorders separated by the genetic effects of different obesity traits on reproductive conditions.** Uniform Manifold Approximation and Projection (UMAP) was used to plot the separation of diseases based on single-SNP genetic effects of obesity instruments, evaluated by two-sample MR. **(B) Number of SNPs in substantial clusters for each obesity x female reproductive condition relationships.** SNPs with  $\geq 80\%$  probability of belonging to a substantial cluster (MRClust) are displayed with their nearest gene as annotated by SNPSnap and “+” or “-” representing causal effect direction. Infert = infertility, endo = endometriosis, UF = uterine fibroids, PE = pre-eclampsia, S. misc. = sporadic miscarriage, M. misc. = multiple consecutive miscarriage.

## 363 Discussion

364 In this first systematic genetics-based causal investigation of the aetiological role of obesity in  
365 female reproductive health, we report evidence that common indices of obesity increase risk of a  
366 broad range of reproductive conditions whose effects may be non-uniform across the obesity  
367 spectrum. The strongest effect of generalised obesity was found for pre-eclampsia, while more  
368 modest effects were observed for nearly all other studied conditions. We identified endocrine  
369 mechanisms, including those related to leptin and insulin resistance, as potential drivers of  
370 aetiological relationships of both generalised and central obesity with female reproductive health.  
371 Finally, we found genetic evidence that certain groups of reproductive conditions, such as UF and  
372 endometriosis, may share a mechanistically similar relationship with obesity.

373 Our findings highlight that the relationships between obesity and female reproductive disorders  
374 are (i) non-uniform in their nature and strength, and (ii) observationally non-linear across the  
375 obesity spectrum. We report substantial differences in the causal effect estimates of BMI on  
376 reproductive diseases, with each S.D. increase in BMI doubling the risk of pre-eclampsia, but more  
377 moderately (ORs = 1.01 - 1.25 for PCOS, miscarriage, UF, HMB) or not at all (infertility,  
378 endometriosis) affecting other conditions. Conversely, central fat distribution independent of BMI  
379 showed substantial genetically predicted effects on both infertility and endometriosis (ORs per 1  
380 S.D. increase in WHRadjBMI = 1.21 - 1.46) as well as on pre-eclampsia and UF (ORs = 1.17 -  
381 1.43), but not on PCOS, HMB, and miscarriage. These findings highlight that the aetiological role  
382 of obesity in female reproductive diseases is heterogeneous in its effect strength, and may be driven  
383 by overall adiposity (PCOS, HMB, and miscarriage), isolated central obesity (infertility and  
384 endometriosis), or by both generalised and central obesity (pre-eclampsia and UF).

385 For several reproductive conditions, we found substantial differences between the observational  
386 and genetically predicted causal effect estimates, which may indicate a bi-directional relationship  
387 between obesity and reproductive health. For instance, while the observational analyses suggested  
388 an 87% increase in PCOS risk per S.D. higher BMI, the MR analyses indicated that each S.D.  
389 higher BMI causally increases PCOS risk by only 13%. Similarly, 1 S.D. higher WHR and  
390 WHRadjBMI causally increase endometriosis risk by 24%, while the observational analyses  
391 suggest a more modest increase in risk of 7% and 2% respectively. This discrepancy may in part  
392 be due to reverse causality, which we were not powered to detect in this study, as the number and  
393 strength of the available genetic instruments for reproductive conditions is substantially lower than  
394 those for BMI and WHR. The obesity traits upon which the observational analyses were based  
395 were measured at ages 44-67, which was for most conditions likely to be several years or decades  
396 after women developed the condition, and often post-menopause. While our observational analyses  
397 adjusted for the effect of age on obesity traits, adjusting for menopause status proved to be  
398 unreliable as up to 42% of women with reproductive diseases in UKB were unsure of their  
399 menopause status, as opposed to 16% of female participants without a diagnosis for any of the  
400 studied conditions. The observational estimates may therefore capture both the effect of obesity  
401 on disease risk as well as any downstream effects of the disease or commonly used treatments on  
402 body weight and fat distribution. For instance, the large observational effect of BMI on PCOS  
403 prevalence may reflect both a causal effect of obesity on disease risk (68), as captured by the  
404 genetically predicted effect, as well as weight gain as a consequence of PCOS (69). Other potential  
405 contributing factors to the differences between genetic and observational estimates are  
406 confounding by unmeasured variables which lead to inflated observational associations (70, 71),  
407 referral bias wherein obesity status affects the likelihood of receiving a diagnosis (72, 73), or

408 differences in pre- and post-menopausal weight and body fat distribution not captured by age (74).  
409 Finally, while the observational relationships between obesity and some female reproductive  
410 disorders were non-linear, we did not find non-linearity in the causal effects of BMI on these  
411 diseases. The non-linear MR analyses were likely under-powered to detect associations with few  
412 cases in each quantile of the BMI spectrum.

413 We noted that genetic estimates for the effect of fat distribution were not similarly attenuated when  
414 compared to BMI effects. This disparity may be due to the differing impacts of overall and  
415 abdominal (central) adiposity, as the latter is thought to be biologically more directly linked to  
416 female reproductive health than generalised obesity, via pathways including insulin resistance and  
417 hyper-androgenaemia (5, 75-77). Supporting the stronger effect of central body fat, we also  
418 reported higher causal effects of waist than hip circumference with HMB, PCOS, pre-eclampsia,  
419 and UF. Genetically predicted visceral adipose tissue mass increased risk of PCOS and pre-  
420 eclampsia, in line with observational studies (78). VAT mass is also observationally associated  
421 with uterine fibroids (76), yet we did not find a significant causal effect of genetically predicted  
422 VAT mass on development of UF, which may suggest a bi-directional or reverse causal  
423 relationship.

424 Endometriosis and infertility were the only reproductive conditions which did not show a  
425 consistently positive link with obesity. The modest observational associations of both BMI and  
426 WHR with higher endometriosis prevalence in UKBB contradict previous studies, including  
427 prospective cohort studies, which reported that lower BMI was associated with increased disease  
428 prevalence (14, 79, 80). The positive association with endometriosis may in part be due to weight  
429 gain as a consequence of the disease, for instance due to hormonal treatments (81-83), chronic pain  
430 (84), inflammation (85), or earlier onset of menopause (86). We however did not find evidence

431 that generalised obesity plays a causal role in the aetiology of endometriosis, which suggests that  
432 the observational finding reflects a reverse causal relationship. Conversely, the positive genetically  
433 predicted effect of WHRadjBMI on endometriosis risk indicates a causal role for abdominal fat  
434 distribution. For infertility, we observe a similar divergence between the observational and  
435 genetically predicted effects of obesity traits, with BMI showing a negative observational  
436 association, but WHRadjBMI a genetically predicted positive association. The causes of female  
437 infertility are multiple, ranging from PCOS (87) and anovulation (88), to tubal disease (89),  
438 endometriosis (90), low oocyte quality (91), hormonal and immunological dysfunction (92-95),  
439 and yet unknown mechanisms. Each of these may have distinct and complex relationships with  
440 obesity, which cannot be captured by studying the links with infertility of any cause. Nonlinear  
441 effects, such as the increased association of under- and overweight with incidence of infertility  
442 (12, 96), may also obscure these estimates, although our observational analyses did not provide  
443 evidence for a non-linear relationship.

444 We conducted the first genetics-based investigation of the mediating hormonal pathways  
445 underlying the causal relationships between obesity and female reproductive health. We identify  
446 mechanisms related to insulin resistance and leptin as mediators of the effects of obesity traits on  
447 UF and pre-eclampsia. The latter is consistent with hypotheses that obese women with metabolic  
448 dysregulation are at highest risk of developing hypertensive disorders of pregnancy via angiogenic  
449 and pro-inflammatory mechanisms. Increased circulating leptin may have a vasoconstrictive,  
450 hypertensive effect, which may be worsened by attenuation of insulin-induced vasorelaxation and  
451 increased levels of TNF-alpha and IL6 (7, 21).

452 Finally, genetic clustering of female reproductive conditions revealed common genetic causes of  
453 obesity on endometriosis, UF, and HMB, which are known to share mechanisms of development

454 (3, 97). The projection of infertility with these diseases merits following up on the genetic basis of  
455 endometriosis-related infertility with an eye to prevention and treatment.

456 The main strength of our work is the systematic approach to characterising the relationship  
457 between a broad range of obesity traits and common female reproductive conditions using both  
458 observational and genetic approaches. All observational associations were estimated in the same  
459 large-scale cohort study, which tends to lead to less biased estimates than case-control studies upon  
460 which most previous results were based. Moreover, we conducted the first genetics-based  
461 mediation analyses to pinpoint the mechanisms driving the causal effect of obesity on risk of  
462 reproductive diseases.

463 Reproductive conditions remain underdiagnosed and underreported in the UK, which was reflected  
464 in their low prevalence among female UKBB participants (**Table A in S1 Table**). This posed a  
465 limitation to our analyses in UK Biobank by reducing power to identify significant associations.  
466 For this reason, we opted to use broad case categories, such as infertility of any cause, as we had  
467 insufficient power and information to examine conditions by sub-types. Secondly, we restricted  
468 our analyses to women of genetically European ancestry, due to a lack of genetic data on women  
469 of other ancestries. Many of the reproductive diseases included here, with uterine fibroids being  
470 the most notable example (98, 99), are more prevalent in non-European populations and our results  
471 may not be transferable to women of other ancestries (100-102), which emphasises the urgent need  
472 to set up large-scale studies similar to UK Biobank on participants of non-European ancestry. We  
473 were further limited in investigations of metabolic, hormonal, and inflammatory mediating  
474 mechanisms by a lack of publicly available GWAS summary statistics for these traits. Finally, the  
475 lack of data on BMI and WHR prior to disease onset, and limited information on the age at which

476 reproductive conditions were first diagnosed, complicated the interpretation of our findings from  
477 observational analyses in UK Biobank.

478 Key priorities for the future are the further exploration and validation of the pathways through  
479 which obesity increases risk of female reproductive disease. Notably, our finding that insulin  
480 resistance may be an important mediating mechanism warrants further attention, as cheap and safe  
481 treatments are available to increase insulin sensitivity. This is demonstrated by the successful use  
482 of metformin as a first-line treatment in women with PCOS (103), but such a treatment strategy  
483 has not yet been explored for other reproductive conditions linked to obesity. More generally,  
484 better and more detailed diagnostic information on reproductive health in large-scale cohort studies  
485 is urgently required for future research on the causes, consequences and aetiological mechanisms  
486 of female reproductive illnesses.

487 In conclusion, we provide genetic evidence that both generalised and central obesity play an  
488 aetiological role in a broad range of female reproductive conditions, but the extent of this link  
489 differs substantially between conditions. Our findings also highlight the importance of hormonal  
490 pathways, notably leptin and insulin resistance, as mediating mechanisms and potential targets for  
491 intervention in the treatment and prevention of common female reproductive conditions.

## 492 **Acknowledgements**

493 We acknowledge the participants and investigators of FinnGen for their contribution to this study.  
494 This research has been conducted using the UK Biobank Resource under Application Number  
495 10844. The views expressed are those of the author(s) and not necessarily those of the NHS, the  
496 NIHR or the Department of Health.

## 497 **Funding**

498 The research was supported by the Wellcome Trust Core Award Grant Number 203141/Z/16/Z  
499 with additional support from the NIHR Oxford BRC.

500 **S.S.V.** is supported by the Rhodes Trust (<https://www.rhodeshouse.ox.ac.uk/>), Clarendon Fund  
501 (<http://www.ox.ac.uk/clarendon/about>), and the Medical Sciences Doctoral Training Centre  
502 (<https://www.medsci.ox.ac.uk/>) at the University of Oxford. **S.B.** is supported by the Li Ka Shing  
503 Foundation. **M.V.H.** works in a unit that receives funding from the UK Medical Research  
504 Council and is supported by a British Heart Foundation Intermediate Clinical Research  
505 Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical  
506 Research Centre. **C.M.L.** is supported by the Li Ka Shing Foundation, NIHR Oxford Biomedical  
507 Research Centre, Oxford, NIH (1P50HD104224-01), Gates Foundation (INV-024200), and a  
508 Wellcome Trust Investigator Award (221782/Z/20/Z). **L.B.L.W.** is supported by the Wellcome  
509 Trust (221651/Z/20/Z).

## 510 **Competing Interests**

511 **C.M.B.** reports grants from Bayer AG, AbbVie Inc, Volition Rx, MDNA Life Sciences, Roche  
512 Diagnostics Inc., and consultancy for Myovant. He is a member of the independent data monitoring  
513 board at ObsEva; **I.G.** reports grants from Bayer AG; **K.T.Z.** reports grants from Bayer AG,  
514 AbbVie Inc, Volition Rx, MDNA Life Sciences, Roche Diagnostics Inc, and non-financial  
515 scientific collaboration with Population Diagnostics Ltd, outside the submitted work; **M.V.H.** has  
516 consulted for Boehringer Ingelheim, and in adherence to the University of Oxford's Clinical Trial  
517 Service Unit & Epidemiological Studies Unit (CSTU) staff policy, did not accept personal

518 honoraria or other payments from pharmaceutical companies; **C.M.L.** reports grants from Bayer  
519 AG and Novo Nordisk and has a partner who works at Vertex; no other relationships or activities  
520 that could appear to have influenced the submitted work.

## 521 **References**

- 522 1. Wei S, Schmidt MD, Dwyer T, Norman RJ, Venn AJ. Obesity and menstrual irregularity:  
523 associations with SHBG, testosterone, and insulin. *Obesity (Silver Spring)*. 2009;17(5):1070-6.
- 524 2. Douchi T, Kuwahata R, Yamamoto S, Oki T, Yamasaki H, Nagata Y. Relationship of  
525 upper body obesity to menstrual disorders. *Acta Obstet Gynecol Scand*. 2002;81(2):147-50.
- 526 3. Gallagher CS, Makinen N, Harris HR, Rahmioglu N, Uimari O, Cook JP, et al. Genome-  
527 wide association and epidemiological analyses reveal common genetic origins between uterine  
528 leiomyomata and endometriosis. *Nat Commun*. 2019;10(1):4857.
- 529 4. Yi KW, Shin JH, Park MS, Kim T, Kim SH, Hur JY. Association of body mass index  
530 with severity of endometriosis in Korean women. *Int J Gynaecol Obstet*. 2009;105(1):39-42.
- 531 5. Diamanti-Kandarakis E. Role of obesity and adiposity in polycystic ovary syndrome. *Int*  
532 *J Obes (Lond)*. 2007;31 Suppl 2:S8-13; discussion S31-2.
- 533 6. Glueck CJ, Goldenberg N. Characteristics of obesity in polycystic ovary syndrome:  
534 Etiology, treatment, and genetics. *Metabolism*. 2019;92:108-20.
- 535 7. Spradley FT. Metabolic abnormalities and obesity's impact on the risk for developing  
536 preeclampsia. *Am J Physiol Regul Integr Comp Physiol*. 2017;312(1):R5-R12.
- 537 8. Lashen H, Fear K, Sturdee DW. Obesity is associated with increased risk of first  
538 trimester and recurrent miscarriage: matched case-control study. *Hum Reprod*.  
539 2004;19(7):1644-6.
- 540 9. Metwally M, Saravelos SH, Ledger WL, Li TC. Body mass index and risk of miscarriage  
541 in women with recurrent miscarriage. *Fertil Steril*. 2010;94(1):290-5.
- 542 10. van der Steeg JW, Steures P, Eijkemans MJ, Habbema JD, Hompes PG, Burggraaff JM,  
543 et al. Obesity affects spontaneous pregnancy chances in subfertile, ovulatory women. *Hum*  
544 *Reprod*. 2008;23(2):324-8.
- 545 11. Wise LA, Rothman KJ, Mikkelsen EM, Sorensen HT, Riis A, Hatch EE. An internet-  
546 based prospective study of body size and time-to-pregnancy. *Hum Reprod*. 2010;25(1):253-64.
- 547 12. Grodstein F, Goldman MB, Cramer DW. Body mass index and ovulatory infertility.  
548 *Epidemiology*. 1994;5(2):247-50.

- 549 13. Missmer SA, Hankinson SE, Spiegelman D, Barbieri RL, Marshall LM, Hunter DJ.  
550 Incidence of laparoscopically confirmed endometriosis by demographic, anthropometric, and  
551 lifestyle factors. *Am J Epidemiol.* 2004;160(8):784-96.
- 552 14. Shah DK, Correia KF, Vitonis AF, Missmer SA. Body size and endometriosis: results  
553 from 20 years of follow-up within the Nurses' Health Study II prospective cohort. *Hum Reprod.*  
554 2013;28(7):1783-92.
- 555 15. Dixon SC, Nagle CM, Thrift AP, Pharoah PD, Pearce CL, Zheng W, et al. Adult body  
556 mass index and risk of ovarian cancer by subtype: a Mendelian randomization study. *Int J*  
557 *Epidemiol.* 2016;45(3):884-95.
- 558 16. Painter JN, O'Mara TA, Marquart L, Webb PM, Attia J, Medland SE, et al. Genetic Risk  
559 Score Mendelian Randomization Shows that Obesity Measured as Body Mass Index, but not  
560 Waist:Hip Ratio, Is Causal for Endometrial Cancer. *Cancer Epidemiol Biomarkers Prev.*  
561 2016;25(11):1503-10.
- 562 17. Brower MA, Hai Y, Jones MR, Guo X, Chen YI, Rotter JI, et al. Bidirectional Mendelian  
563 randomization to explore the causal relationships between body mass index and polycystic ovary  
564 syndrome. *Hum Reprod.* 2019;34(1):127-36.
- 565 18. Mahutte NG, Matalliotakis IM, Goumenou AG, Vassiliadis S, Koumantakis GE, Arici A.  
566 Inverse correlation between peritoneal fluid leptin concentrations and the extent of  
567 endometriosis. *Hum Reprod.* 2003;18(6):1205-9.
- 568 19. Markowska A, Rucinski M, Drews K, Malendowicz LK. Further studies on leptin and  
569 leptin receptor expression in myometrium and uterine myomas. *Eur J Gynaecol Oncol.*  
570 2005;26(5):517-25.
- 571 20. Brannian JD, Schmidt SM, Kreger DO, Hansen KA. Baseline non-fasting serum leptin  
572 concentration to body mass index ratio is predictive of IVF outcomes. *Hum Reprod.*  
573 2001;16(9):1819-26.
- 574 21. Plowden TC, Zarek SM, Rafique S, Sjaarda LA, Schisterman EF, Silver RM, et al.  
575 Preconception leptin levels and pregnancy outcomes: A prospective cohort study. *Obes Sci Pract.*  
576 2020;6(2):181-8.
- 577 22. AlAshqar A, Patzkowsky K, Afrin S, Wild R, Taylor HS, Borahay MA. Cardiometabolic  
578 Risk Factors and Benign Gynecologic Disorders. *Obstet Gynecol Surv.* 2019;74(11):661-73.
- 579 23. Hauth JC, Clifton RG, Roberts JM, Myatt L, Spong CY, Leveno KJ, et al. Maternal  
580 insulin resistance and preeclampsia. *Am J Obstet Gynecol.* 2011;204(4):327 e1-6.

- 581 24. Butler MG, McGuire A, Manzardo AM. Clinically relevant known and candidate genes  
582 for obesity and their overlap with human infertility and reproduction. *J Assist Reprod Genet.*  
583 2015;32(4):495-508.
- 584 25. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open  
585 access resource for identifying the causes of a wide range of complex diseases of middle and old  
586 age. *PLoS Med.* 2015;12(3):e1001779.
- 587 26. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended  
588 alternative to Bonferroni-type adjustments in health studies. *J Clin Epidemiol.* 2014;67(8):850-7.
- 589 27. Carreras-Torres R, Johansson M, Haycock PC, Relton CL, Davey Smith G, Brennan P, et  
590 al. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. *BMJ.*  
591 2018;361:k1767.
- 592 28. Ambler GB, Axel. *mfp: Multivariable Fractional Polynomials*. In: Luecke S, editor. 1.5.2  
593 ed: CRAN; 2015.
- 594 29. Wood S. *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness*  
595 Estimation. In: Wood S, editor. 1.8-31 ed: CRAN; 2021.
- 596 30. H. A. Information Theory and an Extension of the Maximum Likelihood Principle. In:  
597 Parzen E, TK, Kitagawa G., editor. *Selected Papers of Hirotugu Akaike*. Springer Series in  
598 Statistics (Perspectives in Statistics). New York, NY: Springer; 1998.
- 599 31. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-  
600 analysis of genome-wide association studies for body fat distribution in 694 649 individuals of  
601 European ancestry. *Hum Mol Genet.* 2019;28(1):166-74.
- 602 32. Karlsson T, Rask-Andersen M, Pan G, Hoglund J, Wadelius C, Ek WE, et al.  
603 Contribution of genetics to visceral adiposity and its relation to cardiovascular and metabolic  
604 disease. *Nat Med.* 2019;25(9):1390-5.
- 605 33. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, et al. The MRC IEU  
606 OpenGWAS data infrastructure. *bioRxiv.* 2020:2020.08.10.244293.
- 607 34. Lotta LA, Wittemans LBL, Zuber V, Stewart ID, Sharp SJ, Luan J, et al. Association of  
608 Genetic Variants Related to Gluteofemoral vs Abdominal Fat Distribution With Type 2 Diabetes,  
609 Coronary Disease, and Cardiovascular Risk Factors. *JAMA.* 2018;320(24):2553-63.
- 610 35. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when  
611 selection bias can substantially influence observed associations. *Int J Epidemiol.*  
612 2018;47(1):226-35.

- 613 36. Pirastu N, Cordioli M, Nandakumar P, Mignogna G, Abdellaoui A, Hollis B, et al.  
614 Genetic analyses identify widespread sex-differential participation bias. *bioRxiv*.  
615 2021:2020.03.22.001453.
- 616 37. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, et al. Efficiently  
617 controlling for case-control imbalance and sample relatedness in large-scale genetic association  
618 studies. *Nat Genet*. 2018;50(9):1335-41.
- 619 38. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide  
620 association scans. *Bioinformatics*. 2010;26(17):2190-1.
- 621 39. Painter JN, Anderson CA, Nyholt DR, Macgregor S, Lin J, Lee SH, et al. Genome-wide  
622 association study identifies a locus at 7p15.2 associated with endometriosis. *Nat Genet*.  
623 2011;43(1):51-4.
- 624 40. Laisk T, Soares ALG, Ferreira T, Painter JN, Censin JC, Laber S, et al. The genetic  
625 architecture of sporadic and multiple consecutive miscarriage. *Nat Commun*. 2020;11(1):5980.
- 626 41. Deng L, Zhang H, Yu K. Power calculation for the general two-sample Mendelian  
627 randomization analysis. *Genet Epidemiol*. 2020;44(3):290-9.
- 628 42. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian  
629 randomization studies. *Int J Epidemiol*. 2013;42(5):1497-501.
- 630 43. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base  
631 platform supports systematic causal inference across the human phenome. *Elife*. 2018;7.
- 632 44. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A  
633 framework for the investigation of pleiotropy in two-sample summary data Mendelian  
634 randomization. *Stat Med*. 2017;36(11):1783-802.
- 635 45. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, et al.  
636 Improving the visualization, interpretation and analysis of two-sample summary data Mendelian  
637 randomization via the Radial plot and Radial regression. *Int J Epidemiol*. 2018;47(4):1264-78.
- 638 46. Sapkota Y, Steinthorsdottir V, Morris AP, Fassbender A, Rahmioglu N, De Vivo I, et al.  
639 Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes  
640 involved in hormone metabolism. *Nat Commun*. 2017;8:15539.
- 641 47. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, et al. Large-scale genome-wide  
642 meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different  
643 diagnosis criteria. *PLoS Genet*. 2018;14(12):e1007813.

- 644 48. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank  
645 resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-9.
- 646 49. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-  
647 outcome relationship using instrumental variables with application to Mendelian randomization.  
648 *Genet Epidemiol*. 2017;41(4):341-52.
- 649 50. Carter AR, Sanderson E, Hammerton G, Richmond RC, Smith GD, Heron J, et al.  
650 Mendelian randomisation for mediation analysis: current methods and challenges for  
651 implementation. *bioRxiv*. 2020:835819.
- 652 51. Kilpelainen TO, Carli JF, Skowronski AA, Sun Q, Kriebel J, Feitosa MF, et al. Genome-  
653 wide meta-analysis uncovers novel loci influencing circulating leptin levels. *Nat Commun*.  
654 2016;7:10494.
- 655 52. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, et al. A  
656 genome-wide approach accounting for body mass index identifies genetic variants influencing  
657 fasting glycaemic traits and insulin resistance. *Nat Genet*. 2012;44(6):659-69.
- 658 53. Walford GA, Gustafsson S, Rybin D, Stancakova A, Chen H, Liu CT, et al. Genome-  
659 Wide Association Study of the Modified Stumvoll Insulin Sensitivity Index Identifies BCL2 and  
660 FAM19A2 as Novel Insulin Sensitivity Loci. *Diabetes*. 2016;65(10):3200-11.
- 661 54. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between  
662 imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017;13(11):e1007081.
- 663 55. Doob JL. The Limiting Distributions of Certain Statistics. *The Annals of Mathematical*  
664 *Statistics*. 1935;6(3):160-9, 10.
- 665 56. Pers TH, Timshel P, Hirschhorn JN. SNPsnap: a Web-based tool for identification and  
666 annotation of matched SNPs. *Bioinformatics*. 2015;31(3):418-20.
- 667 57. Foley CN, Mason AM, Kirk PDW, Burgess S. MR-Clust: clustering of genetic variants in  
668 Mendelian randomization with similar causal estimates. *Bioinformatics*. 2021;37(4):531-41.
- 669 58. Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting Causal  
670 Pathways Using Mendelian Randomization with Summarized Genetic Data: Application to Age  
671 at Menarche and Risk of Breast Cancer. *Genetics*. 2017;207(2):481-7.
- 672 59. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for  
673 establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol*.  
674 2012;41(1):161-76.

- 675 60. Perry JR, Corre T, Esko T, Chasman DI, Fischer K, Franceschini N, et al. A genome-  
676 wide association study of early menopause and the combined impact of identified variants. *Hum*  
677 *Mol Genet.* 2013;22(7):1465-72.
- 678 61. Jones MR, Mathur R, Cui J, Guo X, Azziz R, Goodarzi MO. Independent confirmation of  
679 association between metabolic phenotypes of polycystic ovary syndrome and variation in the  
680 type 6 17beta-hydroxysteroid dehydrogenase gene. *J Clin Endocrinol Metab.* 2009;94(12):5034-  
681 8.
- 682 62. Ortega-Azorin C, Coltell O, Asensio EM, Sorli JV, Gonzalez JI, Portoles O, et al.  
683 Candidate Gene and Genome-Wide Association Studies for Circulating Leptin Levels Reveal  
684 Population and Sex-Specific Associations in High Cardiovascular Risk Mediterranean Subjects.  
685 *Nutrients.* 2019;11(11).
- 686 63. Miranda-Lora AL, Cruz M, Aguirre-Hernandez J, Molina-Diaz M, Gutierrez J, Flores-  
687 Huerta S, et al. Exploring single nucleotide polymorphisms previously related to obesity and  
688 metabolic traits in pediatric-onset type 2 diabetes. *Acta Diabetol.* 2017;54(7):653-62.
- 689 64. Nielsen TR, Appel EV, Svendstrup M, Ohrt JD, Dahl M, Fonvig CE, et al. A genome-  
690 wide association study of thyroid stimulating hormone and free thyroxine in Danish children and  
691 adolescents. *PLoS One.* 2017;12(3):e0174204.
- 692 65. Ramly B, Afiqah-Aleng N, Mohamed-Hussein ZA. Protein-Protein Interaction Network  
693 Analysis Reveals Several Diseases Highly Associated with Polycystic Ovarian Syndrome. *Int J*  
694 *Mol Sci.* 2019;20(12).
- 695 66. Wu TI, Huang RL, Su PH, Mao SP, Wu CH, Lai HC. Ovarian cancer detection by DNA  
696 methylation in cervical scrapings. *Clin Epigenetics.* 2019;11(1):166.
- 697 67. Makabe T, Arai E, Hirano T, Ito N, Fukamachi Y, Takahashi Y, et al. Genome-wide  
698 DNA methylation profile of early-onset endometrial cancer: its correlation with genetic  
699 aberrations and comparison with late-onset endometrial cancer. *Carcinogenesis.*  
700 2019;40(5):611-23.
- 701 68. Barber TM, Franks S. Obesity and polycystic ovary syndrome. *Clin Endocrinol (Oxf).*  
702 2021.
- 703 69. Teede HJ, Joham AE, Paul E, Moran LJ, Loxton D, Jolley D, et al. Longitudinal weight  
704 gain in women identified with polycystic ovary syndrome: results of an observational study in  
705 young women. *Obesity (Silver Spring).* 2013;21(8):1526-32.

- 706 70. Norgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies  
707 based on large health care databases: problems and potential solutions - a primer for the  
708 clinician. *Clin Epidemiol*. 2017;9:185-93.
- 709 71. VanderWeele TJ, Hernan MA, Robins JM. Causal directed acyclic graphs and the  
710 direction of unmeasured confounding bias. *Epidemiology*. 2008;19(5):720-8.
- 711 72. Ezeh U, Yildiz BO, Azziz R. Referral bias in defining the phenotype and prevalence of  
712 obesity in polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2013;98(6):E1088-96.
- 713 73. Luque-Ramirez M, Alpanes M, Sanchon R, Fernandez-Duran E, Ortiz-Flores AE,  
714 Escobar-Morreale HF. Referral bias in female functional hyperandrogenism and polycystic ovary  
715 syndrome. *Eur J Endocrinol*. 2015;173(5):603-10.
- 716 74. Karvonen-Gutierrez C, Kim C. Association of Mid-Life Changes in Body Size, Body  
717 Composition and Obesity Status with the Menopausal Transition. *Healthcare (Basel)*. 2016;4(3).
- 718 75. McCann SE, Freudenheim JL, Darrow SL, Batt RE, Zielezny MA. Endometriosis and  
719 body fat distribution. *Obstet Gynecol*. 1993;82(4 Pt 1):545-9.
- 720 76. Sun K, Xie Y, Zhao N, Li Z. A case-control study of the relationship between visceral fat  
721 and development of uterine fibroids. *Exp Ther Med*. 2019;18(1):404-10.
- 722 77. Sato F, Nishi M, Kudo R, Miyake H. Body fat distribution and uterine leiomyomas. *J*  
723 *Epidemiol*. 1998;8(3):176-80.
- 724 78. Lord J, Thomas R, Fox B, Acharya U, Wilkin T. The central issue? Visceral fat mass is a  
725 good marker of insulin resistance and metabolic disturbance in women with polycystic ovary  
726 syndrome. *BJOG*. 2006;113(10):1203-9.
- 727 79. Ferrero S, Anserini P, Remorgida V, Ragni N. Body mass index in endometriosis. *Eur J*  
728 *Obstet Gynecol Reprod Biol*. 2005;121(1):94-8.
- 729 80. Hediger ML, Hartnett HJ, Louis GM. Association of endometriosis with body size and  
730 figure. *Fertil Steril*. 2005;84(5):1366-74.
- 731 81. Berlanda N, Somigliana E, Frattaruolo MP, Buggio L, Dridi D, Vercellini P. Surgery  
732 versus hormonal therapy for deep endometriosis: is it a choice of the physician? *Eur J Obstet*  
733 *Gynecol Reprod Biol*. 2017;209:67-71.
- 734 82. Kim SA, Um MJ, Kim HK, Kim SJ, Moon SJ, Jung H. Study of dienogest for  
735 dysmenorrhea and pelvic pain associated with endometriosis. *Obstet Gynecol Sci*.  
736 2016;59(6):506-11.

- 737 83. Jeng CJ, Chuang L, Shen J. A comparison of progestogens or oral contraceptives and  
738 gonadotropin-releasing hormone agonists for the treatment of endometriosis: a systematic  
739 review. *Expert Opin Pharmacother*. 2014;15(6):767-73.
- 740 84. Okifuji A, Hare BD. The association between chronic pain and obesity. *J Pain Res*.  
741 2015;8:399-408.
- 742 85. Engstrom G, Hedblad B, Stavenow L, Lind P, Janzon L, Lindgarde F. Inflammation-  
743 sensitive plasma proteins are associated with future weight gain. *Diabetes*. 2003;52(8):2097-101.
- 744 86. Yasui T, Hayashi K, Mizunuma H, Kubota T, Aso T, Matsumura Y, et al. Association of  
745 endometriosis-related infertility with age at menopause. *Maturitas*. 2011;69(3):279-83.
- 746 87. Sirmans SM, Pate KA. Epidemiology, diagnosis, and management of polycystic ovary  
747 syndrome. *Clin Epidemiol*. 2013;6:1-13.
- 748 88. Laven JS, Imani B, Eijkemans MJ, Fauser BC. New approach to polycystic ovary  
749 syndrome and other forms of anovulatory infertility. *Obstet Gynecol Surv*. 2002;57(11):755-67.
- 750 89. Mardh PA. Tubal factor infertility, with special regard to chlamydial salpingitis. *Curr*  
751 *Opin Infect Dis*. 2004;17(1):49-52.
- 752 90. de Ziegler D, Borghese B, Chapron C. Endometriosis and infertility: pathophysiology and  
753 management. *Lancet*. 2010;376(9742):730-8.
- 754 91. Homer HA. The Role of Oocyte Quality in Explaining "Unexplained" Infertility. *Semin*  
755 *Reprod Med*. 2020;38(1):21-8.
- 756 92. Arojoki M, Jokimaa V, Juuti A, Koskinen P, Irjala K, Anttila L. Hypothyroidism among  
757 infertile women in Finland. *Gynecol Endocrinol*. 2000;14(2):127-31.
- 758 93. Guerin LR, Prins JR, Robertson SA. Regulatory T-cells and immune tolerance in  
759 pregnancy: a new target for infertility treatment? *Hum Reprod Update*. 2009;15(5):517-35.
- 760 94. Luciano AA, Lanzone A, Goverde AJ. Management of female infertility from hormonal  
761 causes. *Int J Gynaecol Obstet*. 2013;123 Suppl 2:S9-17.
- 762 95. Sen A, Kushnir VA, Barad DH, Gleicher N. Endocrine autoimmune diseases and female  
763 infertility. *Nat Rev Endocrinol*. 2014;10(1):37-50.
- 764 96. Ramlau-Hansen CH, Thulstrup AM, Nohr EA, Bonde JP, Sorensen TI, Olsen J.  
765 Subfecundity in overweight and obese couples. *Hum Reprod*. 2007;22(6):1634-7.

- 766 97. Nilufer R, Karina B, Paraskevi C, Rebecca D, Genevieve G, Ayush G, et al. Large-scale  
767 genome-wide association meta-analysis of endometriosis reveals 13 novel loci and genetically-  
768 associated comorbidity with other pain conditions. *bioRxiv*. 2018:406967.
- 769 98. Catherino WH, Eltoukhi HM, Al-Hendy A. Racial and ethnic differences in the  
770 pathogenesis and clinical manifestations of uterine leiomyoma. *Semin Reprod Med*.  
771 2013;31(5):370-9.
- 772 99. Marsh EE, Ekpo GE, Cardozo ER, Brocks M, Dune T, Cohen LS. Racial differences in  
773 fibroid prevalence and ultrasound findings in asymptomatic young women (18-30 years old): a  
774 pilot study. *Fertil Steril*. 2013;99(7):1951-7.
- 775 100. Bougie O, Healey J, Singh SS. Behind the times: revisiting endometriosis and race. *Am J*  
776 *Obstet Gynecol*. 2019;221(1):35 e1- e5.
- 777 101. Kim JJ, Choi YM. Phenotype and genotype of polycystic ovary syndrome in Asia: Ethnic  
778 differences. *J Obstet Gynaecol Res*. 2019;45(12):2330-7.
- 779 102. Nakimuli A, Chazara O, Byamugisha J, Elliott AM, Kaleebu P, Mirembe F, et al.  
780 Pregnancy, parturition and preeclampsia in women of African ancestry. *Am J Obstet Gynecol*.  
781 2014;210(6):510-20 e1.
- 782 103. Tan X, Li S, Chang Y, Fang C, Liu H, Zhang X, et al. Effect of metformin treatment  
783 during pregnancy on women with PCOS: a systematic review and meta-analysis. *Clin Invest*  
784 *Med*. 2016;39(4):E120-31.

785

786

787

788

789

790

791

792

793

## 794 **Supporting Information Captions**

795 **S1 Figure. Comparison of different Mendelian randomisation methods for regression of**  
796 **female reproductive diseases on various obesity traits.** Summary statistics MR was performed  
797 with the R package TwoSampleMR using three different methods whose results are compared.  
798 Effect sizes are displayed as odds ratios (ORs) with 95% confidence intervals (CIs). P-values are  
799 adjusted for multiple testing with the false discovery rate (FDR) correction; solid lines indicate  
800 associations that are significant at an adjusted p-value threshold of 0.05. BMI = body mass index,  
801 IVW = Inverse-variance weighted, PCOS = polycystic ovary syndrome, WHR = waist-hip ratio,  
802 WHRadjBMI = WHR adjusted for BMI, VAT = genetically predicted visceral adipose tissue  
803 mass.

804 **S2 Figure. Mendelian randomisation sensitivity analyses for female reproductive disorders**  
805 **regressed on obesity traits. (A)** Female-specific genetic instruments vs combined-sexes genetic  
806 instruments. **(B)** Reproductive outcomes from meta-analysis of UKBB and FinnGen summary  
807 statistics vs those from FinnGen only. Summary statistics MR performed with TwoSampleMR R  
808 package and best method (displayed, IVW) chosen via Rucker's framework. Odds ratios (ORs)  
809 with 95% confidence intervals (CIs) displayed.

810 **S3 Figure. (A) Non-linear Mendelian randomisation (MR) estimates for relationships**  
811 **between BMI and female reproductive disorders.** Localised average causal estimates (LACE)  
812 are calculated by dividing the IV-free exposure into 100 quantiles (fractional polynomial  
813 method) or 10 quantiles (piecewise linear method) with the odds ratio (OR) and 95% confidence  
814 intervals (CI) displayed. The reference point for OR = 1 is mean BMI, 27.0 kg/m<sup>2</sup>. **(B)**  
815 **Heterogeneity across BMI spectrum in instrument variables used for non-linear MR.**  
816 Instrument variable (IV)-free BMI was divided into quantiles as described in (A). Left: fractional  
817 polynomial 100 quantiles, right: piecewise linear 10 quantiles. Proportion of variance in BMI  
818 explained by instrument SNPs in each quantile is plotted ( $\beta$ ), with 95% confidence intervals  
819 shown as standard error bars. Spont. misc. = spontaneous miscarriage.

820 **S4 Figure. Single-SNP genetic effect estimates for obesity instruments on female**  
821 **reproductive disorders.** SNPs are annotated with their nearest gene by SNPsnap and clustered  
822 by obesity-female reproductive disorder relationship. SNPs with  $\geq 80\%$  probability of  
823 belonging to a substantial cluster (MRClust) are displayed, with an \* if the SNP belongs to  
824 substantial clusters for multiple disorders. Effect size estimates are scaled to a variance of 1  
825 within each disease.

826 **S1 Table. Supplemental Tables A-P.**