The genetics of sex hormones and their effects on mammographic density in women

Cameron B. Haas

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2021

Reading Committee:
Sara Lindström, Chair
Karen Wernli
Li Hsu

Program Authorized to Offer Degree:
Epidemiology

University of Washington

Abstract

The genetics of sex hormones and their effects on mammographic density in women

Cameron B. Haas

Chair of the Supervisory Committee:
Sara Lindström
Department of Epidemiology

In this work we leveraged genomic information from large-scale population-based studies to explore the relationships between three epidemiologic factors associated with breast cancer in women: 1) mammographic density, 2) sex hormone concentrations, and 3) body mass index (BMI). Mammographic density, which describes the proportion of dense (i.e., epithelial and stromal) tissue in the breast, is one of the strongest predictors of breast cancer in women. Women with extremely dense breasts have a 3 to 6-fold increased risk of breast cancer compared to those with primarily fatty breasts. Breast cancer is generally considered to be a primarily hormone-driven cancer, an attribute that has led to the development of effective treatment and prophylactic strategies for hormone receptor positive subtypes and cause for investigating the role of endogenous hormones in breast cancer etiology. Finally, BMI has been consistently observed to have paradoxical associations with breast cancer across menopause, with evidence of preventative effects associated with higher BMI in premenopausal women but increased risk in postmenopausal women.

We first built on recent analyses that investigated the genetic architecture of testosterone and sex hormone binding globulin (SHBG) in men and women of European ancestry by conducting genome-wide association studies (GWAS) of estradiol concentrations in women. Additionally, we investigated the generalizability of previous findings in women of African ancestry. We further conducted menopausal status specific GWAS of these sex hormones to identify loci with heterogeneous effects across menopause. We found that the strongest overall genetic predictor of testosterone concentrations, located in the *CYP3A7* gene, had an effect nearly twice as large in premenopausal women compared to postmenopausal women. Similarly, genetic variants in the *AKR1C4* gene were strongly associated with concentrations of SHBG in premenopausal women, but not in postmenopausal women, with a 5-fold difference in effect estimates between the two. We also estimated the shared heritability across menopausal status specific hormone concentrations, and observed a relatively low genetic correlation between pre- and postmenopausal detectable levels of estradiol, whereas comparisons of pre- and postmenopausal shared heritability for SHBG and testosterone were both close to one, indicating near identical genetic architectures. We performed gene-level tests for enrichment of genetic associations within tissue-specific gene expressions by collapsing multiple SNP-level associations in a gene while accounting for linkage disequilibrium. Using this gene-set analysis for tissue specificity we observed a change from strong adrenal gland tissue specificity of testosterone in premenopausal to adipose tissue specificity in postmenopausal women, suggesting that adiposity may play a more important role in determining circulating concentrations of testosterone after menopause.

To understand the directional relationships between overall and menopausal status specific concentrations of sex hormones and BMI on mammographic dense and non-dense area we performed Mendelian Randomization analyses. We created menopausal status specific genetic instruments for SHBG, testosterone, and estradiol based on our previous work. We obtained single nucleotide polymorphisms (SNP)-specific association statistics from a recent GWAS of mammographic density of up to 27,900 women of European ancestry. Effect estimates for BMI were obtained from the largest meta-GWAS of BMI to date, comprising more than 700,000 individuals. We observed an inverse relationship between overall genetically predicted testosterone and dense area. Increasing genetically predicted BMI was strongly associated with an increase in genetically predicted non-dense area, as previously observed. However, we also observed an inverse association between genetically predicted BMI and absolute dense area, which might explain some of the reduced risk of breast cancer associated with an increase in genetically predicted BMI. Higher genetically predicted BMI was also strongly associated with decreasing SHBG concentrations, as well as increasing concentrations of testosterone. Based on the inverse-variance weighted results, we observed increasing genetically predicted BMI to be associated with a decrease in genetically predicted detectable levels of overall and premenopausal specific estradiol concentrations, but not for postmenopausal only. Multivariable MR approaches for the association of BMI and mammographic density adjusting for sex hormones did not substantively change the effect estimates of BMI.

Building on the strong association between BMI and mammographic density, we sought to identify genetic loci that interact with BMI to alter mammographic density phenotypes. We conducted genome-wide tests for the interaction between SNPs and BMI on percent mammographic density, absolute dense area, and absolute non-dense area in 14,837 women. Despite having the largest sample size to date with genetic and phenotypic data for mammographic density, we did not find any loci that reached standard Bonferroni correction for statistical significance.

This work presents novel findings of the unique genetic architectures of menopausal specific concentrations of sex hormones in women and extends these findings to investigate their associations with mammographic density. We show that BMI plays an important role in determining not only non-dense area, but also dense area and a possibly separate mechanism for breast cancer etiology. Additionally, there is evidence based on our MR approaches of a regulatory role of BMI on endogenous estradiol as yet another possible pathway to tumorigenesis. We did not identify any genetic variant that has a strong modifying effect of BMI on mammographic density phenotypes. It is possible that larger studies are merited to investigate the interactions between germline genetic variants and BMI on mammographic density variation.

Table of Contents

List of Main Tables

List of Main Figures

Acknowledgments

I would like to thank my chair, Dr. Sara Lindström, for her expertise, mentorship, guidance, and constant encouragement. From my first genetic epidemiology course to today I have gained confidence in my own understanding of genomic approaches to understanding disease etiology thanks to her. I'm grateful for the relationship we've developed over these years, the progressive and inclusive environment she's created in our research group, and for providing the academic and emotional resources I needed to maintain my sanity while working on this dissertation.
I would also like to thank the individuals who have provided countless opportunities to do work that is not included in this dissertation. Dr. Karen J. Wernli in particular has been a huge advocate for me since my master's degree, giving me projects and work that have been invaluable for my development as a researcher. I really have gained so much by being associated with her and the cancer researchers at Kaiser Permanente Washington Health Research Institute.
Lastly, thank you to the staff and faculty in the Department of Epidemiology at UW who have answered questions, provided reimbursements, and done so much behind the scenes to help me graduate. Thank you to all of the committees and people doing the unseen work to create a program in which students may thrive.

Dedication

To my husband for being everything a partner should be and doing your best to be exactly the support I needed.

Cross-ancestry genome wide association studies of sex hormones in pre- and postmenopausal women in the UK Biobank

Abstract

Objective: Concentrations of circulating sex hormones have been associated with a variety of diseases in women and are strongly influenced by menopausal status. We investigated the genetic architectures of circulating concentrations of estradiol, testosterone, and sex hormone binding globulin (SHBG) by menopausal status in women of European and African ancestry.
Methods: Using data on 229,966 women from UK Biobank, we conducted genome-wide association studies (GWAS) of circulating concentrations of three sex hormones (estradiol, testosterone, SHBG) in premenopausal and postmenopausal women, respectively. We tested for evidence of heterogeneity of genetic effects by menopausal status and genetic ancestry. We conducted gene-based enrichment analyses to identify tissues in which genes with GWAS-enriched signals were expressed. Finally, we estimated the genetic correlation (r_g) between pre- and postmenopausal specific concentrations of circulating sex hormones.
Results: We identified four loci (5q35.2, 12q14.3, 19q13.42, 20p12.3) that were associated with detectable concentrations of estradiol in both pre- and postmenopausal women of European ancestry. Three of these loci have been previously shown to be associated with ovarian aging. Heterogeneity analysis by menopausal status identified one locus for testosterone (7q22.1) in the *CYP3A7* gene with an effect size in premenopausal women that was twice as large as that in postmenopausal women, and one locus that was strongly associated with concentrations of SHBG in premenopausal women but weakly associated in postmenopausal women (10q15.1) near the *AKR1C4* gene. Our gene-based analysis of testosterone revealed evidence of enrichment of GWAS signals in genes expressed in adrenal gland tissue for premenopausal women and adipose tissue for postmenopausal women. We did not find any evidence of ancestry-specific genetic effects for concentrations of estradiol, testosterone, or SHBG. Lastly, genetic correlation analysis revealed relatively low shared heritability of detectable concentrations of estradiol in pre- versus postmenopausal women (r_g=0.55), but high shared heritability between pre- and postmenopausal concentrations of testosterone (r_g=0.93) and SHBG (r_g=0.89).
Conclusions: Our study provides novel insights into the differential genetic architectures of circulating concentrations of sex hormone in pre- and postmenopausal women. In particular, we identified specific loci that showed genome-wide significant evidence of heterogeneity by menopausal status for testosterone and SHBG. We also observed support for a more prominent role of genetic variants located near genes expressed in adipose tissue in determining testosterone concentrations among postmenopausal women, as compared to premenopausal women.

Introduction

Concentrations of testosterone and estrogens have been implicated for their roles in the development of numerous diseases and disease-related traits, including cancers, cardiovascular disease, diabetes, osteoporosis, and metabolic syndromes(Wagner, Kaplan and Burkman, 2002; Folkerd and Dowsett, 2013; Kim and Halter, 2014). Sex-specific determinants of circulating hormone concentrations as well as sex-specific effects of hormones on disease risk have been widely recognized(Vitale, Mendelsohn and Rosano, 2009; Lauretta *et al*., 2018). In women, concentrations of circulating testosterone and estrogen decline with age, with markedly lower natural production of estrogen with the loss of ovarian follicular activity due to menopause(Burger *et al*., 2007; Flynn *et al*., 2021).

Leveraging the release of biomarker and genome-wide genotype data in the UK Biobank(Bycroft *et al*., 2018), recent work conducted by Ruth and colleagues has provided significant insight into the genetic architecture underlying variation in circulating sex hormones, mainly in men of European ancestry(Ruth *et al*., 2020). Hundreds of independent genetic loci have been associated with concentrations of testosterone and sex hormone binding globulin (SHBG), two highly heritable phenotypes (h^2=13-57% and 20-68%, respectively)(Ring *et al*., 2005; Varghese *et al*., 2012; Ruth *et al*., 2020). However, important gaps remain for understanding the potential impact of menopause on genetic determinants of circulating concentrations of testosterone and SHBG, specifically in women.

To an even greater extent than testosterone and SHBG, the genetic architecture of endogenous estradiol concentrations in women is poorly understood. Estrogens are involved in sex-specific processes and have been associated with numerous sex-linked disease processes including breast cancer development and osteoporosis(Jones *et al*., 2013; Woolcott *et al*., 2013). Detectable concentrations of estradiol in women are predominantly determined by menopausal status, with 97% of postmenopausal women in the UK Biobank having concentrations below the detectable threshold, as compared to 35% of premenopausal women. Consequently, an agnostic genome-wide association study (GWAS) ignoring menopausal status will be biased towards detecting loci associated with age at onset of menopause rather than true variation in estradiol(Ruth *et al*., 2020).

The number of known associations between common genetic variants and human phenotypes continues to grow dramatically, in part due to large-scale biobanks such as the UK Biobank(Bycroft *et al*., 2017). However, genetic association studies still tend to be conducted exclusively among individuals of European ancestry(Popejoy and Fullerton, 2016). The implications of this lack of diversity could result in misidentification of causal variants and overestimation of effects across populations(Carlson *et al*., 2013).

In this study, we conducted GWAS of circulating concentrations of estradiol, testosterone, and SHBG, stratified by menopausal status in women of European and African ancestry. We estimated genetic correlations across menopausal-stratified sex hormone phenotypes and explored the consistency in the genetic architecture of circulating concentrations in sex hormones across European and African ancestries.

Methods

UK Biobank participants and biomarkers

Details regarding survey, genotyping, and biomarker data collection have been presented elsewhere(Bycroft *et al*., 2017, 2018). Briefly, the UK Biobank study collected 34 biomarkers at baseline for ~500,000 participants along with written informed consent and extensive questionnaire information(Bycroft *et al*., 2018). Our analyses focus on measured concentrations of estradiol (field 30800-0.0), testosterone (field 30850-0.0), and SHBG (field 20830-0.0), which were assayed from blood serum at baseline. We excluded women who were pregnant at the time of data collection (field 3140-0.0). In postmenopausal-specific analyses related to estradiol we further excluded women who had reported ever having taken hormone replacement therapy (field 2814-0.0), as exogenous hormones would likely overshadow genetic determinants of circulating estradiol concentrations and reduce our ability to find genetic loci of significance. We also excluded participants who requested to withdraw from the UK Biobank as of 08-20-2020. This research was approved under UK Biobank application 55120.

Defining menopausal status

We defined premenopausal status as having answered "No" to self-reported menopause (field 2724-0.0) and younger than age 60 at the time of reference (field 21003-0.0)(Flynn *et al*., 2021). Women who were pregnant at the time of reference were excluded (3140-0.0). Postmenopausal women were classified as having answered "Yes" to having entered menopause (field 2724-0.0) and at least two years since their last period (field 3581-0.0), which was required to have occur after the age of 40 in order to exclude premature menopause. We excluded women who had ever used hormone replacement therapy in the postmenopausal specific analyses (field 2814-0.0).

Genetic data and empirically assigning ancestry

We use the 'v3' release of the UK Biobank genetic data related with the Haplotype Reference Consortium and 1000 Genome Project (1KGP) imputed variants. We used the 1KGP as our reference population to assign ancestry for all samples with genotype data in the UK Biobank following methods described by Petersen and colleagues(1000 Genomes Project Consortium *et al*., 2015; Peterson *et al*., 2017). Using all samples in the UK Biobank with available genotype data, we used Plink 1.9(Purcell *et al*., 2007) to identify a set of independent variants (--indep-pairwise 10000 10 0.1) and extract a common set of variants from 1KGP with minor allele frequency (MAF)>0.05. We used GenotypeHarmonizer(Deelen *et al*., 2014) to check for strand flips in the UK Biobank genotype data using the 1KGP as a reference before using PLINK1.9 to perform principal components analysis on the harmonized variants that overlap between UK Biobank and 1KGP(Price *et al*., 2006). We then projected those variant weights to UK Biobank samples to calculate ten principal components.

We use the 26 population descriptions in 1KGP to estimate within-population standard deviations for Mahalanobis distances based on the first 10 principal components and removed samples greater than 4 standard deviations away from the mean within their population. We then calculated Mahalanobis distances for all samples in the UK Biobank using the mean and covariance matrix from 1KGP with outliers removed. We assigned all samples to the population with the smallest Mahalanobis distance and their respective superpopulation(Peterson *et al*., 2017). We assigned 431,239 samples as European ancestry and 9,501 samples as African ancestry. We used Plink2.0 to calculate MAF, Hardy-Weinberg equilibrium, and variant missingness within each empirically assigned ancestry population.

Statistical Analyses

Genome-wide association studies

We used PLINK2.0(Purcell and Chang, no date; Chang *et al*., 2015) to perform GWAS analyses. We used GCTA(Yang *et al*., 2011) to identify unrelated individuals within each assigned ancestry population (--king-cutoff 0.125). We modeled all phenotypes using inverse-rank normalized transformations adjusting for age at reference (field 21003-0.0), centre (field 54-0.0), genotyping array (field 22000-0.0), and the first 10 principal components (field 22009-0.1-10) as covariates within empirically assigned ancestry populations, and within menopausal status for menopausal status-specific GWAS. Time since last period, created using age at last period (field 3581-0.0) and age at reference, was also included in the postmenopausal analyses in order to limit any bias for menopause-related associations. Body mass index (BMI) (field 21001-0.0) was additionally included in all SHBG analyses to increase comparability with previous GWAS findings(Ruth *et al*., 2020). We filtered variants with Hardy-Weinberg equilibrium p-value<10-6, minor allele counts<20, MAF<0.01, imputation quality<0.8, genotype missingness>0.1, and excluded individuals who had a genotype missing rate >0.1.

Genome-wide genetic correlations between traits

We estimated the heritability due to genotyped SNPs and genetic correlations between traits using LD Score Regression based on GWAS summary statistics for the European ancestry population(B. K. Bulik-Sullivan, Loh, Finucane, Ripke, Yang, Patterson, *et al*., 2015; Finucane *et al*., 2015). Due to the smaller sample size, we used bivariate GREML analysis from GCTA, which relies a genetic relatedness matrix rather than a reference population, to estimate the heritability and genetic correlation between traits among African ancestry women(Yang *et al*., 2011).

Meta-analysis

We performed fixed effects inverse-variance weighted meta-analysis to conduct cross-ancestry GWAS of each sex hormone using the METAL software (Willer, Li and Abecasis, 2010). We used the Cochran's Q test as implemented in METAL to test for heterogeneity of effects between European and African ancestries. We present the ratio of beta coefficients between African ancestry and European ancestry women as a function of their meta-analysis p-values (Carlson *et al*., 2013; Ruth *et al*., 2020). We additionally conducted genome-wide SNP-specific tests for heterogeneity to identify genetic variants with statistically significant differences in effect size by menopausal status in the European-ancestry GWAS based on the results of the Cochran's *Q*-test as implemented in METAL(Willer, Li and Abecasis, 2010). For all genome-wide analyses, we considered $p<5x10^{-8}$ as statistically significant.

Gene-level association analysis and tissue-specific enrichment

Gene-level association analysis was performed using the MAGMA (multi-marker analysis of genomic annotation)(de Leeuw *et al*., 2015) tool implemented in the Functional Mapping (FUMA) program for GWAS results. MAGMA analyzes SNP-level summary genetic associations while taking linkage disequilibrium (LD) between SNPs into account (Watanabe *et*

al., 2017). SNPs were mapped to a gene if they were located between the start and end sites of the gene. All analyses were based on the human genome build 37.

MAGMA "gene-property analysis" was performed to test for associations between gene-level signals and tissue-specific gene expression profiles. Gene-property analysis uses a multi-variable regression model that includes gene expression in a specific tissue type and the average gene expression across 54 tissue types to evaluate the relationship between tissue specificity and gene-level association. Tissue specific gene expression data were from GTEx version 7(GTEx Consortium, 2013). We performed gene property analysis for tissue specificity of testosterone and SHBG using the GWAS results from women of European ancestry overall and by menopausal status.

Results

Our final sample sizes were 229,966 European ancestry women and 5,336 African ancestry women. Among those empirically assigned as European ancestry, 51,081 were premenopausal and 84,194 postmenopausal. Of those women empirically assigned as African ancestry, 1,994 were premenopausal and 1,200 were postmenopausal. Due to the relatively small sample sizes, results within menopausal status for African ancestry women are only reported in the supplementary data(Haas, no date).

As previously mentioned, we observed a greater proportion of premenopausal women with detectable levels of estradiol compared to postmenopausal women (**Table 1.1**). We also observed greater mean concentrations of testosterone and SHBG in premenopausal women compared to postmenopausal women, and across both pre- and postmenopausal women, we observed right skewed distributions of sex hormones as indicated by the lower median compared to the mean and in relation to the interquartile range (**Table 1**). While women of African ancestry had lower mean concentrations of all three sex hormones, differences were not statistically significant.

Estradiol

We analyzed estradiol as a continuous phenotype (pmol/L) both overall and stratified by menopausal status. Among women with detectable concentrations, 33,032 European ancestry women were classified as premenopausal and 3,759 as postmenopausal (**Supplementary Table 1.1**). We identified one locus (lead SNP rs727428) in the 17p13.1 region near the *SHBG* gene, that was significant in both the overall ($p=2.1x10^{-10}$) analysis and when restricted to premenopausal women only ($p=1.04x10^{-8}$) (**Table 1.2**; **Supplementary Figures 1.1A-B**). Among African ancestry women, 1,706 had detectable concentrations of estradiol, of which 1,285 were classified as premenopausal and only 76 were included in postmenopausal specific GWAS. No variants reached genome-wide significance in the African ancestry GWAS. The lead SNP, rs727428, from the European ancestry analysis only showed moderate association ($p=8x10^{-3}$) in the African ancestry analysis (**Supplementary Figures 1.1C-D**). We observed no evidence of heterogeneity in effect estimates between European and African ancestry women among overlapping variants (**Supplementary Table 1.1**).

Using estradiol as a dichotomized phenotype for detectable concentrations (yes/no) in 229,966 women of European ancestry, we report ten loci that reached genome-wide significance in the overall analysis(**Supplementary Table 1.3**). Among these ten loci, two were significant in postmenopausal but not premenopausal women (4q21.23, 8p11.23), and four were significant in

both pre- and postmenopausal women (5q35.2, 12q14.3, 19q13.42, 20p12.3)(**Table 1.2; Figure 1.1; Supplementary Figure 1.2**). No loci were found to be associated with detectable concentrations of estradiol in premenopausal women but not in postmenopausal women.

We included 5,336 women in the African-ancestry GWAS of detectable concentrations of estradiol. Out of the ten SNPs identified in European ancestry, only four were included in our African ancestry analyses after quality control (**Supplementary Table 1.2; Supplementary Figures 2B-C**). Of those four SNPs, three were directionally consistent with the overall European ancestry GWAS but none had p-values < 0.05 or showed evidence of heterogeneity of effect across ancestries (**Supplementary Table 1.2**).

Testosterone

We conducted GWAS of circulating testosterone concentrations in 182,648 women of European ancestry, of whom 44,401 were premenopausal and 84,003 were postmenopausal. Of the 226 variants reported by Ruth et al. for testosterone in women and observed in our analyses after QC, 103 variants reached statistical significance in either pre- or postmenopausal only analyses. Of those 103 variants, 14 were significant in both premenopausal and postmenopausal specific analyses, nine were significant in premenopausal but not postmenopausal women, and 72 were significant in postmenopausal but not premenopausal women (**Supplementary Table 1.3; Supplementary Figures 1.3A-B**).

We next tested for heterogeneity of SNP effects on testosterone between pre- and postmenopausal women of European ancestry. We identified a genome-wide significant signal for rs45446698 in the 7q22.1 locus near the *CYP3A7* gene, which is involved in cholesterol and steroid metabolism (**Figure 1.2**; **Figure 1.3; Supplementary Figures 1.3C-D**). There was nearly a two-fold increase in the effect estimate of this variant among premenopausal women compared to postmenopausal women (**Table 1.3**).

We observe strong evidence for adrenal gland tissue specificity for testosterone in premenopausal women (β=0.04, s.e.=0.01, $p<3x10^{-5}$), but weaker evidence in postmenopausal women (β=0.008, s.e.=0.01, $p=7.8x10^{-3}$) (**Figure 1.4**; **Supplementary Table 1.8**). We conversely observed strong evidence for adipose tissue specificity in postmenopausal women (β=0.03, s.e.=0.01, p=0.03) but no evidence in premenopausal women $\beta=-6x10^{3}$, s.e.=0.01, p=0.67).

We had GWAS and testosterone data for 4,229 African ancestry women. We observed no variants that reached genome-wide significance (**Supplementary Figures 1.3E-F)**. We had data for 162/226 previously identified variants in European ancestry women(Ruth *et al*., 2020), and of those 96 (59%) had directionally consistent effect estimates with women of European ancestry and 14 had p<0.05, including rs45446698 which tags the *CYP3A7* gene (**Figure 1.5**).

SHBG

We conducted GWAS of circulating SHBG concentrations in 196,901 women of European ancestry, of whom 43,477 were premenopausal, and 92,9211 were postmenopausal. Of the 321 variants reported by Ruth et al. to be associated with SHBG in women and observed in our analyses, 29 were genome-wide significant in both premenopausal specific and postmenopausal specific analyses, two were genome-wide significant in premenopausal but not postmenopausal

women, and 16 were genome-wide significant in postmenopausal but not premenopausal women (**Supplementary Table 1.4; Supplementary Figures 1.4A-B**).

We tested for heterogeneity of SNP effects by menopausal status and identified a signal at 10q15.1 near *AKR1C4* (**Figure 7; Supplementary Figures 1.4C-D**) for which there was a statistically significant effect in premenopausal women (β=-0.10, s.e.=0.009; p=4.3x10^{-26}) but no association in postmenopausal women (β=-0.02, s.e.=0.006; p=2.4x10^{-4}) (**Table 1.3**). The effect estimate was approximately five-fold higher in premenopausal women compared to the effect estimate in postmenopausal women (**Table 1.3**).

Our gene analyses for tissue specificity of SHBG showed strong evidence of liver tissue specificity across both pre- and postmenopausal women. At a significance level of p<0.05, we additionally observed enrichment in muscle tissue in premenopausal women (β=0.01, s.e.=0.008; p=0.04), and in thyroid tissue in postmenopausal women (β=0.03, s.e.=0.01; p=0.02).

We included 4,522 women of African ancestry in our GWAS of concentrations of SHBG. Out of the previously reported 321 variants, we observed 209 in the African-ancestry analyses, of which 148 (71%) were directionally consistent with the European-ancestry effect estimates. Out of the 209 previously reported variants, 36 had p<0.05 (**Supplementary Table 1.4**). In particular, we observed a genome-wide significant signal for the *SHBG* gene with rs727428 as the lead variant (β=0.23, s.e.=0.02; p=7.6x10^{-29}). We present the ratio of effect estimates for the 209 variants observed in European and African ancestry populations (**Figure 1.8**).

Genetic correlations between concentrations of circulating hormones

Using LD Score Regression, we estimated the heritability and genetic correlations of circulating concentrations of sex hormone traits overall and by menopausal status using GWAS summary statistics for women of European ancestry (B. Bulik-Sullivan *et al*., 2015; B. K. Bulik-Sullivan, Loh, Finucane, Ripke, Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, *et al*., 2015). Overall SNP heritability estimates for testosterone and SHBG were similar to those reported by Ruth et al., and generally increased when stratified by menopausal status (**Table 1.4**). Notably, as a binary phenotype indicating detectable concentrations, estradiol had fairly low heritability (1.5%; s.e.=0.002) that increased when stratified by menopausal status to 3.9% (s.e.=0.01) and 3.5% (s.e.=0.006) for pre- and postmenopausal concentrations, respectively. Both SHBG and testosterone showed strong genetic correlations between pre- and postmenopausal concentrations (r_g=0.89, s.e.=0.03 for SHBG; r_g=0.93, s.e. 0.07 for testosterone), while the genetic correlation between pre- and post-menopausal detectable estradiol concentrations was only 0.55 (s.e.=0.16), suggesting that some of the underlying genetic architecture is distinct. We did not estimate SNP heritability and genetic correlations for estradiol as a continuous trait due to the low sample size. We used bivariate GREML to estimate heritability and genetic correlation of the same traits in women of African ancestry. While estimates for estradiol were unstable due to the small sample size with detectable levels, the overall heritability estimates for testosterone and SHBG in the African ancestry population were similar to those of European ancestry (h^2= 0.23, s.e.=0.07 and h^2=0.26, s.e.=0.07, respectively) (**Supplementary Table 1.7**).

Discussion

In this study, we investigated the genetic architecture of circulating concentrations of sex hormones by menopausal status. We identified two loci showing differential effects on circulating hormone concentrations by menopausal status in women of European ancestry. The *CYP3A7* locus, which is the most significant region associated with concentrations of testosterone, showed a stronger effect in premenopausal women compared to postmenopausal women (β=0.44; s.e.=0.02; vs. β=0.28; s.e.=0.01, respectively). While *CYP3A7* is not a novel finding, the significant difference in effect sizes between pre- and postmenopausal women provides important insight into the different biological mechanisms underlying pre and postmenopausal testosterone concentrations. The effect of rs45446698 which tags the CYP3A7*1C allele was almost twice as large in premenopausal women compared to postmenopausal women. This SNP has previously been associated with urinary estrone glucuronide concentrations and breast cancer mortality (Johnson *et al*., 2016).

SNPs close to the aldoketo reductase 1C4 (*AKR1C4*) gene, which is involved in progesterone metabolism showed stronger association with premenopausal concentrations of SHBG compared to postmenopausal concentrations. Previous research conducted in clinical trials of estrogen therapy found that polymorphisms in *AKR1C4* were associated with mammographic density change in the combined estrogen and progestin therapy group but not in the placebo or estrogen only group(Lord *et al*., 2005). Thus, effects of *AKR1C4* may be dependent on higher concentrations of estradiol, and thus we observed little to no effect in postmenopausal women since we excluded anyone with a history of menopausal hormone therapy use.

We found that genes expressed in adipose tissue were enriched in our MAGMA analysis of circulating testosterone in postmenopausal women but not in premenopausal women, where instead adrenal gland tissue were enriched. This suggests that with the loss of ovarian function following menopause, testosterone may be more regulated by adiposity tissue. In contrast, we observed consistent enrichment of liver tissue in our SHBG analyses for both pre- and postmenopausal women, which may be an expected result as ovarian function is likely to have little effect on metabolic processes in the liver.

We did not observe any genome-wide significant associations for concentrations of estradiol or testosterone in women of African ancestry, and this is likely due to our limited sample size (N=5,336). There was a strong signal for SHBG around known variants in the *SHBG* gene(Sato *et al*., 2017). We did not find evidence of any heterogeneity of effects between European ancestry and African ancestry women for SNPs previously associated with concentrations of testosterone and SHBG based on tests of heterogeneity of effects. We compared the effect estimates of SNPs identified in European ancestry women with the estimates in African ancestry women and observed a smaller difference in effect estimates with stronger discovery p-value, suggesting that the strongest genetic determinants are likely the most generalizable across ancestries. Unfortunately, approximately 20% (54/229) of SNPs previously associated with testosterone concentrations were excluded in the quality controls due to high missing rates following standard quality control thresholds(Christopher I Amos *et al*., 2017).

Previous work by Ruth et al. discussed the limitations of analyzing estradiol for women in the UK Biobank due to the relatively low threshold for detection, which resulted in a substantial proportion of women missing measures for estradiol, disproportionately affecting postmenopausal women(Ruth *et al*., 2020). We address the bias for detecting signals associated with menopausal status rather than estradiol by looking at pre- and postmenopausal women separately. In our analysis of postmenopausal women, we adjusted our analyses for time since

menopause and excluded women who reported ever having taken menopausal hormone therapy. We modeled estradiol in two ways. We first restricted our analyses to women with detectable concentrations only, and treated estradiol as a continuous outcome. We observed a genome-wide significant association close to the *SHBG* gene, which was associated with both overall and premenopausal concentrations. The lead SNP rs727428 is located downstream of the *SHBG* gene, and has previously been associated with concentrations of SHBG but not with estradiol(Prescott *et al*., 2012). Second, we modeled estradiol as a binary variable indicating detectable versus undetectable concentrations(Ruth *et al*., 2020). Using this binary phenotype, we focused our results on signals in premenopausal women, limiting the potential bias introduced by differences in menopausal status. Four loci were associated with having detectable concentrations of estradiol in the overall and both pre- and postmenopausal only analyses. Three of these loci are near genes (12q14.3, *HELB*; 19q13.42, *TMEM150B*; and 20p12.3, *MCM8*) that have previously associated with age at natural menopause, primary ovarian insufficiency, and early menopause through mechanisms related to DNA repair(Jiao *et al*., 2018). While we cannot rule out possible residual effects of menopausal aging in our analyses, it is alternatively possible that a mechanism exists through estradiol regulation that leads to early menopause. The lead variant in the 5q35.2 locus is in strong LD with variants previously associated with hormonally driven uterine fibroids and heavy menstrual bleeding(Rafnar *et al*., 2018). Of note, we did not observe associations for variants near the *CYP1B1*, which has been used previously in candidate gene studies of aromatase inhibitors for hormone driven diseases(Dumas and Diorio, 2010; Rudolph *et al*., 2015). Sensitivity analyses adjusting for factors related to timing of menstrual cycle in premenopausal women did not identify any additional signals.

The relatively low genetic correlation between a continuous measure and a binary indicator for detectable concentrations of estradiol point to the importance of improving assay sensitivity, as these are not exchangeable traits. Additionally, unlike the high genetic correlations observed for premenopausal and postmenopausal testosterone, and premenopausal and postmenopausal SHBG, which were close to 1 (i.e., perfectly shared genetic architecture), we saw relatively low genetic correlation for premenopausal detectable concentrations of estradiol and postmenopausal detectable concentrations of estradiol. The low shared genetic architecture suggests that there are genetic determinants of this traits within menopausal status.

We used more stringent thresholds for genetic variant inclusion in our GWAS than in previous studies of sex hormones (e.g. minor allele frequency and removing indels and chromosome X), resulting in a higher proportion of missing SNPs as compared to a previous publication(Ruth *et al*., 2020). For example, of 358 variants identified by Ruth and colleagues for SHBG in women, only 321 were included in our analyses. Additionally, we observed some discrepancies in findings between our GWAS and theirs. This could be due to different adjusting factors, as Ruth and colleagues additionally adjusted for time since fasting. Previous work suggested that inclusion of fasting or dilution factors did not improve bias estimates(Neale, 2019).

In summary, we identified novel genetic variants associated with concentrations of estradiol, and also genetic variants showing differential effects on circulating testosterone and SHBG according to menopausal status. Our work provides support for stratifying analyses within relevant sample characteristics, such as menopausal status, when investigating the genetic architecture of biomarkers.

Tables and Figures

Table 1.1.1. Descriptive characteristics of sex hormone concentrations in pre- and postmenopausal women in the UK Biobank population.

Premenopausal (N=62,587)		Postmenopausal (N=124,820)	
Mean (standard deviation)	Median (interquartile range)	Mean (standard deviation)	Median (interquartile range)
Estradiol – n=39,586 (63.2%)		Estradiol – n=4,262 (3.4%)	
571 (448)	432 (290, 683)	339 (280)	257 (204, 376)
Testosterone – n=53,179 (85.0%)		Testosterone – n=95,031 (76.1%)	
1.22 (0.60)	1.13 (0.82, 1.49)	1.08 (0.67)	0.97 (0.69, 1.32)
SHBG – n=52,127 (83.3%)		SHBG – n=105,289 (84.4%)	
68.0 (33.1)	62.5 (44.2, 84.9)	59.5 (28.1)	54.8 (39.4, 74.2)

*n represents the number available for inclusion in GWAS of hormone concentrations as a continuous trait. Estradiol as a dichotomized trait indicating detectable measures of estradiol was used as an additional phenotype for GWAS.

Table 1.2. Independent signals from GWAS of estradiol in women of European ancestry.

Phenotype	Lead variant	MAF*	Region	Nearest gene	Overall		Premenopausal		Postmenopausal	
					Effect estimate (SE**)	P-value	Effect estimate (SE)	P-value	Effect estimate (SE)	P-value
Estradiol continuous (pmol/L)	rs727428	0.45	17p13.1	SHBG	-0.04 (0.006)	2.1E-10	-0.05 (0.008)	1.04E-8	-0.04 (0.02)	0.08
Detectable estradiol (binary)	rs2454949	0.49	5q35.2	ZNF346	0.02 (0.003)	5.9E-15	0.04 (0.006)	3.3E-9	0.03 (0.005)	2.4E-8
	rs75770066	0.03	12q14.3	HELB	0.06 (0.008)	5.4E-14	0.11 (0.02)	6.9E-11	0.10 (0.01)	9.1E-13
	rs71181755	0.47	19q13.42	TMEM150B	-0.03 (0.003)	7.5E-19	-0.04 (0.006)	2.9E-11	-0.03 (0.005)	8.2E-12
	rs16991615	0.06	20p12.3	MCM8	0.06 (0.006)	3.9E-25	0.10 (0.01)	1.8E-17	0.10 (0.01)	4.1E-24

*MAF=Minor allele frequency
**SE=Standard error

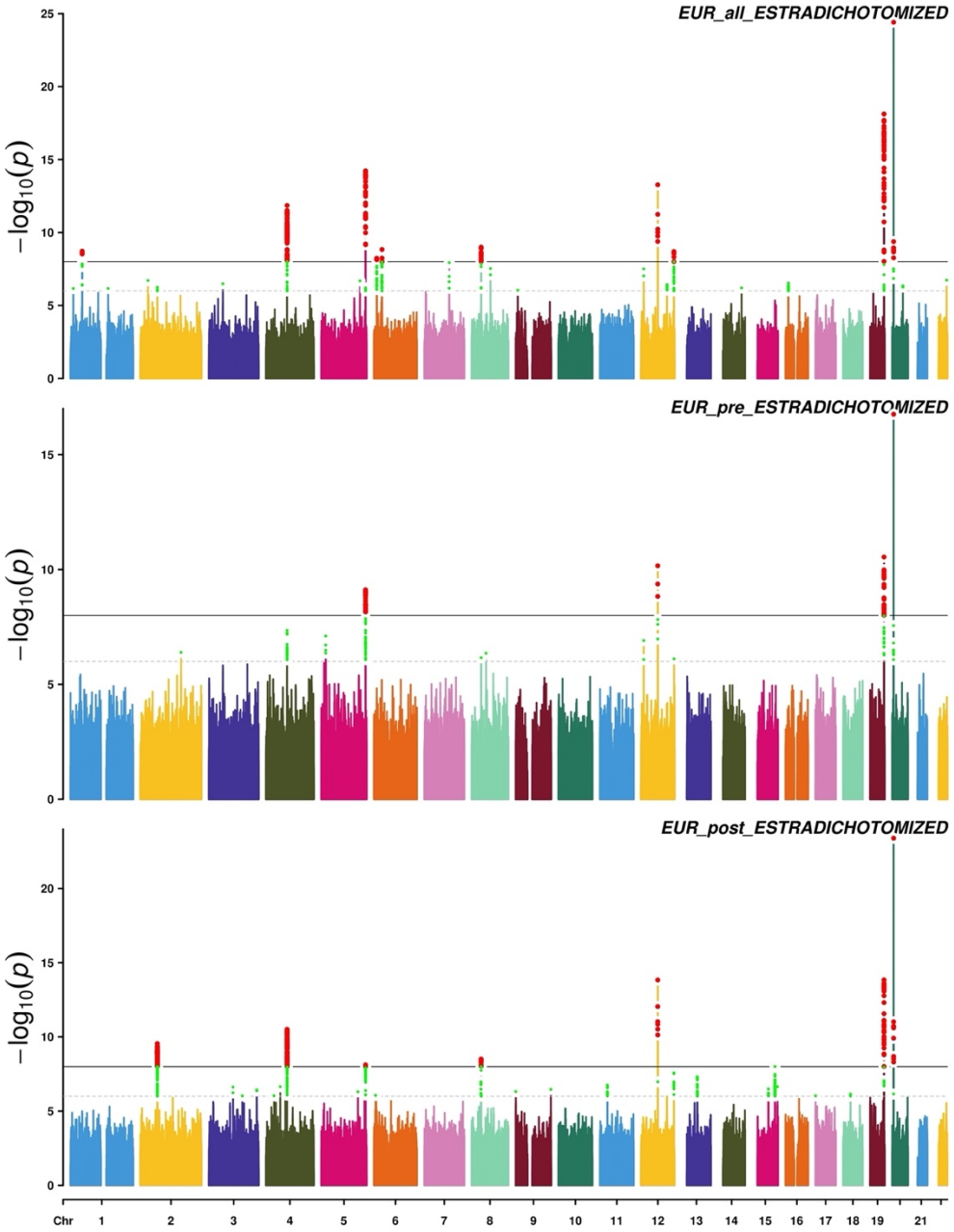


Figure 1.1. Genome-wide analysis of detectable concentrations of estradiol in women of European ancestry overall, and by menopausal status.
Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

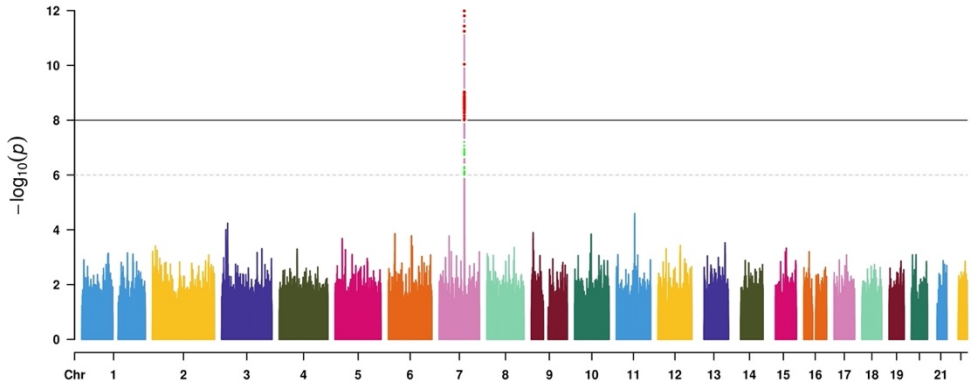


Figure 1.2. Genome-wide analysis of heterogeneity of effect for testosterone in women of European ancestry between premenopausal and postmenopausal. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

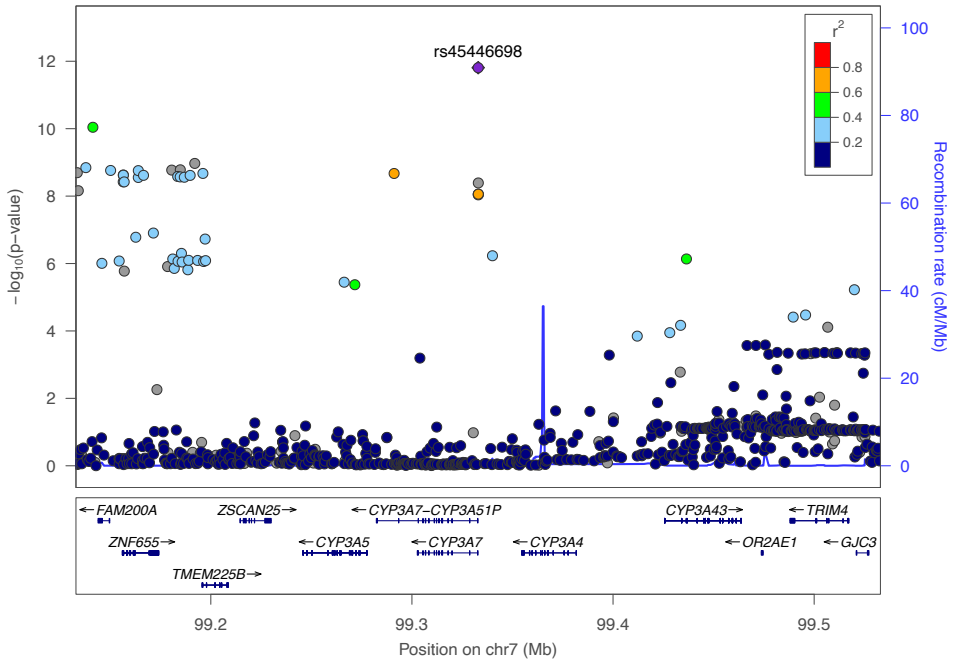


Figure 1.3. Regional association plot of the chromosome 7 region (CYP3A7) showing heterogeneous effects in pre- and postmenopausal concentrations of testosterone.

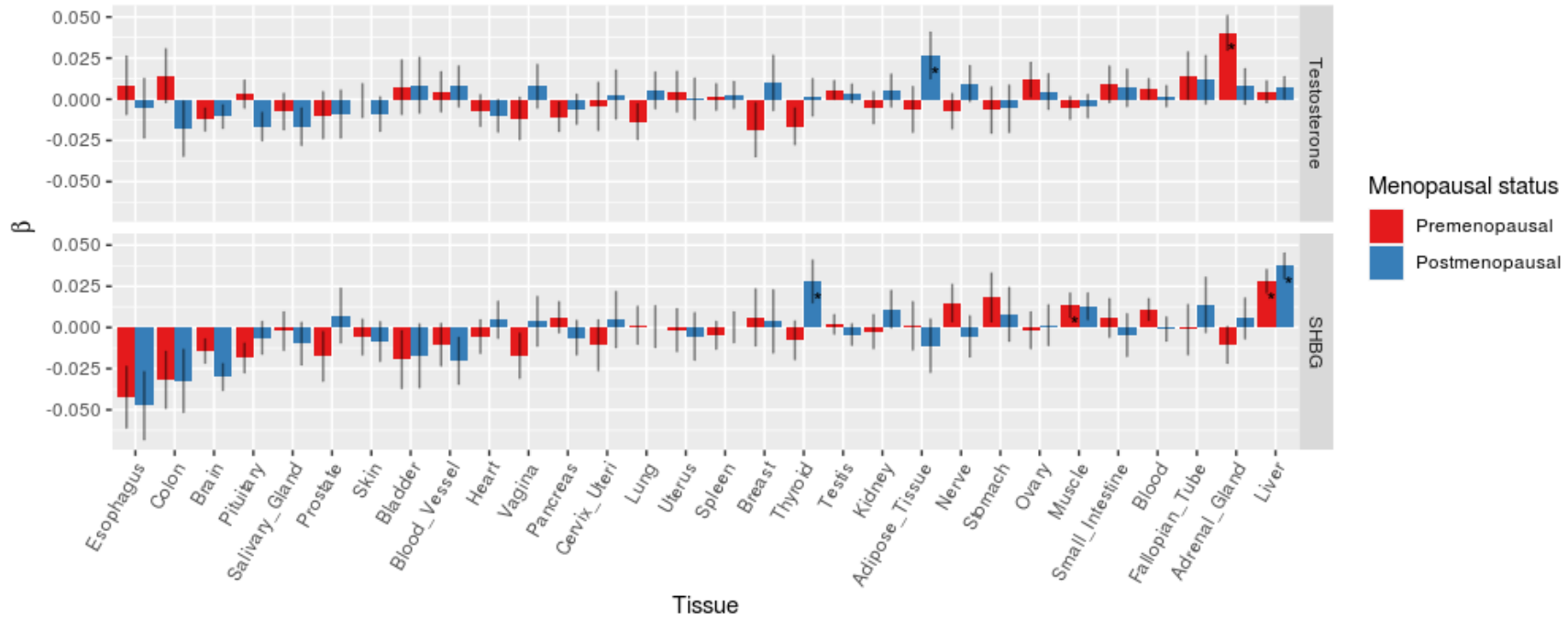


Figure 1.4. Gene property analysis for tissue specificity of testosterone and SHBG in women of European ancestry by menopausal status.

*Statistical significance at p-value<0.05

Table 1.3. Independent significant signals from GWAS of heterogeneity of effect between pre- and postmenopausal women of European ancestry for concentrations of testosterone and SHBG.

					Premenopausal		Postmenopausal	
Phenotype	Lead variant	MAF*	Loci	Nearest gene	Effect estimate (SE)	P-value	Effect estimate (SE)	P-value
Testosterone (ng/dL)	rs45446698	0.04	7q22.1	CYP3A7	0.44 (0.02)	6.3E-133	0.28 (0.01)	1.2E-104
SHBG (ng/dL)	rs3812617	0.16	10p15.1	AKR1C4	-0.10 (0.009)	4.3E-26	-0.02 (0.006)	2.4E-4

*MAF=Minor allele frequency
**SE=Standard error

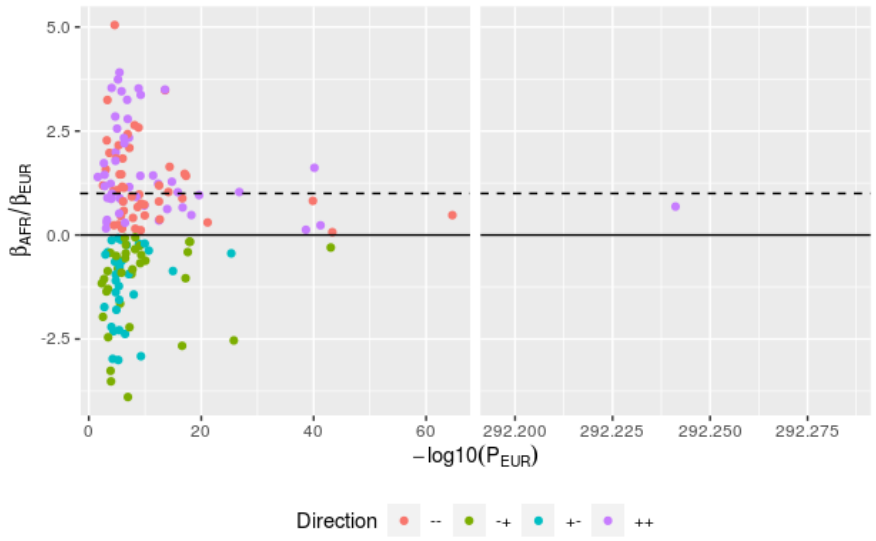


Figure 1.5. African and European cross-ancestry consistency of direction and ratio of size of effect for previously identified variants of circulating testosterone in women overall.

Points below the x axis indicate opposite directions of effect between African and European ancestry populations. The dashed horizontal line indicates the expected ratio of equal effect sizes between ancestry populations. X axis p-value is the meta-analyzed p-value using GWAS's from European and African ancestries in women overall.

Note: The x-axis was segmented because the smallest p-value in the overall ancestry of women of European ancestry obscured the observations for variants with significant but relatively larger p-values.

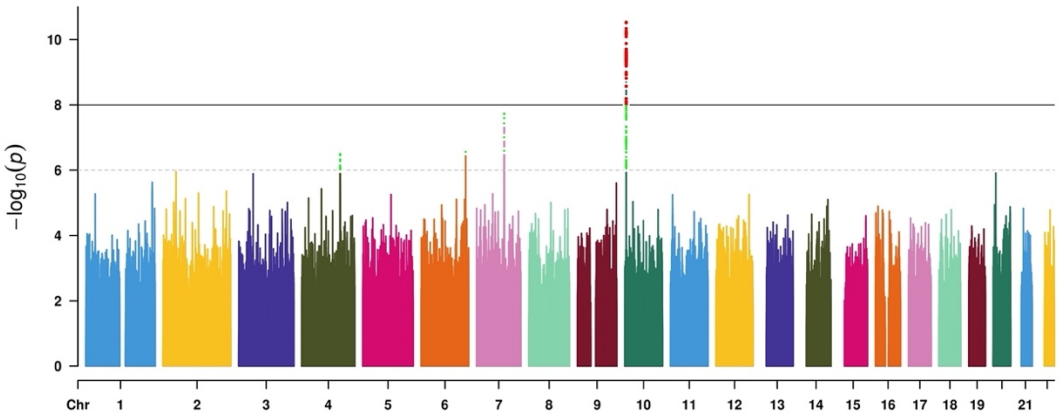


Figure 1.6. Genome-wide analysis of heterogeneity of effect for SHBG in women of European ancestry between premenopausal and postmenopausal.
Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

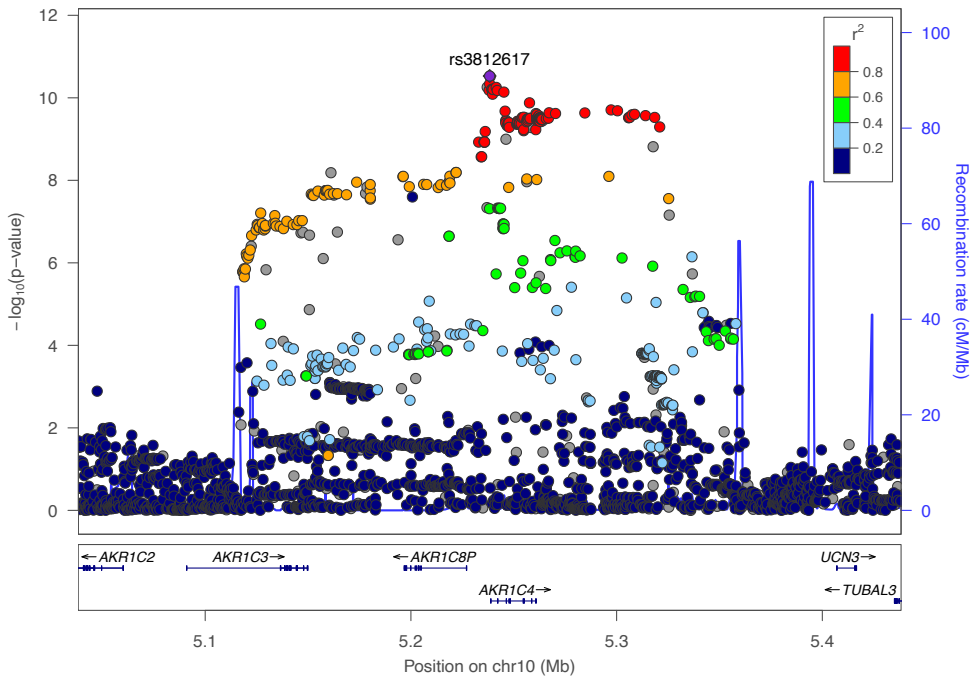


Figure 1.7. Regional association plot of the chromosome 10 region showing heterogeneous effects in pre- and postmenopausal concentrations of SHBG.

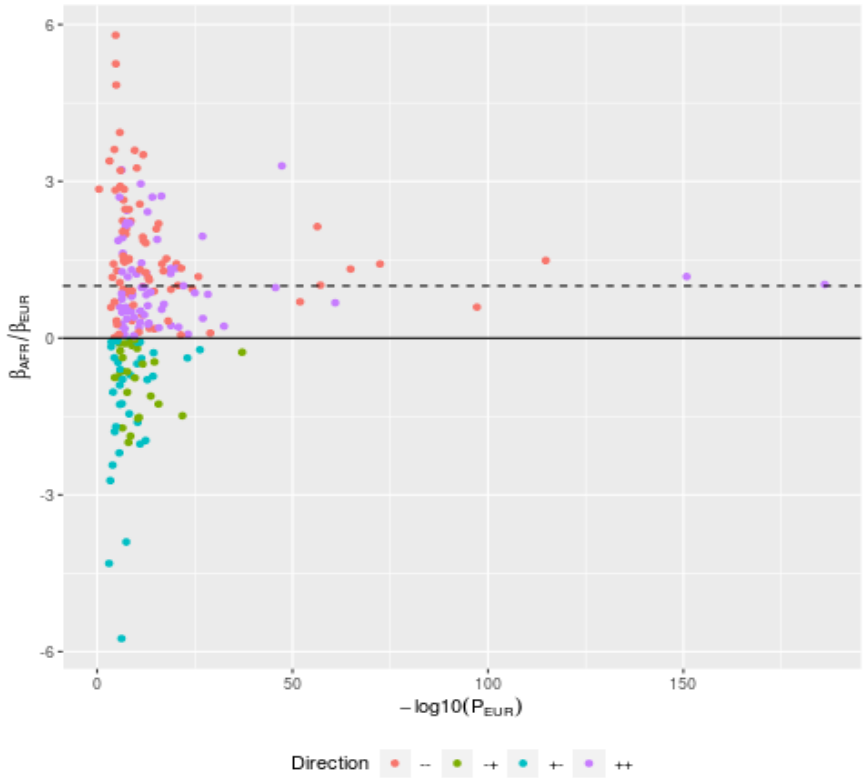


Figure 1.8. Ratio of beta coefficients between African and European GWAS of SHBG in all women by meta-analysis p-value for previous reported variants.

Table 1.4. Genetic correlation between overall and menopausal status specific hormone concentrations among European ancestry women.

Trait	H2 % (s.e.m%)	E2	Pre	Post	E	Pre	T	Pre	Post	SHBG	Pre	Post
Estradichotomized (E2)	1.5 (0.2)	-	0.91 (0.08)	0.48 (0.10)	0.51 (0.32)	0.03 (0.26)	0.09 (0.05)	0.18 (0.08)	-0.003 (0.07)	0.19 (0.05)	0.2 (0.07)	0.14 (0.06)
Premenopausal	3.9 (1.0)		-	0.55 (0.16)	0.09 (0.32)	-0.34 (0.34)	0.05 (0.07)	-0.03 (0.11)	0.006 (0.09)	0.18 (0.07)	0.21 (0.08)	0.16 (0.08)
Postmenopausal	3.5 (0.6)			-	0.45 (0.35)	-0.08 (0.26)	0.21 (0.05)	0.31 (0.08)	0.13 (0.08)	0.08 (0.06)	0.13 (0.08)	0.07 (0.06)
Estradiol (E)	1.1 (1.0)				-	0.86 (0.21)	-0.14 (0.14)	-0.31 (0.28)	-0.17 (0.19)	0.62 (0.31)	0.8 (0.37)	0.68 (0.33)
Premenopausal	1.6 (1.3)					-	-0.05 (0.12)	-0.06 (0.20)	-0.05 (0.17)	0.6 (0.28)	0.76 (0.32)	0.67 (0.31)
Postmenopausal	NA						-	-	-	-	-	-
Testosterone (T)	11 (0.9)						-	0.98 (0.03)	0.99 (0.02)	0.03 (0.04)	-0.07 (0.05)	-0.03 (0.04)
Premenopausal	12 (1.8)							-	0.93 (0.07)	0.08 (0.05)	-0.03 (0.07)	-0.02 (0.06)
Postmenopausal	11 (1.1)								-	-0.002 (0.04)	-0.10 (0.05)	-0.05 (0.05)
SHBG	19 (1.9)									-	0.75 (0.03)	**0.85 (0.02)**
Premenopausal	23 (2.8)										-	**0.89 (0.03)**
Postmenopausal	22 (2.6)											-

Bold correlation estimates indicate Bonferroni statistical significance ($p<0.05/50$)

Disentangling the relationship of body mass index, menopausal status, and sex hormones in mammographic density using Mendelian Randomization

Abstract

Background: Mammographic density, composed of dense and non-dense area, is strongly determined by body mass index (BMI) and a strong predictor of breast cancer. The effects of circulating concentrations of sex hormones on dense and non-dense area remain to be determined. In this study, we investigate the relationship between sex hormones, BMI, and dense and non-dense areas of the breast.
Methods: We used Mendelian Randomization (MR) approaches with instruments created from large European ancestry-based genome-wide association studies (GWAS) for circulating concentrations of sex hormones, BMI, and breast density phenotypes. We applied multiple MR approaches including inverse variance weighted (IVW), multivariable, weighted median, and MR-Egger regression to assure robustness of our results.
Results: Increased genetically predicted testosterone was associated with decreased dense area (IVW: b=-0.22; 95% CI=-0.38, -0.053; p=0.009), and this association was also significant in multivariable MR adjusting for BMI. We observed an association between increasing genetically predicted estradiol concentrations and decreasing non-dense area (MR-Egger: b=-3.32; 95% CI –5.83, -0.82; p= 0.009), although this association was not supported by other MR approaches. Genetically predicted BMI was associated with increased non-dense area (IVW: b=1.79; 95% CI=1.58, 2.00; p=9.57E-63) and decreased dense area (IVW: b = -0.37; 95% CI= -0.51, -0.23; p= 4.7E-7).
Conclusions: We found strong evidence supporting previous observations that BMI is associated with increased non-dense area, and new evidence that it works to also decrease mammographic dense area. In contrast, we observed weak evidence that specific circulating sex hormone concentrations are associated with mammographic density phenotypes.

Introduction

Mammographic density is most often characterized as the proportion of absolute dense area (composed of epithelial and stromal tissue) versus absolute non-dense area (composed of adipose tissue), two independent and oppositely associated risk factors for breast cancer(Assi *et al.*, 2012). Compared to women with primarily fatty breasts (less than 10% epithelial and stromal tissue), those with greater than 75% dense breasts have a 3-6 fold increased risk for breast cancer(Assi *et al.*, 2012). Body mass index (BMI) is a strong predictor of breast density, with a consistent inverse association, as women with higher BMI tend to have lower breast density(Butler *et al.*, 2008; Tehranifar *et al.*, 2011; Nguyen *et al.*, 2013). In contrast, the effects of higher BMI on breast cancer risk are modified by menopausal status, showing a protective effect before menopause but increasing breast cancer risk after(Bremnes *et al.*, 2007; Pettersson *et al.*, 2014; Bertrand *et al.*, 2018; Ooi *et al.*, 2019). One potential explanation for these observations is that because the primary source of endogenous estrogens in postmenopausal women is through the aromatization of androgens in adipose tissue, BMI is a strong positive predictor of circulating estrogen concentrations after menopause (Tamimi *et al.*, 2005), whereas premenopausal estrogen concentrations are primarily driven by ovarian function. Along with estrogen, sex hormone binding globulin (SHBG) and testosterone have been proposed as potential risk factors for breast cancer given their associations with breast cancer risk factors

such as obesity, smoking, and alcohol consumption(Cauley *et al.*, 1999; Key *et al.*, 2002; Tamimi *et al.*, 2005; Hankinson and Eliassen, 2010).

Observational studies of the association between endogenous sex hormones and mammographic density have shown varied results. While some studies have observed no associations between endogenous hormones and breast density, particularly after accounting for BMI(Boyd *et al.*, 2006; Warren *et al.*, 2006), others have found positive associations between SHBG and mammographic density, mixed results for estradiol, and an inverse association with testosterone(N F Boyd *et al.*, 2002; Greendale *et al.*, 2005; Bremnes *et al.*, 2007; Yong *et al.*, 2009; Borugian *et al.*, 2014). Numerous reasons could explain the discrepancies in findings. While most studies analyzed pre- and postmenopausal women separately, sample sizes have been limited and adjusting factors and modeling of parameters have been inconsistent across studies. However, the effects of adiposity are widely recognized among all of the studies, often with attenuating findings in BMI-adjusted models(Sprague *et al.*, 2011).

To get a better understanding of the complex relationship between BMI, circulating hormone concentrations, menopausal status and mammographic density, we developed a direct acyclic graph (DAG) as a guiding framework for this study (**Figure 2.1)**. We then conducted two-sample Mendelian Randomization studies, leveraging genome-wide genotype data from UK Biobank, GIANT and the MODE/BCAC consortia.

Methods

Summary statistics of sex hormones, BMI, and mammographic density

Single nucleotide polymorphism (SNP)-specific beta coefficients and standard errors for concentrations of estradiol, testosterone, and SHBG were estimated using women of European ancestry from the UK Biobank, as previously described(Haas *et al.*, no date; Ruth *et al.*, 2020). Briefly, we used inverse-ranked normalized transformations of hormone concentrations, adjusted for age at reference, study centre, and the first ten principal components. Estradiol was modeled as a binary trait to indicate detectable concentrations (yes/no). We conducted GWAS among all women (excluding women who were pregnant or users of menopausal hormone therapy or oral contraceptives at the time of blood draw) and stratified by menopausal status. SNPs for testosterone and SHBG were selected based on previously published variants reported by Ruth et al.(Ruth *et al.*, 2020) for women, using results from our overall and menopausal status specific analyses subset to overlapping SNPs with p-values below 0.05 in the UK Biobank analyses.

We used published beta coefficients and standard errors for BMI based on findings from the largest meta-GWAS of anthropometric traits from the GIANT consortium, which included nearly 700 genome-wide significant variants associated with BMI (Loh *et al.*, 2015; Pulit *et al.*, 2019; Tang *et al.*, 2021). Estimated effect sizes for breast density phenotypes (absolute dense area and non-dense area) were based on GWAS conducted in the MODE/BCAC consortium among women of European ancestry as previously described, which adjusted for age at mammography and the first 10 principal components (Chen *et al.*, 2021). We conducted additional analyses to obtain pre- and postmenopausal specific effect estimates for dense and non-dense area. Study populations and sample sizes for GWAS summary statistics are presented in **Table 2.1**.

To ensure consistency between the UK Biobank, GIANT and MODE/BCAC data, we removed any palindromic variants (A/T or C/G) as previously recommended (Burgess *et al.*, 2019).

Cross-trait genetics correlation

We estimated the genetic correlation between sex hormone phenotypes and dense area and non-dense area using LD Score Regression (LDSC)(B. K. Bulik-Sullivan, Loh, Finucane, Ripke, Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, *et al.*, 2015; Finucane *et al.*, 2015). We additionally used the LDHub resource to estimate the genetic correlation between sex hormone phenotypes and anthropometric traits, including BMI(Zheng *et al.*, 2017). Heritability estimates for mammographic density phenotypes have been previously reported by Chen et al.(Chen *et al.*, 2021) and BMI as reported by the Neale Lab(Neale, 2019).

Mendelian Randomization

We conducted two-sample MR analyses to estimate the association between circulating concentrations of three sex hormones (estradiol, testosterone, SHBG) overall and by menopausal status and mammographic density traits. We used effect estimates for variants associated with BMI as previously published by Tang et al. to test for the effect of BMI on dense and non-dense area based on summary statistics from MODE, as well as overall and menopausal status specific sex hormone concentrations. We report the results of the inverse-variance weighted (IVW), simple median, weighted median, and MR-Egger approaches as implemented using the MendelianRandomization (Yavorska, Olena; Staley, 2017) package in R version 4.0.1(Team, no date).

We performed additional multivariable MR (MVMR) analyses for which effect estimates for BMI as estimated in UK Biobank GWAS were included to account for possible confounding due to pleiotropy. As an example, when analyzing the association of circulating concentrations of testosterone and dense area, we used variants associated with testosterone among women in the UK Biobank as the primary exposure. For the variants included in our genetic instrument for testosterone, we also extracted their effects on BMI as estimated among women in UK Biobank as the secondary exposure similar to previous studies (Burgess and Thompson, 2015). UK Biobank estimates were used because summary statistics were available for all variants included for instruments of the exposure of interest.

We considered p-values less than 0.05 to be weak evidence of an association, and p-values less than 0.01 to be strong evidence of an association to account for the number of statistical tests performed.

Results

Genetic correlations

We estimated the genetic correlation between three sex hormones concentrations (SHBG, testosterone, and estradiol) overall and within pre- and postmenopausal women only, breast density phenotypes (absolute dense area, and absolute non-dense area), and BMI using LDSC.

Absolute dense area (r_g=0.50, s.e.=0.16; p=0.002) showed strong shared heritability with premenopausal detectable concentrations of estradiol. Premenopausal SHBG was genetically correlated with non-dense area (r_g=-0.22, s.e.=0.08; p=0.008), (**Table 2.2**), whereas testosterone was not genetically correlated with any of the mammographic density phenotypes. BMI showed positive genetic correlation with testosterone (r_g=0.12, s.e.= 0.04; p=2.1E-5) and negative genetic correlation with SHBG (r_g=-0.35, s.e.=0.03; p=2.7E-35).

Effects of sex hormones on breast density

SHBG: Using MR analyses, we did not find any evidence of an effect of SHBG on either dense area or non-dense area based on any of the methods (**Figure 2.2**; **Supplementary Table 2.1**). There was no evidence in either pre- or postmenopausal specific analyses, nor in the multivariable MR estimates adjusting for BMI.
Testosterone: An increase in overall genetically predicted testosterone in women was associated with a decrease in dense area (IVW: β=-0.27; 95% CI=-0.38, -0.053; p=0.009) without evidence of strong directional pleiotropy based on the MR-Egger intercept (p=0.14). A similar association was observed for the simple median, but not statistically significant for the weighted median or MR-Egger approaches (**Supplementary Table 2.1**). While directionally consistent, the strength of the association for within pre- and postmenopausal testosterone concentrations did not reach statistical significance (**Supplementary Table 2.1**). The multivariable MR IVW adjusted for BMI was also significant with a near identical effect estimate as the unadjusted IVW estimate (MVMR IVW for testosterone: β=-0.21; 95% CI=-0.39, -0.03; p=0.02).
Estradiol: Genetically predicted detectable levels of estradiol in women overall showed evidence of an association with decreased non-dense area (MR-Egger: β=-3.32; 95% CI -5.83, -0.82; p= 0.009) and the MR-Egger intercept did suggest directional pleiotropic effects (p= 0.038). When incorporating the effect of BMI, we observed inconsistent findings. While the multivariable MR IVW of genetically predicted detectable levels of estradiol and non-dense area was not significant, the multivariable MR-Egger approach was significant with some evidence of unbalanced pleiotropy (β= -3.16; 95% CI= -5.73, -0.60; p=0.016; intercept p-value=0.049) (**Supplementary Table 2.1**). The premenopausal association between genetically predicted detectable levels of estradiol and non-dense area was statistically significant based on the MR-Egger test with evidence of pleiotropy (MR-Egger: β=-2.03; 95% CI= -3.74, -0.31; p=0.02; intercept p-value=0.039), but this association was not observed by other MR methods or in postmenopausal women.

Effects of BMI on breast density and sex hormones

Genetically predicted BMI was strongly associated with an increase in non-dense area (IVW: β =1.79; 95% CI=1.58, 2.00; p= 9.57E-63) without evidence of directional pleiotropy (MR-Egger intercept: p=0.18) (**Table 2.3**). Increasing genetically predicted BMI was also associated with a decrease in dense area (IVW: β = -0.37; 95% CI= -0.51, -0.23; p= 2.09E-7) without evidence of directional pleiotropy (MR-Egger intercept: p=0.06). We additionally conducted multivariable MR for the effects of genetically predicted BMI on dense and non-dense area with adjustments for genetically predicted SHBG, testosterone, and estradiol each separately but the estimated effects and statistical significance for BMI did not substantively change (**Supplementary Table 2.2**).

Genetically predicted BMI was associated with a decrease in overall SHBG (IVW: β = -0.35; 95% CI= -0.39, -0.32; p= 2.38E-75), premenopausal SHBG (IVW: β = -0.43; 95% CI = -0.49, -0.39; p= 1.20 E-62) and postmenopausal SHBG (IVW: β = -0.33; 95% CI= -0.38, -0.29; p= 1.66E-47). These results remained consistent across all MR approaches. We observed associations for genetically predicted BMI and overall, pre-, and postmenopausal testosterone using IVW (Overall: β = 0.08.; 95% CI =0.05, 0.12; p= 2.87E-7) (**Table 2.3**), without evidence of pleiotropy based on the MR-Egger intercept (p=0.19). Genetically predicted BMI was associated with detectable concentrations of estradiol both overall and among premenopausal women (Premenopausal: β=-0.05; 95% CI = -0.09, -0.015; p=0.006) without evidence of pleiotropy based on the MR-Egger intercept (p=0.97). However, we observed no evidence of an association between genetically predicted BMI and postmenopausal detectable concentrations of estradiol, although the standard errors were large.

Discussion

In this study we leveraged genetic data to investigate the relationships between menopause, concentrations of SHBG, testosterone, estradiol, BMI and breast density phenotypes. We first looked at genetic correlations between the traits of interest and observed significant negative genetic correlations between SHBG and BMI, as well as between premenopausal SHBG and non-dense area, and also a positive genetic correlation between testosterone and BMI. We observed a shared heritability between premenopausal detectable levels of estradiol and percent mammographic density, which is likely driven by the positive genetic correlation between premenopausal detectable levels of estradiol and dense area, and also a non-significant negative genetic correlation with non-dense area.

Numerous observational studies have supported the effects of BMI on non-dense area (Wong *et al.*, 2011; Nguyen *et al.*, 2013; Busana *et al.*, 2016). Recent Mendelian Randomization has also validated the causal effects of BMI on breast size, but did not delineate dense and non-dense composition (Ooi *et al.*, 2019). We here demonstrate that increased genetically predicted BMI is strongly associated with increased non-dense area and decreased dense area. We performed a comprehensive set of analyses to investigate whether BMI confounded any of our MR analysis of sex hormone phenotypes and mammographic density phenotypes but did not find evidence of such.

The effects of BMI on circulating concentrations of sex hormones have important implications for understanding their potential involvement in breast cancer development. As expected, higher genetically predicted BMI was associated with lower SHBG and higher testosterone concentrations. Our results further suggest that BMI might have a larger effect on testosterone concentrations in postmenopausal women, as compared to premenopausal women. This observation is consistent with previous work that has found greater tissue specificity of postmenopausal testosterone in adipose tissue gene expression while premenopausal testosterone had greater adrenal gland tissue specificity (Haas *et al.*, no date). We also observed associations between genetically predicted increased BMI and lower overall and premenopausal detectable levels of estradiol, but this was not observed for postmenopausal estradiol. These observations can help explain the paradoxical effects of BMI on breast cancer risk. Estradiol has long been suspected to cause hormone receptor positive breast cancers(Arslan *et al.*, 2009; Quan *et al.*, 2014; Shawky *et al.*, 2017). A previous MR study of BMI on breast cancer demonstrated an inverse association across both pre- and postmenopausal women (Guo *et al.*, 2016; Shu *et al.*,

2019). Our findings suggest that BMI may work to reduce breast cancer risk through decreased concentrations of estradiol, particularly in premenopausal women, and this mechanism does not necessarily work through modifying breast density. Our results also suggest that the observed reduction in estradiol caused by increased BMI may not occur after menopause, resulting in the loss of the protective effects of BMI for breast cancer.

We used an MR approach to test the associations between circulating concentrations of sex hormones and absolute dense area and absolute non-dense area and found no strong evidence of directional associations. Previous work has suggested a mediating role of estradiol between adiposity and breast cancer (Key *et al.*, 2002; Warren *et al.*, 2006; Baglietto *et al.*, 2010; Endogenous Hormones and Breast Cancer Collaborative Group *et al.*, 2011). Much of previous work has been conducted exclusively among postmenopausal women, limiting the generalizability to premenopausal women. In this study, we observed evidence of a negative association between genetically predicted estradiol concentrations in premenopausal women and non-dense area, suggesting that estradiol may play a role in decreasing non-dense area independent of any effect on dense area or BMI as observed by the multivariable MR approaches. Previous studies investigating the associations of estrogen-related genetic variation and mammographic density relied on candidate genes and showed no association (Dumas and Diorio, 2010). Our approach for estimating the effects of BMI on sex hormones may also violate some assumptions of MR, namely the independence between gene-exposure and gene-outcome association due to sample overlap as all women in the UK Biobank were included in the BMI estimates, but were only approximately a quarter of the total sample population. However, recent work has shown most MR approaches to be unbiased in this setting except for MR-Egger, in which case we interpret those estimates with caution(Burgess, Davies and Thompson, 2016).

While recent releases of the UK Biobank biomarker data and large-scale GWAS on circulating concentrations of sex hormones have strengthened the instruments available for conducting MR analyses, we recognize limitations in our ability to detect causal associations. We were particularly limited by our instrument for detectable concentrations of estradiol due to the low heritability and potential bias due to any menopausal effects not captured by our binary classification for menopausal status. Our previous work identified only one locus associated with estradiol as a continuous trait. Since postmenopausal women in the UK Biobank had a disproportionately high proportion of non-detectable concentrations of estradiol we used a binary phenotype to indicate detectable levels (Haas *et al.*, no date). We also relied on circulated plasma concentrations of hormones rather than possibly more etiological relevant tissue specific measures, as plasma concentrations may not reflect breast tissue exposure to the hormones of interest. However, previous work has shown significant correlation between plasma estradiol and benign tissue estradiol concentrations in both pre- and postmenopausal women (Stanczyk, Mathews and Sherman, 2015).

We leveraged large datasets to delineate the genetic architectures of pre- from postmenopausal concentrations of sex hormones to develop instruments for our MR analyses, furthering the body of research investigating the differences in breast density and breast cancer risk factors by menopausal status. We found evidence that genetically predicted BMI not only increases non-dense area but is also associated with decreased dense area, an additional benefit that could help explain the lower risk of breast cancer observed among premenopausal women with high BMI, and this risk decrease may be mediated by decreased estradiol concentrations in premenopausal but not postmenopausal women.

Tables and Figures

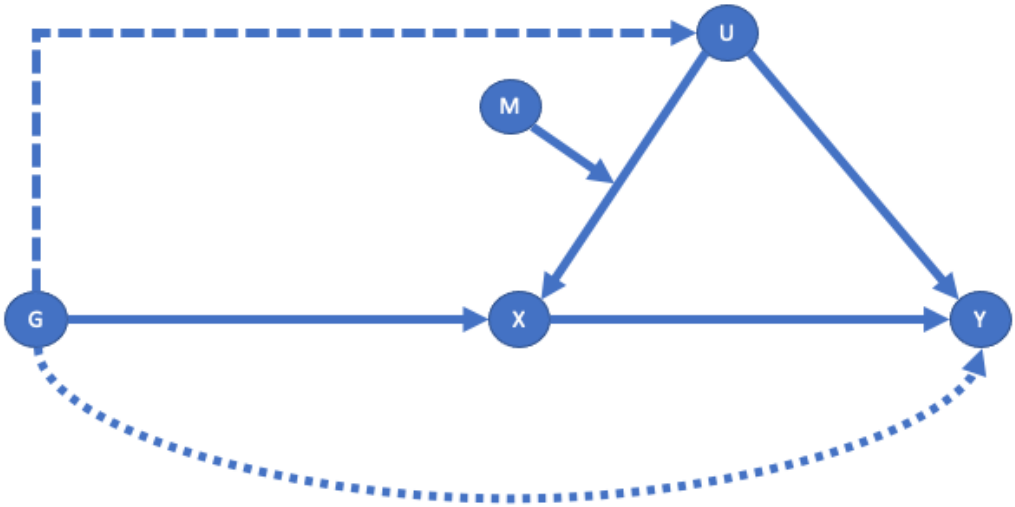


Figure 2.1. Directed acyclic graph for the genetic instrumental variable analysis of the association between sex hormones (X) and breast density traits (Y).
X=Sex hormone phenotypes (i.e. estradiol, testosterone, SHBG)
Y=Breast density phenotypes (i.e. dense area and non-dense area)
G=Genetic instrument for sex hormone phenotypes
U=Confounders and pleiotropic effects (e.g. BMI)
M=Effect modifiers of confounders and pleiotropic effects (e.g. menopause)

Table 2.1. Summary of GWAS's used for MR instrument effect estimates.

Trait		Study Population	N
SHBG			196,901
	Premenopausal	UK Biobank	43,477
	Postmenopausal		92,911
Testosterone			182,648
	Premenopausal	UK Biobank	44,401
	Postmenopausal		84,003
Detectable Estradiol			229,966
	Premenopausal	UK Biobank	51,081
	Postmenopausal		84,194
DA/NDA			14,036
	Premenopausal	BCAC/MODE	4,464
	Postmenopausal		9,245
BMI		UK Biobank/GIANT	806,810

Table 2.2. Genetic correlations between circulating hormone concentrations, mammographic density phenotypes and BMI among European ancestry women, overall and by menopausal status.

			DA	NDA	BMI
Trait		H^2 (s.e.m)	0.32 (0.04)	0.24 (0.03)	0.25 (0.0008)
SHBG		0.19 (1.9)	0.05 (0.06) p=0.42	-0.06 (0.06) p=0.29	**-0.35 (0.03) p=2.7E-35**
	Premenopausal	0.23 (2.8)	0.06 (0.08) p=0.50	**-0.22 (0.08) p=0.008**	**-0.34 (0.038) p=2.7E-19**
	Postmenopausal	0.22 (2.6)	0.07 (0.06) p=0.28	-0.08 (0.06) p=0.24	**-0.29 (0.03) p=5.6E-20**
Testosterone		0.11 (0.9)	-0.02 (0.06) p=0.78	-0.02 (0.06) p=0.66	**0.12 (0.04) p=2.1E-5**
	Premenopausal	0.12 (1.8)	-0.13 (0.10) p=0.20	0.03 (0.10) p=0.78	0.14 (0.05) p=0.003
	Postmenopausal	0.11 (1.1)	0.12 (0.08) p=0.13	-0.03 (0.07) p=0.68	**0.13 (0.03) p=2.0E-4**
Detectable levels of estradiol		0.015 (0.02)	0.23 (0.12) p=0.047	-0.16 (0.11) p=0.17	-0.040 (0.06) p=0.45
	Premenopausal	0.039 (0.01)	**0.50 (0.16) p=0.002**	-0.21 (0.15) p=0.14	-0.090 (0.076) p=0.24
	Postmenopausal	0.035 (0.06)	-0.02 (0.11) p=0.89	0.04 (0.11) p=0.70	0.075 (0.062) p=0.22

Bold correlation estimates indicate Bonferroni statistical significance ($p<0.05/48$)

PMD= Percent mammographic density, DA= dense area, NDA=non-dense area, BMI=body-mass index

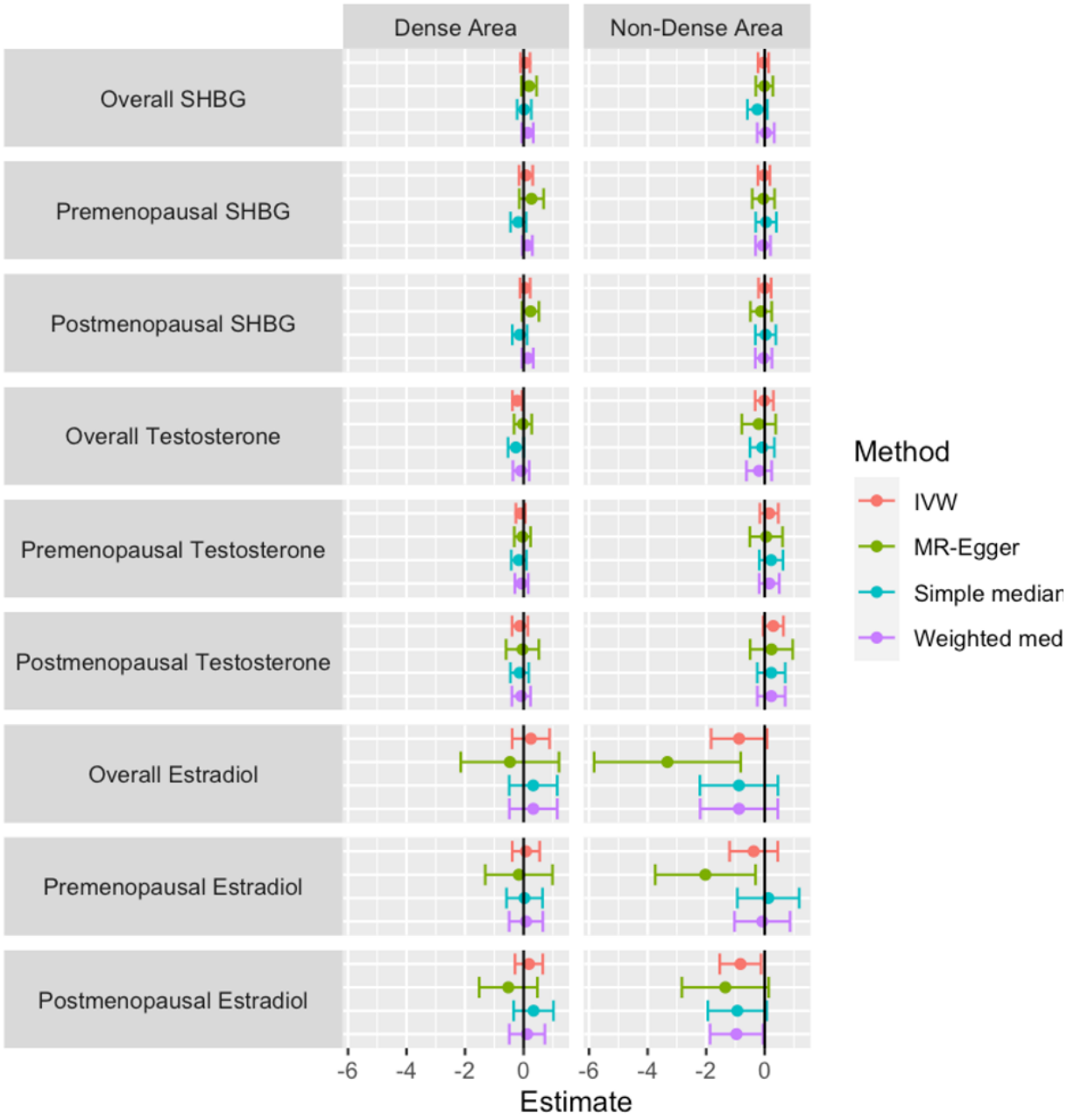


Figure 2.2. Estimated effect and 95% confidence intervals (CI) for the association between overall, pre- and postmenopausal specific sex hormone concentrations and two breast density phenotypes (absolute dense area and absolute non-dense area) based on the different Mendelian randomization approaches used in this study.

Table 2.3. Mendelian Randomization effect estimates of BMI on breast density phenotypes and overall, pre- and postmenopausal specific sex hormones based on inverse-variance weighted and MR-Egger approaches.

Outcome	IVW Estimate	95% CI	P-value	MR-Egger Estimate	95% CI	P-value
Dense Area	-0.37	(-0.51, -0.23)	2.09E-07	-0.70	(-1.08, -0.33)	2.6E-04
Non-Dense Area	1.79	(1.58, 2.00)	9.56E-63	2.15	(1.59, 2.72)	8.02E-14
Overall SHBG	-0.35	(-0.39, -0.31)	2.38E-75	-0.36	(-0.46, -0.26)	2.25E-12
Premenopausal SHBG	-0.44	(-0.49, -0.39)	1.21E-62	-0.54	(-0.68, -0.41)	9.28E-14
Postmenopausal SHBG	-0.33	(-0.38, -0.29)	1.66E-47	-0.30	(-0.41, -0.18)	4.32E-07
Overall Testosterone	0.08	(0.05, 0.12)	2.87E-07	0.03	(-0.06, 0.12)	0.51
Premenopausal Testosterone	0.13	(0.07, 0.18)	1.88E-08	0.04	(-0.10, 0.18)	0.61
Postmenopausal Testosterone	0.08	(0.05, 0.12)	1.59E-05	0.08	(-0.02, 0.11)	0.11
Overall Estradiol	-0.02	(-0.04, 0.00)	0.03	-0.06	(-0.12, -0.02)	0.01
Premenopausal Estradiol	-0.05	(-0.09, -0.02)	0.006	-0.05	(-0.18, 0.02)	0.33
Postmenopausal Estradiol	0.003	(-0.03, 0.04)	0.83	0.003	(-0.08, 0.09)	0.94

Genome-wide gene-environment interaction study of body mass index and mammographic density phenotypes

Abstract

Background: Body mass index (BMI) is the strongest predictor of mammographic density, in turn one of the strongest predictors of breast cancer in women. We performed genome-wide gene-environment (G-E) analyses of BMI on percent mammographic density (PMD), dense area (DA), and non-dense (NDA).
Methods: We used samples from the Marker of Density (MODE) consortium and the Breast Cancer Association Consortium (BCAC) with both genetic and phenotypic data to test the interaction effect estimate between SNPs and BMI on PD, DA, and NDA.
Results: Nearly 15,000 women were included in our G-E analyses for BMI and mammographic density phenotypes. No SNPs reach Bonferroni correction for statistical significance. One locus for PMD did reach the suggestive threshold: locus 7q33 near the *AKR1B1* gene, although there was evidence of greater genomic inflation when including variants with minor allele frequencies less than 5% for NDA.
Conclusions: Larger studies are merited to provide sufficient power to further valid suggestive findings of an interaction between SNPs in the *ARK1B1* gene and PMD.

Introduction

Mammographic density describes the proportions of epithelial and stromal, and fat tissue in the breast, and is one of the strongest predictors of breast cancer risk in women(Assi et al., 2012). Almost 5% of the female population has extremely dense breasts (percent dense area ≥75%) and, compared to women with primarily fatty breasts (<10% dense tissue), these women have a three- to six-fold increase in risk(Assi et al., 2012),(Byrne et al., 1995). As a strong risk factor for breast cancer, mammographic density serves as an appropriate endpoint and target for reducing the risk of breast cancer in women(Yaffe et al., 2005). Up to 67% of the variation in mammographic density has been estimated to be due to genetic factors in twin studies and it has shown to have a shared genetic bases with breast cancer(Norman F Boyd et al., 2002; Stevens et al., 2012; Vachon et al., 2012; Lindström et al., 2014; Stone et al., 2015). Based on GWAS data, the estimated variation of percent mammographic density explained by common genetic variants is 27%(Chen et al., 2021), leaving a substantial amount of heritability left to be explained.

Body mass index (BMI) is one of the strongest predictors of mammographic density, with a consistent inverse association, as women with higher BMI tend to have lower mammographic density(Butler et al., 2008; Tehranifar et al., 2011; Nguyen et al., 2013). In contrast, the effects of higher BMI on breast cancer risk are modified by menopausal status, showing a protective effect before menopause but increasing breast cancer risk after(Bremnes et al., 2007; Pettersson et al., 2014; Bertrand et al., 2018; Ooi et al., 2019). One hypothesis for these observations is that because the primary source of endogenous estrogens in postmenopausal women is through the aromatization of androgens in adipose tissue, BMI is a strong positive predictor of circulating estrogen levels(Tamimi et al., 2005).

Performing gene-environment (G-E) interactions studies may contribute to our understanding of the underlying genetic architecture contributing to variation in mammographic density phenotypes. The underlying motivation for conducting G-E interaction analysis is to identify SNPs for which the association with the outcome depends on the strata of an environmental variable. For mammographic density and breast cancer, it is likely that there exist SNPs which are only associated with the outcome within specific environmental strata and would not reach genome-wide significance in traditional marginal GWAS, particularly if the relevant stratum is uncommon(Travis et al., 2010; Nickels et al., 2013). G-E interactions can provide new insights into causal pathways to mammographic density and potentially point to novel biological targets that could diminish the effects of mammographic density on breast cancer risk(Rudolph, Chang-Claude and Schmidt, 2016).

To our knowledge, no previous studies have analyzed genome-wide G-E interactions on mammographic density, likely due to limited sample sizes and access to genome-wide data. With genetic, epidemiologic, and mammographic density data for 15,000 women, we conducted the largest G-E interaction analyses for BMI and mammographic density among women of European ancestry to date.

Methods

Study design and phenotype data

The BCAC/MODE consortia are a cross-study collaborative effort that includes more than 25,000 women of European ancestry with both GWAS and mammographic density data. We also leveraged data on an additional ~2,000 women from the Mayo Mammography Healthy Study (MMHS) which were included in our population following MODE protocols(Olson et al., 2012). We excluded studies with fewer than 150 samples with phenotype data.

Mammographic density phenotypes were ascertained using Cumulus measures, a computer-assisted thresholding method for estimating mammographic density(Byng et al., 1994, 1998; Boyd et al., 1995; Ciatto et al., 2012). Briefly, the reader uses a sliding scale to outline the breast edge and then the dense area based on pixel brightness. We performed GWAS of three mammographic density phenotypes: 1) Percent mammographic density (i.e., the proportion of dense area relative to the total area of the breast), 2) absolute dense area (cm^2), and 3) absolute non-dense area (cm^2). Our analyses included 19,543 women, among which there are 6,492 pre/perimenopausal women and 13,051 postmenopausal women. Age and BMI (kg/m^2) at the time of mammogram collection were included as covariates in the GWAS. For the participants with missing BMI at mammogram (N=1,767), self-reported BMI within five years of mammogram collection was used as an approximate.

Genotyping, imputation, and quality control

Genotyping was generated with the iCOGs(Bahcall, 2019) and OncoArray(Christopher I. Amos *et al.*, 2017) platforms manufactured by Illumina, depending on the study and sample. OncoArray data was preferentially included for samples genotyped on both arrays. Standard quality control for genotype missingness, Hardy-Weinberg equilibrium, and heterozygosity were applied prior to imputation. Genotype dosage was generated from imputation using 1000 Genomes phase 3 version 5(1000 Genomes Project Consortium *et al.*, 2015) using IMPUTE2(Howie *et al.*, 2012). Results were filtered to variants with imputation quality greater than 0.8 and minor allele frequency greater than 5%.

Statistical analysis

We first conducted meta-analyses across studies with at least 100 samples with BMI information for the effects of BMI on PMD, DA, and NDA using the meta package(Schwarzer, Carpenter and Rücker, 2015) in R (version 4.0.1), adjusting for age at the time of mammography. We conducted genome-wide interaction analyses to test the interaction between imputed dosage and BMI on three mammographic density phenotypes (PMD, DA, NDA) separately for samples genotyped on OncoArray and iCOGs. We tested for interaction by including a SNPxBMI product term in a linear regression analysis. We used within-study normalized transformations of the square root of each outcome. We used within-study quantiles for BMI as an ordered factor in regression analyses. All analyses were adjusted for age at mammography and the first 10 principal components. We pooled data from iCOGS and OncoArray through meta-analyses using METAL(Willer, Li and Abecasis, 2010).

Results

Final sample sizes were 14,837, 11,516, and 11,586 for PMD, DA, and NDA, respectively. We present the fixed effect and random effects estimates for the association of BMI and the

mammographic density phenotypes of interest (**Figure 3.1**). The random effects estimate confirmed an association between BMI and absolute dense and non-dense area ($\beta = -0.036$, 95% CI = -0.05, -0.02; $\beta = 0.36$, 95% CI = 0.32, 0.40, respectively). Between the two arrays, 4,017,148 variants were included after filtering to variants with MAF>5% and at least 10,000 samples genotyped. No variant reached standard genome-wide significance $5x10^{-8}$ or a suggestive threshold of p-value $1x10^{-6}$ for PMD, DA, or NDA (**Figure 3.2**; **Figure 3.3**).

Discussion

We did not discover any variants that reached genome-wide significance for interacting with BMI to alter any of the mammographic density phenotypes. Despite the largest sample size available to date, we were likely still underpowered to detect such variants, as G-E studies require even larger sample sizes than GWAS(Luan et al., 2001). We used within-study quantiles of BMI in order to reduce the influence of extreme values of BMI in our model. However, alternative modeling of BMI may be warranted to detect an association. Our quality control thresholds based on within study sample sizes were conservative in order to reduce genomic inflation, although we saw little difference when using a more lenient MAF of 1% when comparing lambdas, particularly for PMD and DA.

Genetic interactions with BMI on mammographic density are important to understand in order to achieve the aims of personalized medicine. Future work to increase sample sizes for G-E studies of this relationship should be considered.

Tables and Figures

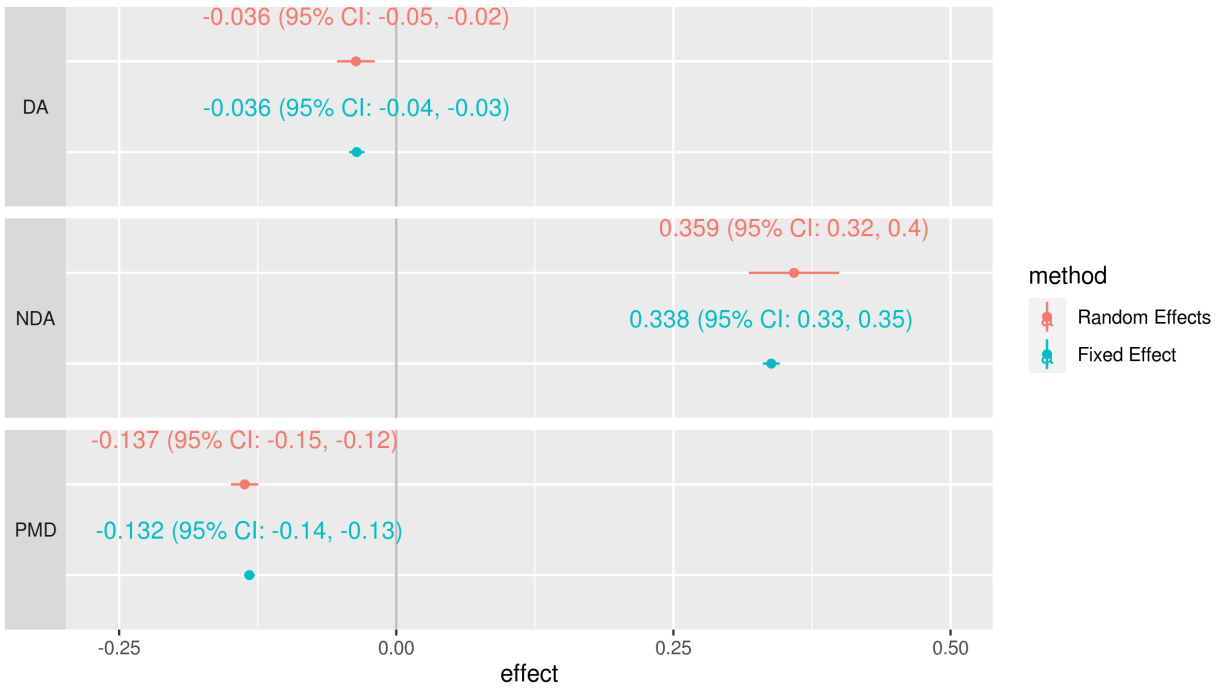


Figure 3. Meta-analysis summary estimates of the effect of BMI on mammographic density phenotypes in the BCAC/MODE population.

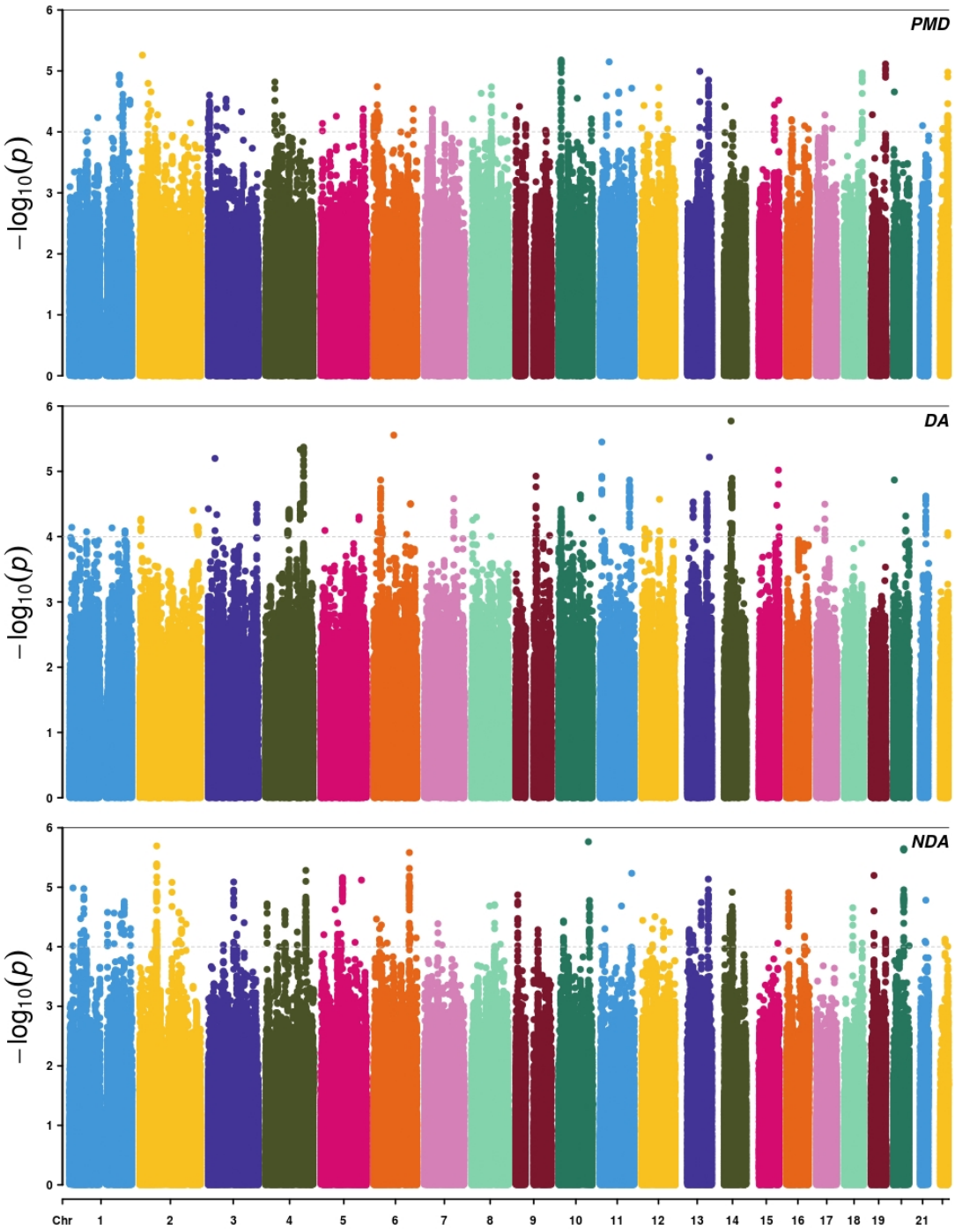


Figure 3.4. Manhattan plot of genome-wide interaction analysis for body mass index and percent mammographic density (PMD), dense area (DA), and non-dense area (NDA).

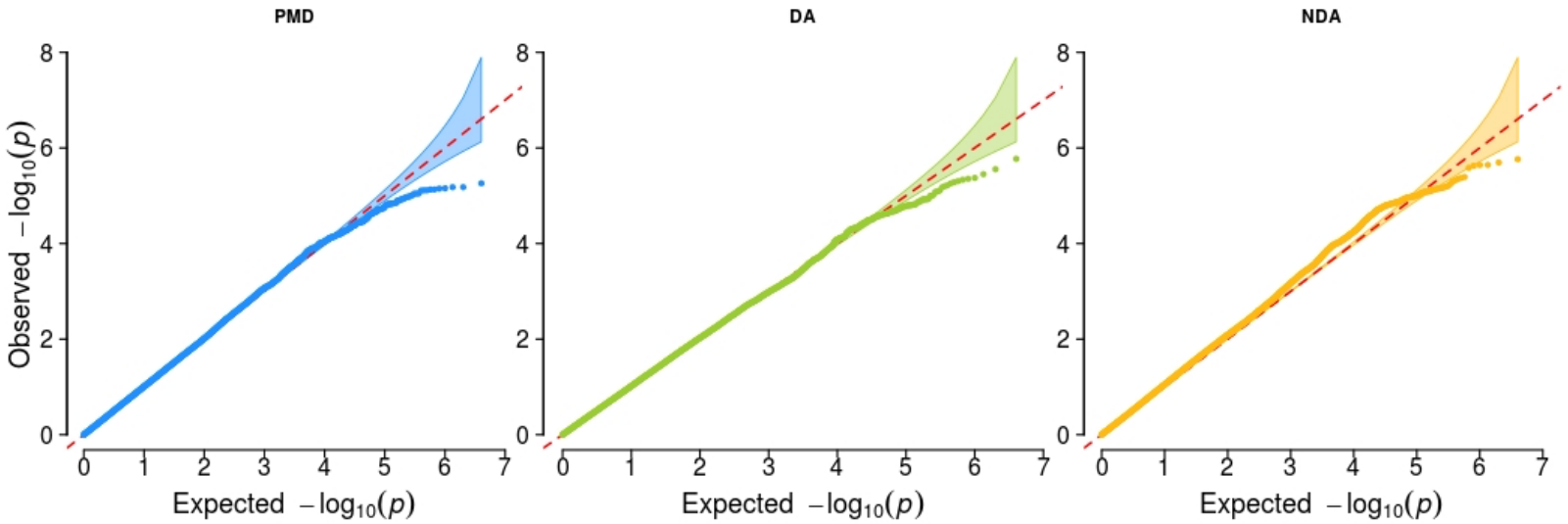


Figure 3.5. QQ-plot of genome-wide interaction analyses of body mass index and percent mammographic density (PMD), dense area (DA), and non-dense area (NDA).

Conclusions

Mammographic density is a strong predictor of breast cancer in women, but the mechanisms through which this risk is conferred is unclear. This work investigated the complex relationships between BMI and sex hormones with mammographic density and differences across menopause. Our findings suggest that BMI and adipose tissue may play unique roles in determining both estradiol and testosterone concentrations in premenopausal versus postmenopausal women. Using genome-wide analyses we found loci with significantly attenuated effects after menopause on concentrations of testosterone and SHBG. We leveraged these findings to show a change from adrenal gland tissue-specificity for SNPs associated with testosterone before menopause to greater adipose tissue specificity after menopause. Our Mendelian Randomization analyses show a positive association between genetically predicted BMI and testosterone in both pre- and postmenopausal women. As a risk factor for numerous diseases, future research is merited to understand how modifying excess adiposity may help regulate testosterone differentially across menopause.

We present evidence that the previously observed protective effects of BMI on breast cancer may work through its reduction on estradiol in premenopausal women, which is not sustained after menopause based on our Mendelian Randomization approach. This would be consistent with previous findings of preventative effects of BMI on hormone receptor positive forms of breast cancer. We also posit that BMI not only increases percent mammographic density through a positive effect on absolute non-dense area, but also independently decreases absolute dense area. Although we did not find loci that interact with BMI to modify mammographic density, further research with larger sample sizes is merited as such findings could be leveraged to identify women for which targeted interventions to alter excess adiposity may be more beneficial for preventing certain subtypes of breast cancer, namely those cancers related to mammographic density.

Breast cancer is a heterogeneous disease with numerous pathways. Our findings help to elucidate possible mechanisms through, and independently of, mammographic density to help develop preventative strategies for breast cancer in women.

References

1000 Genomes Project Consortium *et al*. (2015) 'A global reference for human genetic variation.', *Nature*, 526(7571), pp. 68–74. doi: 10.1038/nature15393.

Amos, Christopher I *et al*. (2017) 'The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 26(1), pp. 126–135. doi: 10.1158/1055-9965.EPI-16-0106.

Amos, Christopher I. *et al*. (2017) 'The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers', *Cancer Epidemiology Biomarkers & Prevention*, 26(1), pp. 126–135. doi: 10.1158/1055-9965.EPI-16-0106.

Arslan, A. A. *et al*. (2009) 'Circulating estrogen metabolites and risk for breast cancer in premenopausal women.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive*

Oncology, 18(8), pp. 2273–9. doi: 10.1158/1055-9965.EPI-09-0312.
Assi, V. *et al*. (2012) 'Clinical and epidemiological issues in mammographic density', *Nature Reviews Clinical Oncology*, 9(1), pp. 33–40. doi: 10.1038/nrclinonc.2011.173.
Baglietto, L. *et al*. (2010) 'Circulating Steroid Hormone Levels and Risk of Breast Cancer for Postmenopausal Women', *Cancer Epidemiology Biomarkers & Prevention*, 19(2), pp. 492–502. doi: 10.1158/1055-9965.EPI-09-0532.
Bahcall, O. (2019) 'COGS project and design of the iCOGS array', *Nature Genetics*. doi: 10.1038/ngicogs.4.
Bertrand, K. A. *et al*. (2018) 'Circulating Hormones and Mammographic Density in Premenopausal Women.', *Hormones & cancer*, 9(2), pp. 117–127. doi: 10.1007/s12672-017-0321-6.
Borugian, M. J. *et al*. (2014) 'Fasting insulin and endogenous hormones in relation to premenopausal breast density (Canada).', *Cancer causes & control : CCC*, 25(3), pp. 385–94. doi: 10.1007/s10552-014-0339-9.
Boyd, N. F. *et al*. (1995) 'Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study.', *Journal of the National Cancer Institute*, 87(9), pp. 670–5. doi: 10.1093/jnci/87.9.670.
Boyd, Norman F *et al*. (2002) 'Heritability of mammographic density, a risk factor for breast cancer.', *The New England journal of medicine*, 347(12), pp. 886–94. doi: 10.1056/NEJMoa013390.
Boyd, N F *et al*. (2002) 'The association of breast mitogens with mammographic densities.', *British journal of cancer*, 87(8), pp. 876–82. doi: 10.1038/sj.bjc.6600537.
Boyd, N. F. *et al*. (2006) 'Mammographic density: a hormonally responsive risk factor for breast cancer.', *The journal of the British Menopause Society*, 12(4), pp. 186–93. doi: 10.1258/136218006779160436.
Bremnes, Y. *et al*. (2007) 'Endogenous sex hormones, prolactin and mammographic density in postmenopausal Norwegian women.', *International journal of cancer*, 121(11), pp. 2506–11. doi: 10.1002/ijc.22971.
Bulik-Sullivan, B. *et al*. (2015) 'An atlas of genetic correlations across human diseases and traits.', *Nature genetics*, 47(11), pp. 1236–41. doi: 10.1038/ng.3406.
Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, *et al*. (2015) 'LD Score regression distinguishes confounding from polygenicity in genome-wide association studies.', *Nature genetics*, 47(3), pp. 291–5. doi: 10.1038/ng.3211.
Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., *et al*. (2015) 'LD Score regression distinguishes confounding from polygenicity in genome-wide association studies', *Nature Genetics*, 47(3), pp. 291–295. doi: 10.1038/ng.3211.
Burger, H. G. *et al*. (2007) 'A review of hormonal changes during the menopausal transition: focus on findings from the Melbourne Women's Midlife Health Project', *Human Reproduction Update*, 13(6), pp. 559–565. doi: 10.1093/humupd/dmm020.
Burgess, S. *et al*. (2019) 'Guidelines for performing Mendelian randomization investigations', *Wellcome Open Research*, 4, p. 186. doi: 10.12688/wellcomeopenres.15555.1.
Burgess, S., Davies, N. M. and Thompson, S. G. (2016) 'Bias due to participant overlap in two-sample Mendelian randomization', *Genetic Epidemiology*, 40(7). doi: 10.1002/gepi.21998.
Burgess, S. and Thompson, S. G. (2015) 'Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects.', *American journal of epidemiology*,

181(4), pp. 251–60. doi: 10.1093/aje/kwu283.
Busana, M. C. *et al*. (2016) 'Impact of type of full-field digital image on mammographic density assessment and breast cancer risk estimation: a case-control study.', *Breast cancer research : BCR*, 18(1), p. 96. doi: 10.1186/s13058-016-0756-7.
Butler, L. M. *et al*. (2008) 'Menstrual and reproductive factors in relation to mammographic density: the Study of Women's Health Across the Nation (SWAN).', *Breast cancer research and treatment*, 112(1), pp. 165–74. doi: 10.1007/s10549-007-9840-0.
Bycroft, C. *et al*. (2017) 'Genome-wide genetic data on ~500,000 UK Biobank participants', *bioRxiv*, p. 166298. doi: 10.1101/166298.
Bycroft, C. *et al*. (2018) 'The UK Biobank resource with deep phenotyping and genomic data.', *Nature*, 562(7726), pp. 203–209. doi: 10.1038/s41586-018-0579-z.
Byng, J. W. *et al*. (1994) 'The quantitative analysis of mammographic densities.', *Physics in medicine and biology*, 39(10), pp. 1629–38. doi: 10.1088/0031-9155/39/10/008.
Byng, J. W. *et al*. (1998) 'Analysis of mammographic density and breast cancer risk from digitized mammograms.', *Radiographics : a review publication of the Radiological Society of North America, Inc*, 18(6), pp. 1587–98. doi: 10.1148/radiographics.18.6.9821201.
Byrne, C. *et al*. (1995) 'Mammographic features and breast cancer risk: effects with time, age, and menopause status.', *Journal of the National Cancer Institute*, 87(21), pp. 1622–9. doi: 10.1093/jnci/87.21.1622.
Carlson, C. S. *et al*. (2013) 'Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study.', *PLoS biology*, 11(9), p. e1001661. doi: 10.1371/journal.pbio.1001661.
Cauley, J. A. *et al*. (1999) 'Elevated serum estradiol and testosterone concentrations are associated with a high risk for breast cancer. Study of Osteoporotic Fractures Research Group.', *Annals of internal medicine*, 130(4 Pt 1), pp. 270–7. doi: 10.7326/0003-4819-130-4_part_1-199902160-00004.
Chang, C. C. *et al*. (2015) 'Second-generation PLINK: rising to the challenge of larger and richer datasets', *GigaScience*, 4(1), p. 7. doi: 10.1186/s13742-015-0047-8.
Chen, H. *et al*. (2021) 'Genome-wide Association Study Reveals Novel Loci Associated with Mammographic Density and Further Establishes its Shared Genetic Basis with Breast Cancer', *Under review*.
Ciatto, S. *et al*. (2012) 'A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification.', *Breast (Edinburgh, Scotland)*, 21(4), pp. 503–6. doi: 10.1016/j.breast.2012.01.005.
Deelen, P. *et al*. (2014) 'Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration', *BMC Research Notes*, 7(1), p. 901. doi: 10.1186/1756-0500-7-901.
Dumas, I. and Diorio, C. (2010) 'Polymorphisms in genes involved in the estrogen pathway and mammographic density', *BMC Cancer*, 10(1), p. 636. doi: 10.1186/1471-2407-10-636.
Endogenous Hormones and Breast Cancer Collaborative Group *et al*. (2011) 'Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies.', *British journal of cancer*, 105(5), pp. 709–22. doi: 10.1038/bjc.2011.254.
Finucane, H. K. *et al*. (2015) 'Partitioning heritability by functional annotation using genome-wide association summary statistics.', *Nature genetics*, 47(11), pp. 1228–35. doi: 10.1038/ng.3404.
Flynn, E. *et al*. (2021) 'Sex-specific genetic effects across biomarkers', *European Journal of*

Human Genetics, 29(1), pp. 154–163. doi: 10.1038/s41431-020-00712-w.
Folkerd, E. and Dowsett, M. (2013) 'Sex hormones and breast cancer risk and prognosis', *The Breast*, 22, pp. S38–S43. doi: 10.1016/j.breast.2013.07.007.
Greendale, G. A. *et al*. (2005) 'The association of endogenous sex steroids and sex steroid binding proteins with mammographic density: results from the Postmenopausal Estrogen/Progestin Interventions Mammographic Density Study.', *American journal of epidemiology*, 162(9), pp. 826–34. doi: 10.1093/aje/kwi286.
GTEx Consortium (2013) 'The Genotype-Tissue Expression (GTEx) project.', *Nature genetics*, 45(6), pp. 580–5. doi: 10.1038/ng.2653.
Guo, Y. *et al*. (2016) 'Genetically Predicted Body Mass Index and Breast Cancer Risk: Mendelian Randomization Analyses of Data from 145,000 Women of European Descent', *PLOS Medicine*. Edited by A. H. Beck, 13(8), p. e1002105. doi: 10.1371/journal.pmed.1002105.
Haas, C. B. *et al*. (no date) 'Cross-ancestry genome wide association studies of sex hormones in pre- and postmenopausal women in the UK Biobank.', *Journal of Clinical Metabolism and Endocrinology*.
Haas, C. B. (no date) *Supplementary Materials for UKB GWAS*, *GitHub Repository*. Available at: https://github.com/cbhaas/UKB_HormonesMenopauseGWAS.git.
Hankinson, S. E. and Eliassen, A. H. (2010) 'Circulating sex steroids and breast cancer risk in premenopausal women.', *Hormones & cancer*, 1(1), pp. 2–10. doi: 10.1007/s12672-009-0003-0.
Howie, B. *et al*. (2012) 'Fast and accurate genotype imputation in genome-wide association studies through pre-phasing.', *Nature genetics*, 44(8), pp. 955–9. doi: 10.1038/ng.2354.
Jiao, X. *et al*. (2018) 'Molecular Genetics of Premature Ovarian Insufficiency', *Trends in Endocrinology & Metabolism*, 29(11), pp. 795–807. doi: 10.1016/j.tem.2018.07.002.
Johnson, N. *et al*. (2016) 'Cytochrome P450 Allele CYP3A7*1C Associates with Adverse Outcomes in Chronic Lymphocytic Leukemia, Breast, and Lung Cancer', *Cancer Research*, 76(6), pp. 1485–1493. doi: 10.1158/0008-5472.CAN-15-1410.
Jones, M. E. *et al*. (2013) 'Changes in estradiol and testosterone levels in postmenopausal women after changes in body mass index.', *The Journal of clinical endocrinology and metabolism*, 98(7), pp. 2967–74. doi: 10.1210/jc.2013-1588.
Key, T. *et al*. (2002) 'Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies.', *Journal of the National Cancer Institute*, 94(8), pp. 606–16. doi: 10.1093/jnci/94.8.606.
Kim, C. and Halter, J. B. (2014) 'Endogenous Sex Hormones, Metabolic Syndrome, and Diabetes in Men and Women', *Current Cardiology Reports*, 16(4), p. 467. doi: 10.1007/s11886-014-0467-6.
Lauretta, R. *et al*. (2018) 'Gender in Endocrine Diseases: Role of Sex Gonadal Hormones', *International Journal of Endocrinology*, 2018, pp. 1–11. doi: 10.1155/2018/4847376.
de Leeuw, C. A. *et al*. (2015) 'MAGMA: Generalized Gene-Set Analysis of GWAS Data', *PLOS Computational Biology*. Edited by H. Tang, 11(4), p. e1004219. doi: 10.1371/journal.pcbi.1004219.
Lindström, S. *et al*. (2014) 'Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk.', *Nature communications*, 5, p. 5303. doi: 10.1038/ncomms6303.
Loh, P.-R. *et al*. (2015) 'Efficient Bayesian mixed-model analysis increases association power in large cohorts', *Nature Genetics*, 47(3), pp. 284–290. doi: 10.1038/ng.3190.
Lord, S. J. *et al*. (2005) 'Polymorphisms in genes involved in estrogen and progesterone

metabolism and mammographic density changes in women randomized to postmenopausal hormone therapy: results from a pilot study', *Breast Cancer Research*, 7(3), p. R336. doi: 10.1186/bcr999.
Luan, J. A. *et al*. (2001) 'Sample size determination for studies of gene-environment interaction.', *International journal of epidemiology*, 30(5), pp. 1035–40. doi: 10.1016/j.hoc.2005.02.009.
Neale, B. (2019) *GWAS OF UK BIOBANK BIOMARKER MEASUREMENTS*. Available at: http://www.nealelab.is/uk-biobank.
Nguyen, T. L. *et al*. (2013) 'Explaining variance in the cumulus mammographic measures that predict breast cancer risk: a twins and sisters study.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 22(12), pp. 2395–403. doi: 10.1158/1055-9965.EPI-13-0481.
Nickels, S. *et al*. (2013) 'Evidence of Gene–Environment Interactions between Common Breast Cancer Susceptibility Loci and Established Environmental Risk Factors', *PLoS Genetics*. Edited by M. S. Horwitz, 9(3), p. e1003284. doi: 10.1371/journal.pgen.1003284.
Olson, J. E. *et al*. (2012) 'The influence of mammogram acquisition on the mammographic density and breast cancer association in the Mayo Mammography Health Study cohort.', *Breast cancer research : BCR*, 14(6), p. R147. doi: 10.1186/bcr3357.
Ooi, B. N. S. *et al*. (2019) 'The genetic interplay between body mass index, breast size and breast cancer risk: a Mendelian randomization analysis', *International Journal of Epidemiology*, 48(3), pp. 781–794. doi: 10.1093/ije/dyz124.
Peterson, R. E. *et al*. (2017) 'The utility of empirically assigning ancestry groups in cross-population genetic studies of addiction', *The American Journal on Addictions*, 26(5), pp. 494–501. doi: 10.1111/ajad.12586.
Pettersson, A. *et al*. (2014) 'Mammographic density phenotypes and risk of breast cancer: a meta-analysis.', *Journal of the National Cancer Institute*, 106(5). doi: 10.1093/jnci/dju078.
Popejoy, A. B. and Fullerton, S. M. (2016) 'Genomics is failing on diversity.', *Nature*, 538(7624), pp. 161–164. doi: 10.1038/538161a.
Prescott, J. *et al*. (2012) 'Genome-Wide Association Study of Circulating Estradiol, Testosterone, and Sex Hormone-Binding Globulin in Postmenopausal Women', *PLoS ONE*. Edited by O. Y. Gorlova, 7(6), p. e37815. doi: 10.1371/journal.pone.0037815.
Price, A. L. *et al*. (2006) 'Principal components analysis corrects for stratification in genome-wide association studies', *Nature Genetics*, 38(8), pp. 904–909. doi: 10.1038/ng1847.
Pulit, S. L. *et al*. (2019) 'Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry', *Human Molecular Genetics*, 28(1), pp. 166–174. doi: 10.1093/hmg/ddy327.
Purcell, S. *et al*. (2007) 'PLINK: a tool set for whole-genome association and population-based linkage analyses.', *American journal of human genetics*, 81(3), pp. 559–75. doi: 10.1086/519795.
Purcell, S. and Chang, C. (no date) 'PLINK v2.0'. Available at: www.cog-genomics.org/plink/2.0/.
Quan, L. *et al*. (2014) 'Variants of estrogen-related genes and breast cancer risk in European and African American women.', *Endocrine-related cancer*, 21(6), pp. 853–64. doi: 10.1530/ERC-14-0250.
Rafnar, T. *et al*. (2018) 'Variants associating with uterine leiomyoma highlight genetic

background shared by various cancers and hormone-related traits', *Nature Communications*, 9(1), p. 3636. doi: 10.1038/s41467-018-05428-6.
Ring, H. Z. *et al*. (2005) 'Heritability of plasma sex hormones and hormone binding globulin in adult male twins.', *The Journal of clinical endocrinology and metabolism*, 90(6), pp. 3653–8. doi: 10.1210/jc.2004-1025.
Rudolph, A. *et al*. (2015) 'A comprehensive evaluation of interaction between genetic variants and use of menopausal hormone therapy on mammographic density.', *Breast cancer research : BCR*, 17, p. 110. doi: 10.1186/s13058-015-0625-9.
Rudolph, A., Chang-Claude, J. and Schmidt, M. K. (2016) 'Gene-environment interaction and risk of breast cancer.', *British journal of cancer*, 114(2), pp. 125–33. doi: 10.1038/bjc.2015.439.
Ruth, K. S. *et al*. (2020) 'Using human genetics to understand the disease impacts of testosterone in men and women', *Nature Medicine*, 26(2), pp. 252–258. doi: 10.1038/s41591-020-0751-5.
Sato, Y. *et al*. (2017) 'An independent validation study of three single nucleotide polymorphisms at the sex hormone-binding globulin locus for testosterone levels identified by genome-wide association studies', *Human Reproduction Open*, 2017(1). doi: 10.1093/hropen/hox002.
Schwarzer, G., Carpenter, J. R. and Rücker, G. (2015) *Meta-Analysis with R*. Cham: Springer International Publishing (Use R!). doi: 10.1007/978-3-319-21416-0.
Shawky, M. S. *et al*. (2017) 'Mammographic density: a potential monitoring biomarker for adjuvant and preventative breast cancer endocrine therapies.', *Oncotarget*, 8(3), pp. 5578–5591. doi: 10.18632/oncotarget.13484.
Shu, X. *et al*. (2019) 'Associations of obesity and circulating insulin and glucose with breast cancer risk: a Mendelian randomization analysis.', *International journal of epidemiology*, 48(3), pp. 795–806. doi: 10.1093/ije/dyy201.
Sprague, B. L. *et al*. (2011) 'Circulating sex hormones and mammographic breast density among postmenopausal women.', *Hormones & cancer*, 2(1), pp. 62–72. doi: 10.1007/s12672-010-0056-0.
Stanczyk, F. Z., Mathews, B. W. and Sherman, M. E. (2015) 'Relationships of sex steroid hormone levels in benign and cancerous breast tissue and blood: A critical appraisal of current science', *Steroids*, 99, pp. 91–102. doi: 10.1016/j.steroids.2014.12.011.
Stevens, K. N. *et al*. (2012) 'Identification of a novel percent mammographic density locus at 12q24.', *Human molecular genetics*, 21(14), pp. 3299–305. doi: 10.1093/hmg/dds158.
Stone, J. *et al*. (2015) 'Novel Associations between Common Breast Cancer Susceptibility Variants and Risk-Predicting Mammographic Density Measures.', *Cancer research*, 75(12), pp. 2457–67. doi: 10.1158/0008-5472.CAN-14-2012.
Tamimi, R. M. *et al*. (2005) 'Endogenous sex hormone levels and mammographic density among postmenopausal women.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 14(11 Pt 1), pp. 2641–7. doi: 10.1158/1055-9965.EPI-05-0558.
Tang, B. *et al*. (2021) 'Obesity-Related Traits and the Development of Rheumatoid Arthritis: Evidence From Genetic Data', *Arthritis & Rheumatology*, 73(2), pp. 203–211. doi: 10.1002/art.41517.
Team, R. C. (no date) *R: A language and environment for statistical computing*. Vienna, Austria. Available at: http://www.r-project.org/.
Tehranifar, P. *et al*. (2011) 'Reproductive and menstrual factors and mammographic density in African American, Caribbean, and white women.', *Cancer causes & control : CCC*, 22(4), pp. 599–610. doi: 10.1007/s10552-011-9733-8.

Travis, R. C. *et al*. (2010) 'Gene–environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study', *The Lancet*, 375(9732), pp. 2143–2151. doi: 10.1016/S0140-6736(10)60636-8.
Vachon, C. M. *et al*. (2012) 'Common breast cancer susceptibility variants in LSP1 and RAD51L1 are associated with mammographic density measures that predict breast cancer risk.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 21(7), pp. 1156–66. doi: 10.1158/1055-9965.EPI-12-0066.
Varghese, J. S. *et al*. (2012) 'The heritability of mammographic breast density and circulating sex-hormone levels: two independent breast cancer risk factors.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 21(12), pp. 2167–75. doi: 10.1158/1055-9965.EPI-12-0789.
Vitale, C., Mendelsohn, M. E. and Rosano, G. M. C. (2009) 'Gender differences in the cardiovascular effect of sex hormones', *Nature Reviews Cardiology*, 6(8), pp. 532–542. doi: 10.1038/nrcardio.2009.105.
Wagner, J. D., Kaplan, J. R. and Burkman, R. T. (2002) 'Reproductive hormones and cardiovascular disease', *Obstetrics and Gynecology Clinics of North America*, 29(3), pp. 475–493. doi: 10.1016/S0889-8545(02)00011-6.
Warren, R. *et al*. (2006) 'Associations among mammographic density, circulating sex hormones, and polymorphisms in sex hormone metabolism genes in postmenopausal women.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 15(8), pp. 1502–8. doi: 10.1158/1055-9965.EPI-05-0828.
Watanabe, K. *et al*. (2017) 'Functional mapping and annotation of genetic associations with FUMA', *Nature Communications*, 8(1), p. 1826. doi: 10.1038/s41467-017-01261-5.
Willer, C. J., Li, Y. and Abecasis, G. R. (2010) 'METAL: fast and efficient meta-analysis of genomewide association scans.', *Bioinformatics (Oxford, England)*, 26(17), pp. 2190–1. doi: 10.1093/bioinformatics/btq340.
Wong, C. S. *et al*. (2011) 'Mammographic density and its interaction with other breast cancer risk factors in an Asian population.', *British journal of cancer*, 104(5), pp. 871–4. doi: 10.1038/sj.bjc.6606085.
Woolcott, C. G. *et al*. (2013) 'Association between sex hormones, glucose homeostasis, adipokines, and inflammatory markers and mammographic density among postmenopausal women.', *Breast cancer research and treatment*, 139(1), pp. 255–65. doi: 10.1007/s10549-013-2534-x.
Yaffe, M. *et al*. (2005) 'Is mammographic density, as currently measured, a robust surrogate marker for breast cancer?', *Gynecological Endocrinology*, 21(sup1), pp. 17–21. doi: 10.1080/09513590400030004.
Yang, J. *et al*. (2011) 'GCTA: a tool for genome-wide complex trait analysis.', *American journal of human genetics*, 88(1), pp. 76–82. doi: 10.1016/j.ajhg.2010.11.011.
Yavorska, Olena; Staley, J. <james. staley@bristol. ac. uk. (2017) 'MendelianRandomization: Mendelian Randomization Package', *R package version 0.4.1*. Available at: https://cran.r-project.org/package=MendelianRandomization.
Yong, M. *et al*. (2009) 'Associations between endogenous sex hormone levels and mammographic and bone densities in premenopausal women', *Cancer Causes & Control*, 20(7),

pp. 1039–1053. doi: 10.1007/s10552-009-9321-3.
Zheng, J. *et al*. (2017) 'LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis', *Bioinformatics*, 33(2), pp. 272–279. doi: 10.1093/bioinformatics/btw613.

Supplementary Figures

Supplementary Figure 1.1A-B: GWAS results for estradiol as a continuous trait for women of European ancestry overall and in pre- and postmenopausal women only.

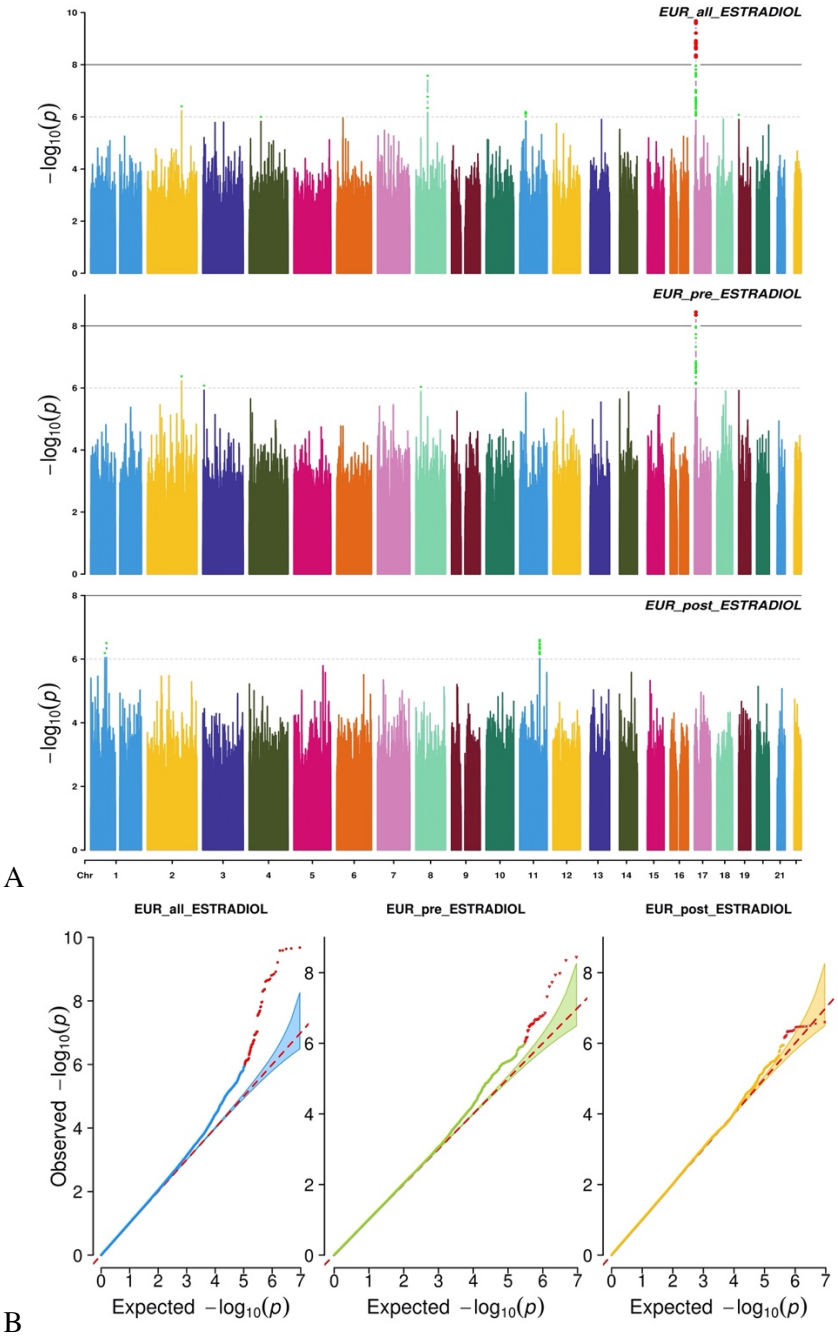
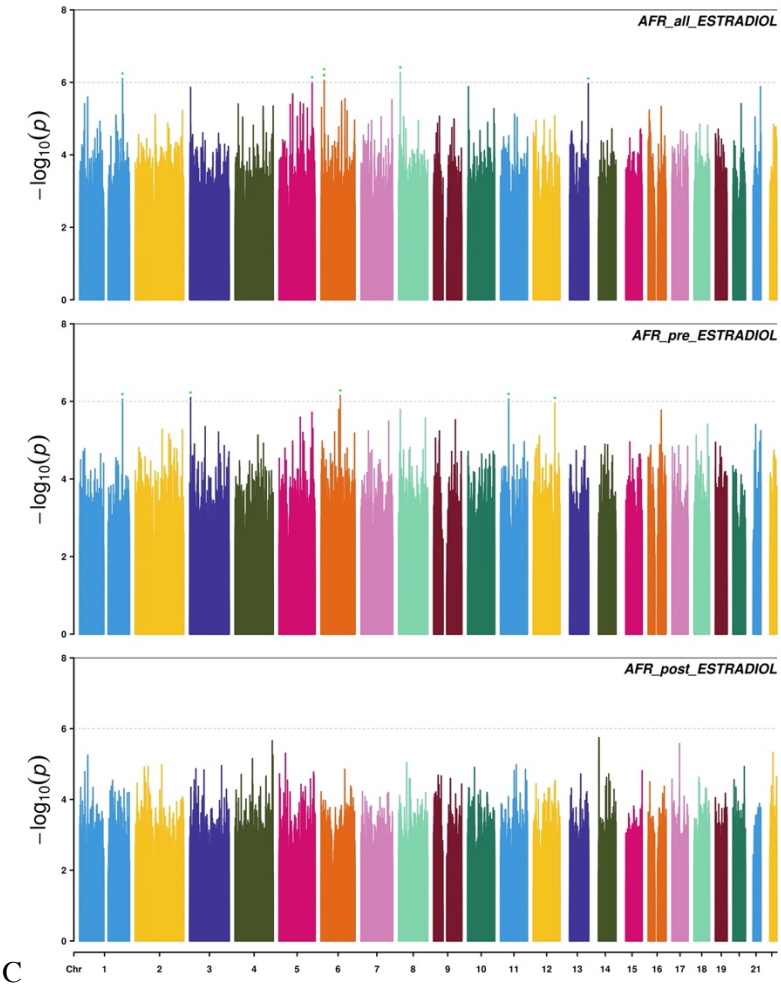


Figure A shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

Supplementary Figure 1.1C-D: GWAS results for estradiol as a continuous trait for women of African ancestry overall and in pre- and postmenopausal women only.

C



D

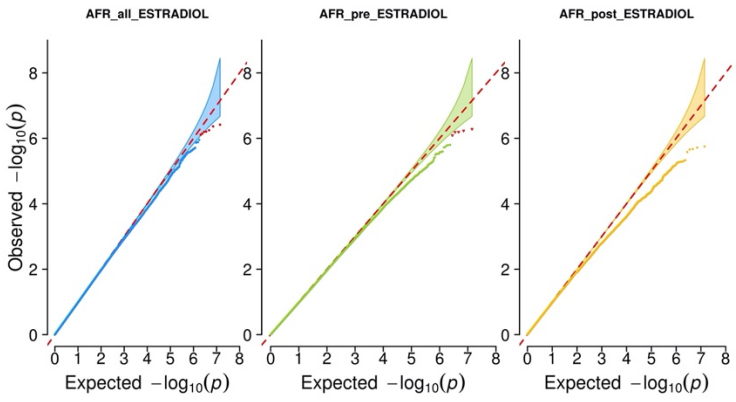
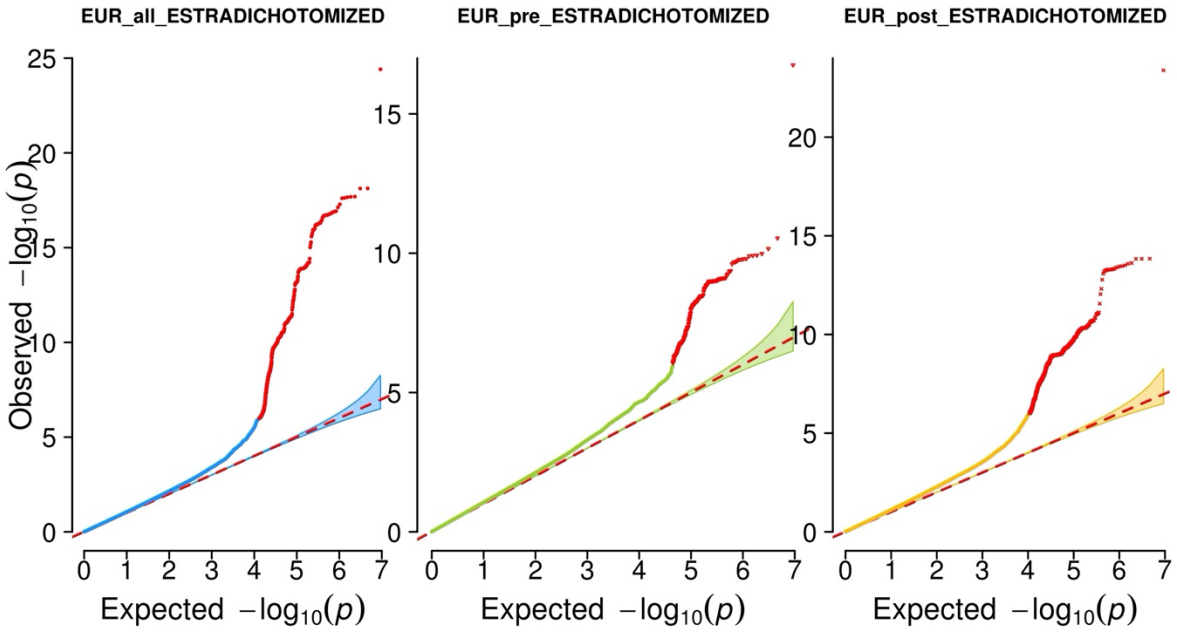


Figure C shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

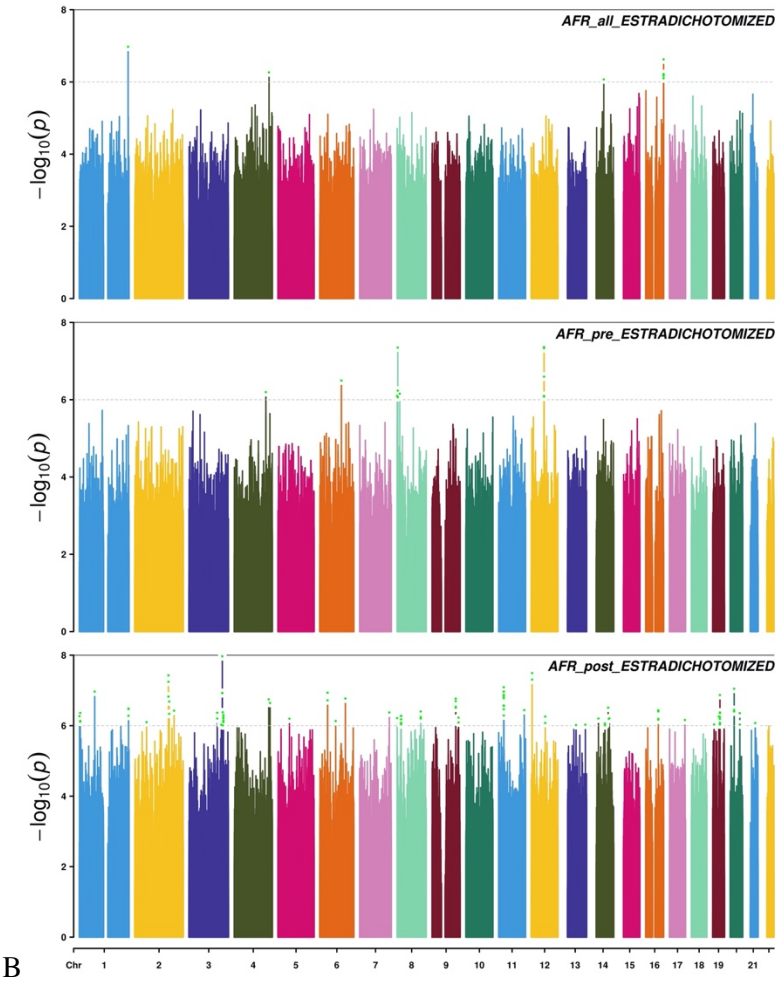
Supplementary Figure 1.2A: GWAS results for detectable levels of estradiol as a binary trait for women of European ancestry overall and in pre- and postmenopausal women only.



A

Supplementary Figure 2.2B-C: GWAS results for SHBG for women of African ancestry overall and in pre- and postmenopausal women only.

B



C

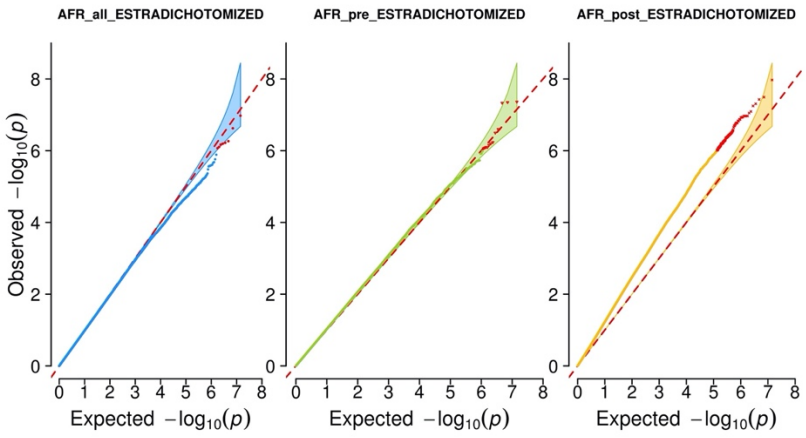
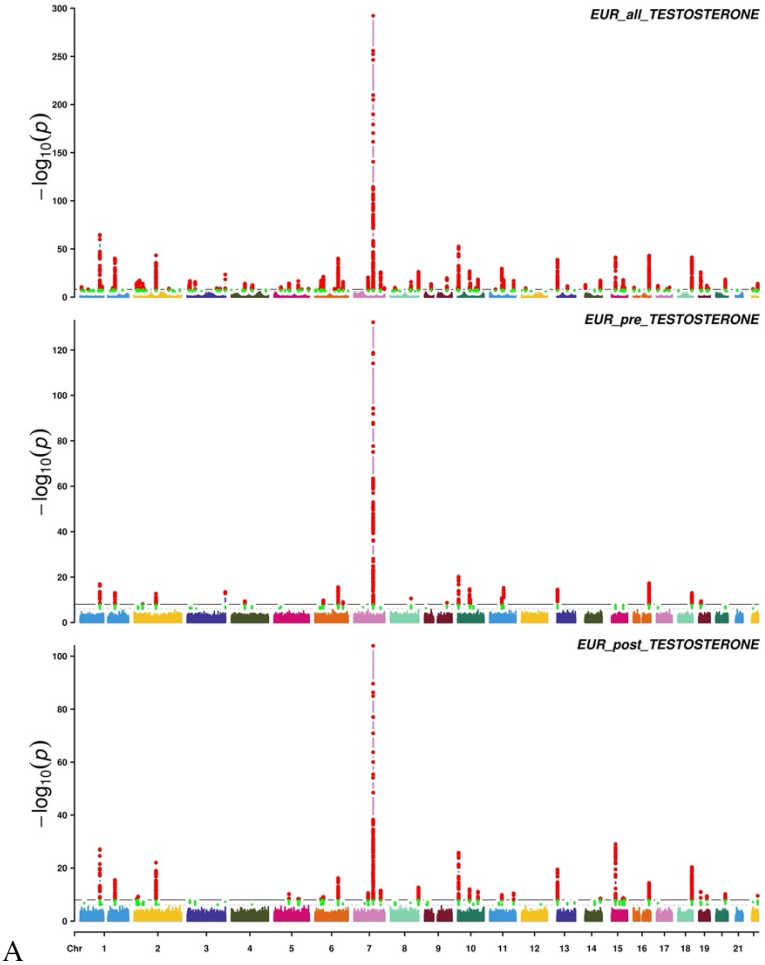
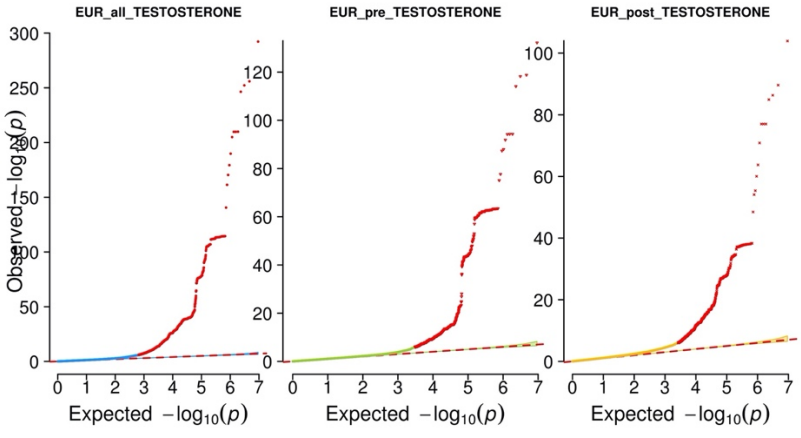


Figure A shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

Supplementary Figure 1.3A-B: GWAS results for testosterone for women of European ancestry overall and in pre- and postmenopausal women only.



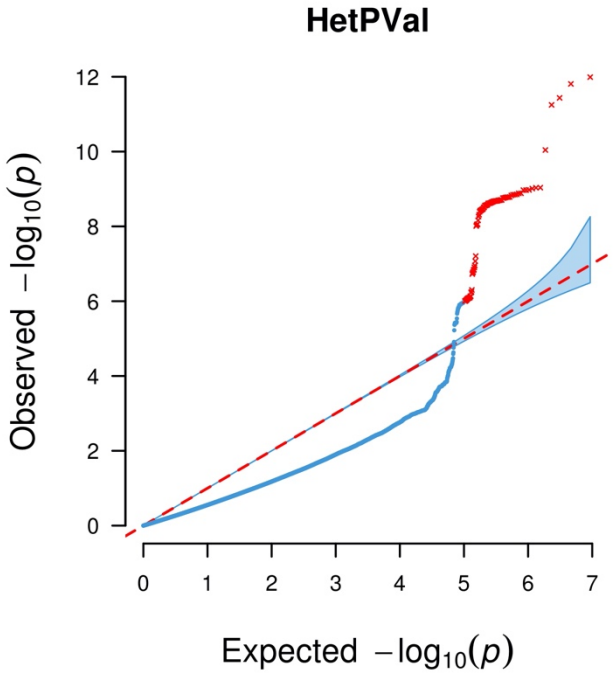
A



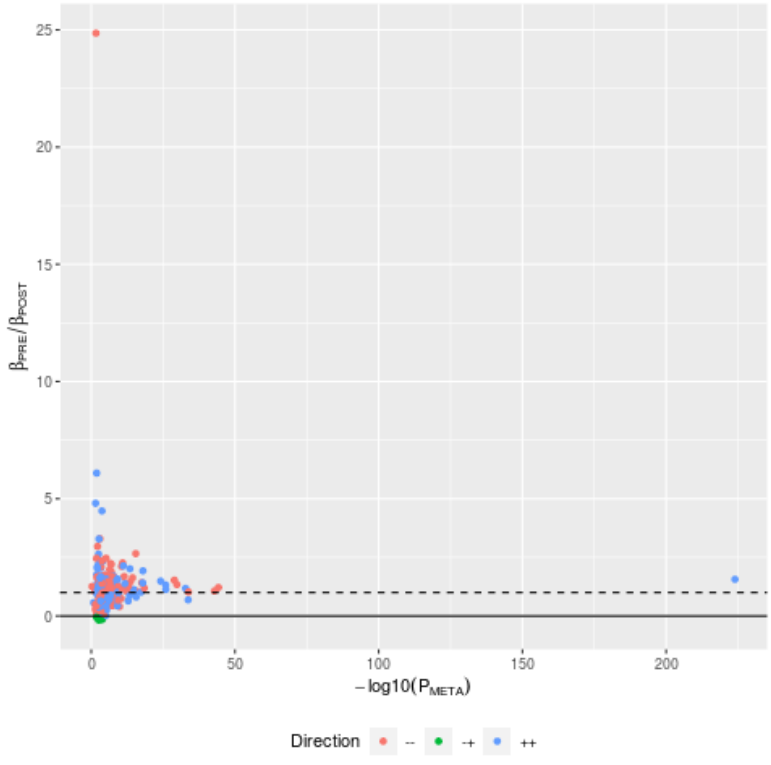
B

Figure A shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

Supplementary Figure 1.3C-D: GWAS for tests of heterogeneity of effects between pre- and postmenopausal status estimates for testosterone in women of European ancestry.

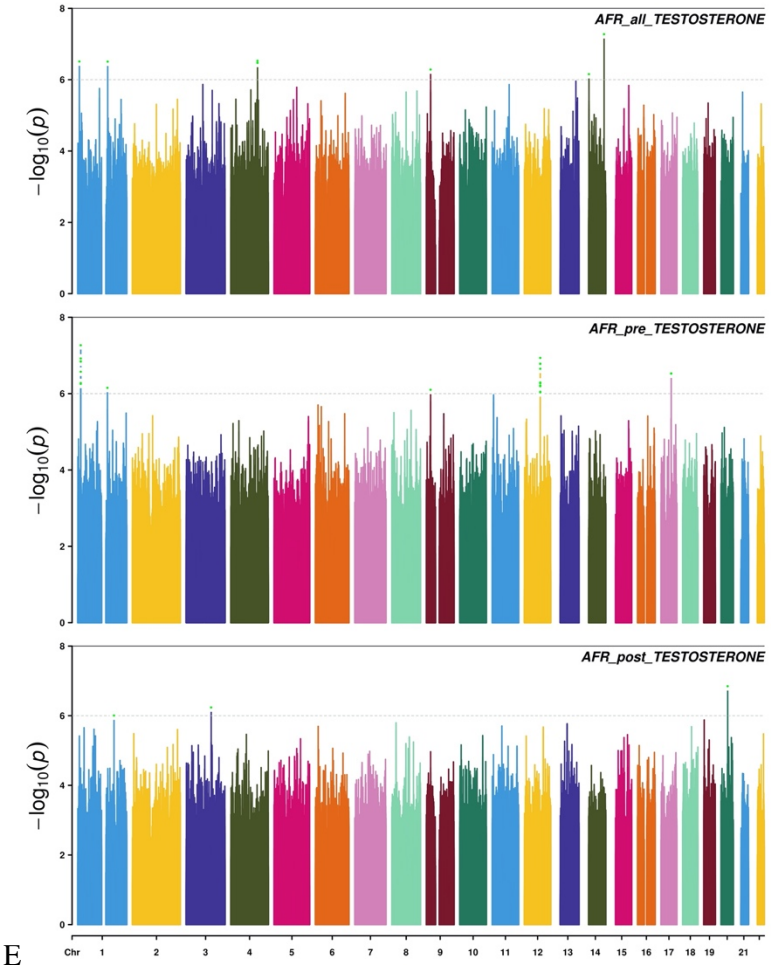


C

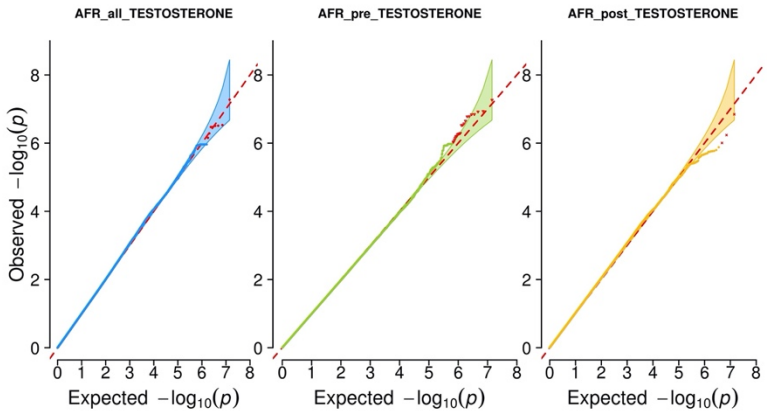


D

Supplementary Figure 1.3E-F: GWAS results for testosterone for women of African ancestry overall and in pre- and postmenopausal women only.



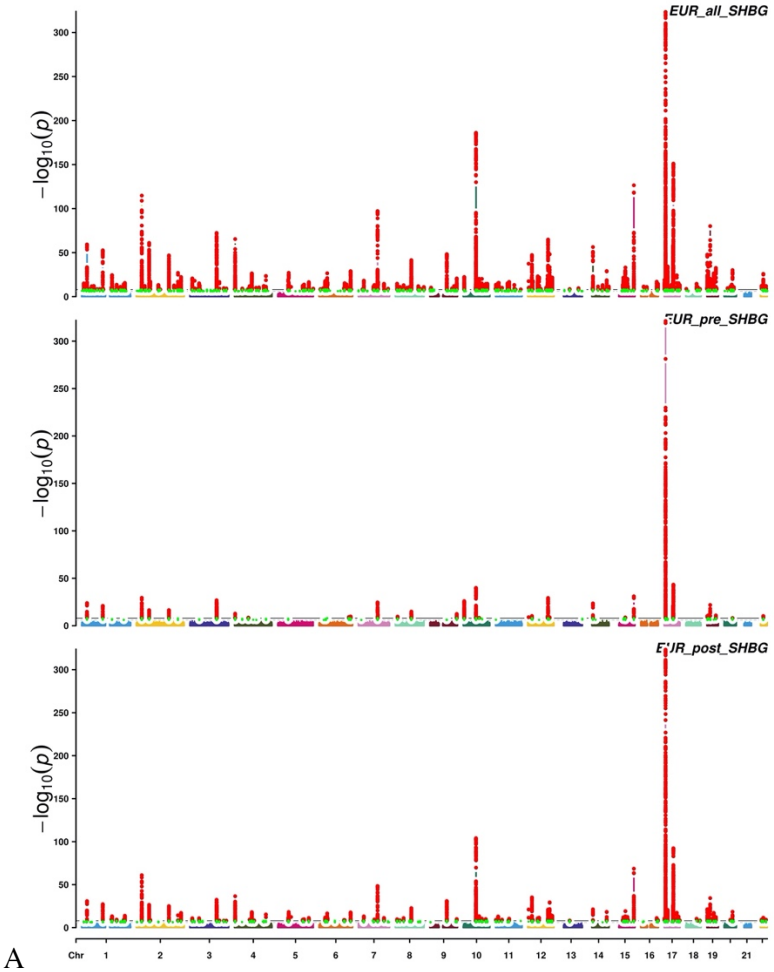
E



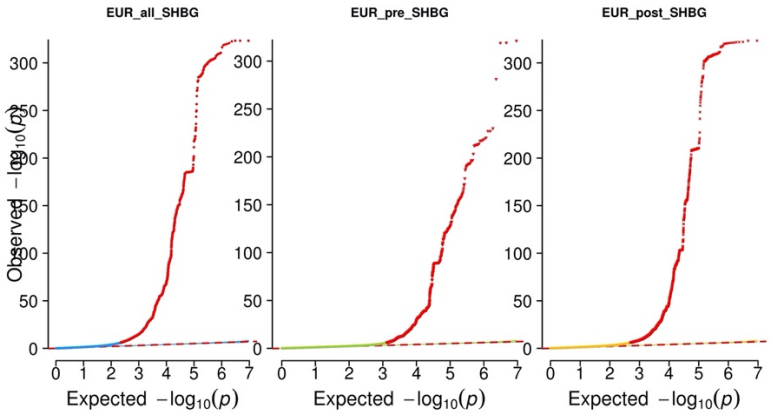
F

Figure E shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

Supplementary Figure 1.4A-B: GWAS results for SHBG for women of European ancestry overall and in pre- and postmenopausal women only.



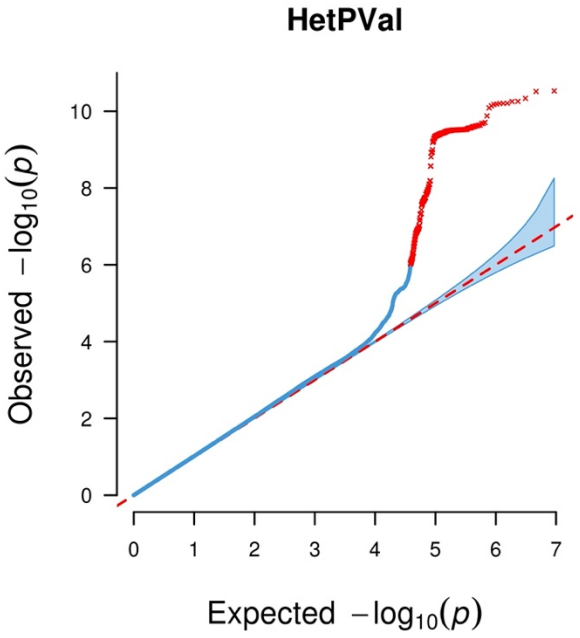
A



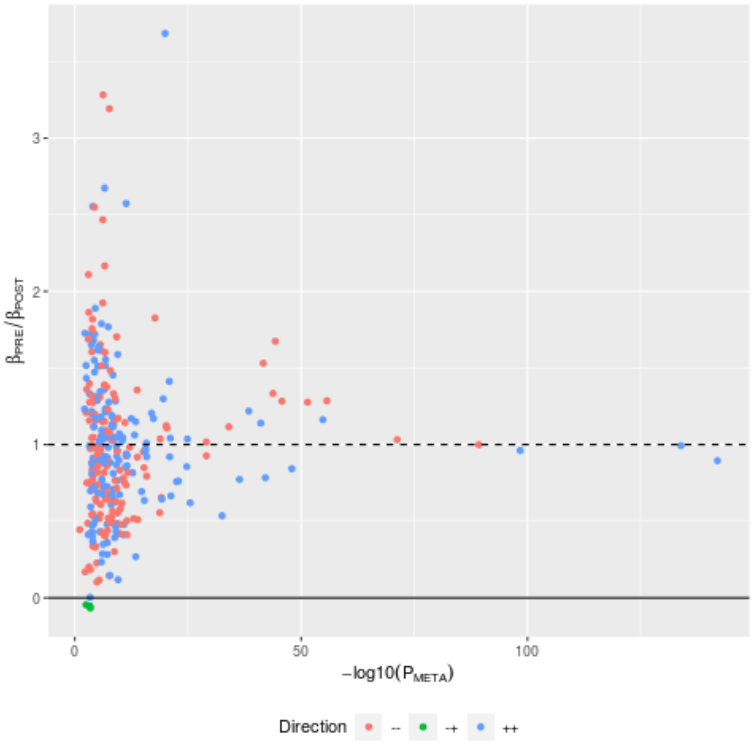
B

Figure A shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

Supplementary Figure 1.4C-D: GWAS for tests of heterogeneity of effects between pre- and postmenopausal status estimates for SHBG in women of European ancestry.

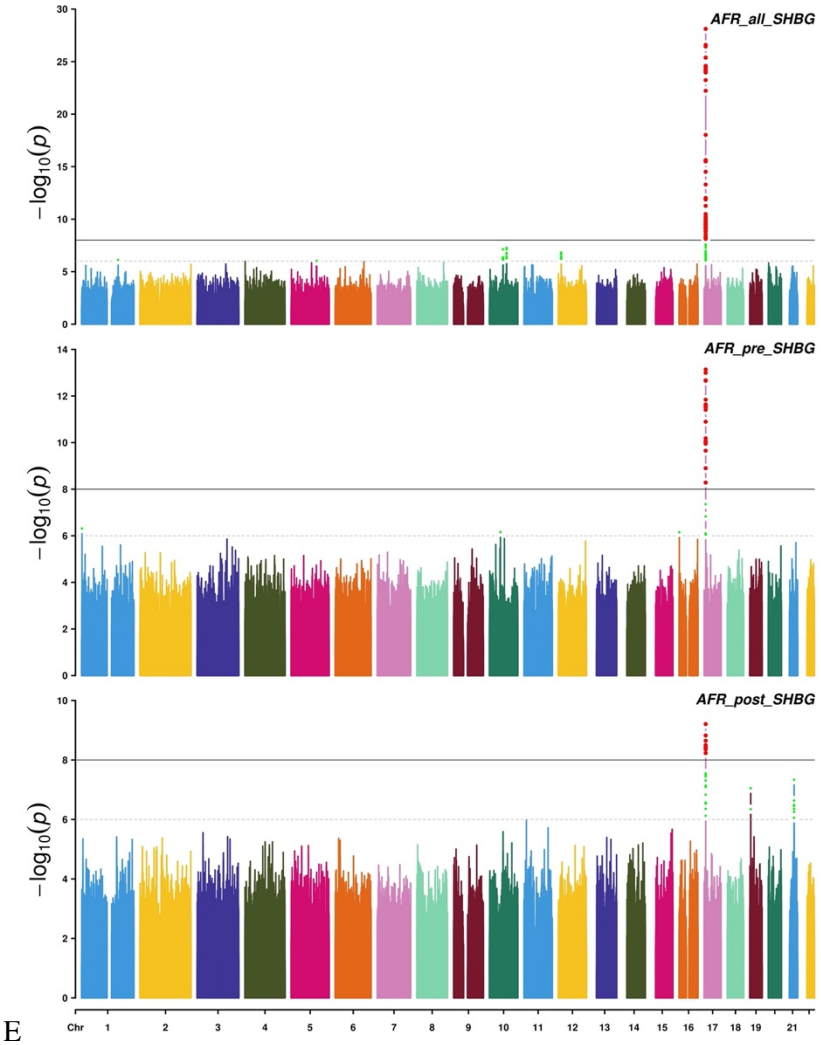


C

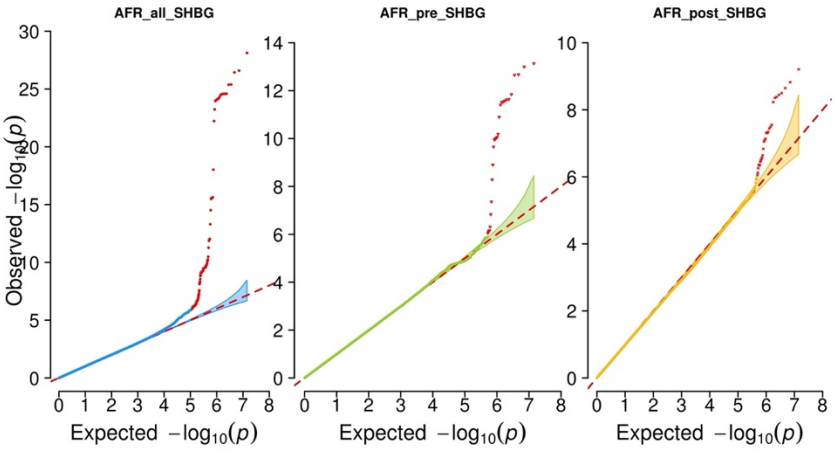


D

Supplementary Figure 1.4E-F: GWAS results for SHBG for women of African ancestry overall and in pre- and postmenopausal women only.



E



F

Figure E shows the Manhattan plots. Horizontal black line indicates Bonferroni correction threshold at p-value < 5E-8. Points in green indicate variants with p-value below a suggestive threshold of 1E-6, and red points indicate variants below Bonferroni correction p-value.

Supplementary Tables

Supplementary Table 1.1 - Independently significant signals from GWAS of continuous estradiol overall, by menopausal status, and across European and African ancestries
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.2 - Independently significant signals from GWAS of detectable estradiol overall, by menopausal status, and across European and African ancestries
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.3 - Previously identified signals from GWAS of testosterone in women overall, by menopausal status, and across European and African ancestries
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.4 - Independent significant signals from test of heterogeneity of effects for testosterone between pre- and postmenopausal women of European ancestry
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.5 - Gene set tissue specificity from GWAS results of testosterone in women of European ancestry
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.6 - Previously identified signals from GWAS of SHBG in women overall, by menopausal status, and across European and African ancestries
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.7 - Independent significant signals from test of heterogeneity of effects for SHBG between pre- and postmenopausal women of European ancestry
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.8 - Gene set tissue specificity from GWAS results of SHBG in women of European ancestry
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 1.9 - Genetic correlation estimates in women of African ancestry across sex hormone traits
https://github.com/cbhaas/UKB_HormonesMenopauseGWAS

Supplementary Table 2.1 Effect estimates based on main Mendelian Randomization approaches for overall and menopausal status specific sex hormone traits on mammographic density phenotypes.

Exposure	Outcome	Method	Estimate	Std Error	95% CI		P-value
Overall SHBG	Non-Dense Area	Simple median	-0.068889116	0.135319955	-0.33411135	0.196333123	0.610693907
Overall SHBG	Non-Dense Area	Weighted median	0.058995298	0.132303393	-0.20031459	0.318305183	0.655662847
Overall SHBG	Non-Dense Area	IVW	0.034367584	0.085755926	-0.13371094	0.202446111	0.688596492
Overall SHBG	Non-Dense Area	MR-Egger	0.069915021	0.141362677	-0.20715074	0.346980776	0.620897283
Overall SHBG	Non-Dense Area	(intercept)	-0.001810303	0.005714886	-0.01301127	0.009390668	0.751418284
Overall SHBG	Dense Area	Simple median	0.00301992	0.092200157	-0.17768907	0.183728906	0.973870795
Overall SHBG	Dense Area	Weighted median	0.122155371	0.088228541	-0.05076939	0.295080133	0.166195156
Overall SHBG	Dense Area	IVW	0.047082211	0.059200368	-0.06894838	0.163112801	0.426437465
Overall SHBG	Dense Area	MR-Egger	0.09269858	0.097405713	-0.09821311	0.283610271	0.341261843
Overall SHBG	Dense Area	(intercept)	-0.002327209	0.003942086	-0.01005355	0.005399137	0.554956307
Premenopausal SHBG	Non-Dense Area	Simple median	0.034548833	0.147824384	-0.25518164	0.324279301	0.815205943
Premenopausal SHBG	Non-Dense Area	Weighted median	-0.05160541	0.116520853	-0.27998208	0.176771266	0.657848474
Premenopausal SHBG	Non-Dense Area	IVW	-0.016271696	0.097207917	-0.20679571	0.174252321	0.867062687
Premenopausal SHBG	Non-Dense Area	MR-Egger	-0.006733686	0.170754352	-0.34140607	0.327938695	0.968543633
Premenopausal SHBG	Non-Dense Area	(intercept)	-0.001078899	0.015720998	-0.03189149	0.029733692	0.945285828
Premenopausal SHBG	Dense Area	Simple median	-0.006487474	0.103589791	-0.20951973	0.196544786	0.950063868
Premenopausal SHBG	Dense Area	Weighted median	0.11145355	0.077965679	-0.04135637	0.264263472	0.152854636
Premenopausal SHBG	Dense Area	IVW	0.053600306	0.087413549	-0.1177271	0.224927715	0.539756828
Premenopausal SHBG	Dense Area	MR-Egger	0.165820168	0.150884877	-0.12990876	0.461549092	0.271774746
Premenopausal SHBG	Dense Area	(intercept)	-0.012709603	0.013905397	-0.03996368	0.014544473	0.360714172
Postmenopausal SHBG	Non-Dense Area	Simple median	0.083771276	0.139068455	-0.18879789	0.356340439	0.546924953
Postmenopausal SHBG	Non-Dense Area	Weighted median	-0.032347645	0.128408462	-0.28402361	0.219328316	0.80110901
Postmenopausal SHBG	Non-Dense Area	IVW	0.019546598	0.085590018	-0.14820676	0.187299951	0.819354931
Postmenopausal SHBG	Non-Dense Area	MR-Egger	-0.073176058	0.146697807	-0.36069848	0.214346361	0.617904958
Postmenopausal SHBG	Non-Dense Area	(intercept)	0.006312044	0.008101178	-0.00956597	0.022190061	0.435890513
Postmenopausal SHBG	Dense Area	Simple median	-0.056836569	0.098376909	-0.24965177	0.13597863	0.563437656

Postmenopausal SHBG	Dense Area	Weighted median	0.119779932	0.085550644	-0.04789625	0.287456113	0.16148171
Postmenopausal SHBG	Dense Area	IVW	0.033424637	0.069106474	-0.10202156	0.168870838	0.628621026
Postmenopausal SHBG	Dense Area	MR-Egger	0.158316419	0.117595459	-0.07216645	0.388799283	0.178212221
Postmenopausal SHBG	Dense Area	(intercept)	-0.008515828	0.00650131	-0.02125816	0.004226507	0.190242109
Overall Testosterone	Non-Dense Area	Simple median	-0.085947182	0.204864965	-0.48747514	0.315580771	0.674828181
Overall Testosterone	Non-Dense Area	Weighted median	-0.19179681	0.219330755	-0.62167719	0.23808357	0.381865714
Overall Testosterone	Non-Dense Area	IVW	-0.0072579	0.152123492	-0.30541446	0.290898666	0.961946905
Overall Testosterone	Non-Dense Area	MR-Egger	-0.224502506	0.28795725	-0.78888835	0.339883334	0.435603797
Overall Testosterone	Non-Dense Area	(intercept)	0.009633938	0.010837069	-0.01160633	0.030874203	0.374013813
Overall Testosterone	Dense Area	Simple median	-0.265083093	0.131411889	-0.52264566	-0.00752052	0.043675416
Overall Testosterone	Dense Area	Weighted median	-0.086479378	0.139581826	-0.36005473	0.187095973	0.535547215
Overall Testosterone	Dense Area	IVW	-0.237096031	0.08134803	-0.39653524	-0.07765682	0.003561578
Overall Testosterone	Dense Area	MR-Egger	-0.034296432	0.15203702	-0.33228352	0.263690651	0.8215285
Overall Testosterone	Dense Area	(intercept)	-0.009011399	0.00572735	-0.0202368	0.002213999	0.115626726
Premenopausal Testosterone	Non-Dense Area	Simple median	0.142266738	0.203077465	-0.25575778	0.540291255	0.483581372
Premenopausal Testosterone	Non-Dense Area	Weighted median	0.144672137	0.175756673	-0.19980461	0.489148886	0.410429065
Premenopausal Testosterone	Non-Dense Area	IVW	0.08509571	0.161347246	-0.23113908	0.401330502	0.597910797
Premenopausal Testosterone	Non-Dense Area	MR-Egger	0.081164749	0.292276958	-0.49168756	0.654017061	0.781244137
Premenopausal Testosterone	Non-Dense Area	(intercept)	3.43E-04	0.021039991	-0.04089465	0.041580594	0.98699437
Premenopausal Testosterone	Dense Area	Simple median	-0.108356188	0.127172484	-0.35760968	0.1408973	0.394191259
Premenopausal Testosterone	Dense Area	Weighted median	-0.062745354	0.116200392	-0.29049394	0.165003229	0.589213991
Premenopausal Testosterone	Dense Area	IVW	-0.091773671	0.079774253	-0.24812833	0.064580991	0.249972088
Premenopausal Testosterone	Dense Area	MR-Egger	-0.052371039	0.140692135	-0.32812256	0.223380478	0.70971523

Premenopausal Testosterone	Dense Area	(intercept)	-0.003452301	0.010153762	-0.02335331	0.016448708	0.733854924
Postmenopausal Testosterone	Non-Dense Area	Simple median	0.222270741	0.245996357	-0.25987326	0.704414741	0.366232485
Postmenopausal Testosterone	Non-Dense Area	Weighted median	0.223174352	0.24310949	-0.25331149	0.699660197	0.35861921
Postmenopausal Testosterone	Non-Dense Area	IVW	0.23085278	0.175803942	-0.11371662	0.575422175	0.189140394
Postmenopausal Testosterone	Non-Dense Area	MR-Egger	0.264319774	0.378015343	-0.47657668	1.005216232	0.48440813
Postmenopausal Testosterone	Non-Dense Area	(intercept)	-0.001918917	0.019090565	-0.03933574	0.035497903	0.919934279
Postmenopausal Testosterone	Dense Area	Simple median	-0.083396268	0.157613213	-0.39231249	0.225519952	0.596722362
Postmenopausal Testosterone	Dense Area	Weighted median	-0.081482641	0.158873914	-0.39286979	0.229904508	0.608037957
Postmenopausal Testosterone	Dense Area	IVW	-0.13217041	0.128672892	-0.38436464	0.120023825	0.304335025
Postmenopausal Testosterone	Dense Area	MR-Egger	-0.025233415	0.274881518	-0.56399129	0.513524461	0.926858993
Postmenopausal Testosterone	Dense Area	(intercept)	-0.006141823	0.013893019	-0.03337164	0.021087995	0.658431487
Overall Estradiol	Non-Dense Area	Simple median	-0.916995028	0.68295261	-2.25555755	0.42156749	0.179371715
Overall Estradiol	Non-Dense Area	Weighted median	-0.897929781	0.648308021	-2.16859015	0.37273059	0.166041538
Overall Estradiol	Non-Dense Area	IVW	-0.944955065	0.472467998	-1.87097533	-0.0189348	0.045495906
Overall Estradiol	Non-Dense Area	MR-Egger	-3.126680558	1.264260969	-5.60458652	-0.64877459	0.013393576
Overall Estradiol	Non-Dense Area	(intercept)	0.059513828	0.031988187	-0.00318187	0.122209521	0.062815674
Overall Estradiol	Dense Area	Simple median	0.317161507	0.422742357	-0.51139829	1.145721301	0.453105498
Overall Estradiol	Dense Area	Weighted median	0.325207139	0.402359104	-0.46340221	1.113816491	0.418946116
Overall Estradiol	Dense Area	IVW	0.253041739	0.316361483	-0.36701537	0.873098852	0.423797697
Overall Estradiol	Dense Area	MR-Egger	-0.445817113	0.847612851	-2.10710778	1.215473548	0.598910541
Overall Estradiol	Dense Area	(intercept)	0.019057815	0.021443979	-0.02297161	0.061087242	0.374150517
Premenopausal Estradiol	Non-Dense Area	Simple median	0.122190274	0.539654755	-0.93551361	1.179894158	0.820872405

Premenopausal Estradiol	Non-Dense Area	Weighted median	-0.080938318	0.483986236	-1.02953391	0.867657273	0.867186966
Premenopausal Estradiol	Non-Dense Area	IVW	-0.376860704	0.419995643	-1.20003704	0.44631563	0.369560618
Premenopausal Estradiol	Non-Dense Area	MR-Egger	-2.026422691	0.87437467	-3.74016555	-0.31267983	0.020472815
Premenopausal Estradiol	Non-Dense Area	(intercept)	0.090128478	0.04362584	0.0046234	0.175633552	0.038833906
Premenopausal Estradiol	Dense Area	Simple median	0.030286451	0.313513211	-0.58418815	0.644761054	0.923041333
Premenopausal Estradiol	Dense Area	Weighted median	0.081355104	0.293926571	-0.49473039	0.657440598	0.781943533
Premenopausal Estradiol	Dense Area	IVW	0.079958013	0.238611424	-0.38771178	0.54762781	0.737551787
Premenopausal Estradiol	Dense Area	MR-Egger	-0.161148709	0.586567356	-1.3107996	0.988502182	0.783522293
Premenopausal Estradiol	Dense Area	(intercept)	0.013148885	0.029222391	-0.04412595	0.070423718	0.652739834
Postmenopausal Estradiol	Non-Dense Area	Simple median	-0.689034682	0.470021793	-1.61026047	0.232191103	0.142658295
Postmenopausal Estradiol	Non-Dense Area	Weighted median	-0.822400016	0.460195786	-1.72436718	0.079567149	0.073926901
Postmenopausal Estradiol	Non-Dense Area	IVW	-0.589849659	0.332188427	-1.24092701	0.061227694	0.075790981
Postmenopausal Estradiol	Non-Dense Area	MR-Egger	-1.284196164	0.755784031	-2.76550564	0.197113316	0.089289497
Postmenopausal Estradiol	Non-Dense Area	(intercept)	0.031151906	0.030457415	-0.02854353	0.090847344	0.306401458
Postmenopausal Estradiol	Dense Area	Simple median	0.448973979	0.300622594	-0.14023548	1.038183435	0.135311457
Postmenopausal Estradiol	Dense Area	Weighted median	0.228942984	0.292328458	-0.34401026	0.801896233	0.433527077
Postmenopausal Estradiol	Dense Area	IVW	0.258102474	0.222669887	-0.17832248	0.694527433	0.246404795
Postmenopausal Estradiol	Dense Area	MR-Egger	-0.506870923	0.507369144	-1.50129617	0.487554326	0.317785957
Postmenopausal Estradiol	Dense Area	(intercept)	0.034272664	0.020425267	-0.00576012	0.074305451	0.093356023

Supplementary Table 2.2 Effect estimates based on main Mendelian Randomization approaches for BMI on mammographic density phenotypes and overall and menopausal status specific sex hormone traits.

Outcome	Method	Estimate	Std Error	95% CI		P-value
Dense Area	Simple median	-0.2941032	0.1040339	-0.4980059	-0.0902005	0.00469872
Dense Area	Weighted median	-0.4983469	0.11076431	-0.7154409	-0.2812528	6.82E-06
Dense Area	IVW	-0.3730216	0.07185945	-0.5138635	-0.2321797	2.09E-07
Dense Area	MR-Egger	-0.7039285	0.19272032	-1.0816533	-0.3262036	2.60E-04
Dense Area	(intercept)	0.00570286	0.00308289	-3.39E-04	0.01174522	0.06433611
Premenopausal Dense Area	Simple median	-0.2001717	0.17675017	-0.5465956	0.14625231	0.25741932
Premenopausal Dense Area	Weighted median	-0.5743782	0.18812499	-0.9430964	-0.2056601	0.00226435
Premenopausal Dense Area	IVW	-0.2615932	0.11504023	-0.4870679	-0.0361185	0.02297033
Premenopausal Dense Area	MR-Egger	-1.0393585	0.30860478	-1.6442127	-0.4345043	7.57E-04
Premenopausal Dense Area	(intercept)	0.01337394	0.0049285	0.00371426	0.02303363	0.00665579
Postmenopausal Dense Area	Simple median	-0.3424353	0.12985925	-0.5969548	-0.0879159	0.00836495
Postmenopausal Dense Area	Weighted median	-0.5822819	0.13997192	-0.8566218	-0.3079419	3.18E-05
Postmenopausal Dense Area	IVW	-0.3321426	0.12934202	-0.5856483	-0.0786369	0.01023047
Postmenopausal Dense Area	MR-Egger	-0.6725835	0.34721224	-1.353107	0.00793996	0.05273358
Postmenopausal Dense Area	(intercept)	0.00586926	0.00555526	-0.0050188	0.01675737	0.29072883
Non-Dense Area	Simple median	1.96065217	0.15565443	1.65557509	2.26572925	2.22E-36
Non-Dense Area	Weighted median	2.10953354	0.17505645	1.76642919	2.45263788	1.93E-33
Non-Dense Area	IVW	1.79369829	0.10728651	1.58342059	2.00397599	9.57E-63
Non-Dense Area	MR-Egger	2.15420357	0.28837475	1.58899945	2.71940769	8.02E-14
Non-Dense Area	(intercept)	-0.0062107	0.00461196	-0.01525	0.00282858	0.17809312
Premenopausal Non-Dense Area	Simple median	1.35930247	0.29623135	0.77869968	1.93990525	4.46E-06

Premenopausal Non-Dense Area	Weighted median	1.81799284	0.31916602	1.19243893	2.44354676	1.23E-08
Premenopausal Non-Dense Area	IVW	1.74650773	0.1997748	1.35495633	2.13805914	2.28E-18
Premenopausal Non-Dense Area	MR-Egger	2.73519684	0.53752187	1.68167333	3.78872035	3.61E-07
Premenopausal Non-Dense Area	(intercept)	-0.0169983	0.00858306	-0.0338208	-1.76E-04	0.04765302
Postmenopausal Non-Dense Area	Simple median	1.74258226	0.18494292	1.38010079	2.10506372	4.41E-21
Postmenopausal Non-Dense Area	Weighted median	1.75784502	0.22911714	1.30878367	2.20690636	1.69E-14
Postmenopausal Non-Dense Area	IVW	1.74189786	0.12908173	1.4889023	1.99489341	1.68E-41
Postmenopausal Non-Dense Area	MR-Egger	1.64315755	0.34702162	0.96300768	2.32330743	2.19E-06
Postmenopausal Non-Dense Area	(intercept)	0.00170276	0.00555419	-0.0091833	0.01258878	0.75916971
Overall SHBG	Simple median	-0.4196252	0.01667884	-0.4523152	-0.3869353	1.12E-139
Overall SHBG	Weighted median	-0.3746688	0.01807198	-0.4100892	-0.3392483	1.78E-95
Overall SHBG	IVW	-0.353036	0.01922041	-0.3907074	-0.3153647	2.38E-75
Overall SHBG	MR-Egger	-0.3667364	0.05225458	-0.4691535	-0.2643193	2.25E-12
Overall SHBG	(intercept)	2.35E-04	8.34E-04	-0.0013992	0.00186948	0.77795429
Premenopausal SHBG	Simple median	-0.442325	0.03258182	-0.5061842	-0.3784658	5.57E-42
Premenopausal SHBG	Weighted median	-0.4351106	0.04063467	-0.5147531	-0.3554682	9.35E-27
Premenopausal SHBG	IVW	-0.436888	0.02615314	-0.4881472	-0.3856288	1.21E-62
Premenopausal SHBG	MR-Egger	-0.5291261	0.07101783	-0.6683185	-0.3899337	9.28E-14
Premenopausal SHBG	(intercept)	0.00158272	0.0011331	-6.38E-04	0.00380356	0.16247495
Postmenopausal SHBG	Simple median	-0.3837929	0.02339745	-0.4296511	-0.3379348	1.81E-60
Postmenopausal SHBG	Weighted median	-0.3005568	0.02504881	-0.3496516	-0.2514621	3.60E-33
Postmenopausal SHBG	IVW	-0.3324032	0.02295865	-0.3774014	-0.2874051	1.66E-47
Postmenopausal SHBG	MR-Egger	-0.3153915	0.06240284	-0.4376988	-0.1930842	4.32E-07

Postmenopausal SHBG	(intercept)	-2.92E-04	9.96E-04	-0.0022439	0.00165991	0.76936086
Overall Testosterone	Simple median	0.08841171	0.0163729	0.05632142	0.12050199	6.67E-08
Overall Testosterone	Weighted median	0.04070935	0.01775899	0.00590237	0.07551633	0.02188698
Overall Testosterone	IVW	0.08455736	0.01647773	0.0522616	0.11685312	2.87E-07
Overall Testosterone	MR-Egger	0.02957158	0.04473069	-0.058099	0.11724211	0.50854642
Overall Testosterone	(intercept)	9.44E-04	7.14E-04	-4.55E-04	0.0023427	0.18613098
Premenopausal Testosterone	Simple median	0.17208535	0.03197753	0.10941055	0.23476015	7.39E-08
Premenopausal Testosterone	Weighted median	0.11901061	0.03498181	0.05044752	0.18757371	6.69E-04
Premenopausal Testosterone	IVW	0.12652301	0.02654308	0.07449954	0.17854649	1.87E-06
Premenopausal Testosterone	MR-Egger	0.03719428	0.07208864	-0.1040969	0.17848543	0.60588789
Premenopausal Testosterone	(intercept)	0.00153274	0.00115015	-7.22E-04	0.00378701	0.1826494
Postmenopausal Testosterone	Simple median	0.07925636	0.02350757	0.03318237	0.12533036	7.48E-04
Postmenopausal Testosterone	Weighted median	0.06554453	0.0291011	0.00850743	0.12258164	0.02430304
Postmenopausal Testosterone	IVW	0.08463504	0.01960865	0.0462028	0.12306729	1.59E-05
Postmenopausal Testosterone	MR-Egger	0.08493949	0.05331293	-0.0195519	0.1894309	0.11110972
Postmenopausal Testosterone	(intercept)	-5.23E-06	8.51E-04	-0.0016727	0.00166222	0.99509952
Overall Estradiol	Simple median	-0.0212902	0.013707	-0.0481554	0.00557502	0.12036678
Overall Estradiol	Weighted median	-0.0231973	0.0172759	-0.0570574	0.01066287	0.17935196
Overall Estradiol	IVW	-0.0210805	0.00941557	-0.0395347	-0.0026263	0.02516248
Overall Estradiol	MR-Egger	-0.064702	0.02552395	-0.114728	-0.0146759	0.01124631
Overall Estradiol	(intercept)	7.49E-04	4.07E-04	-4.96E-05	0.00154696	0.06604506
Premenopausal Estradiol	Simple median	-0.0620475	0.0288171	-0.118528	-0.005567	0.031307
Premenopausal Estradiol	Weighted median	-0.042603	0.03505616	-0.1113118	0.02610581	0.22425978

Premenopausal Estradiol	IVW	-0.05342	0.01945786	-0.0915567	-0.0152833	0.00604333
Premenopausal Estradiol	MR-Egger	-0.0517755	0.05292895	-0.1555143	0.05196334	0.3279718
Premenopausal Estradiol	(intercept)	-2.82E-05	8.44E-04	-0.0016833	0.0016269	0.97334396
Postmenopausal Estradiol	Simple median	0.01030967	0.02282986	-0.034436	0.05505537	0.65156644
Postmenopausal Estradiol	Weighted median	0.01861768	0.02828235	-0.0368147	0.07405006	0.51035881
Postmenopausal Estradiol	IVW	0.00346422	0.01657947	-0.0290309	0.03595939	0.83448987
Postmenopausal Estradiol	MR-Egger	0.00317536	0.04507591	-0.0851718	0.09152252	0.94383965
Postmenopausal Estradiol	(intercept)	4.96E-06	7.19E-04	-0.0014049	0.00141478	0.99450072