Am J Cancer Res 2023;13(3):1128-1136
www.ajcr.us /ISSN:2156-6976/ajcr0148122

Brief Communication

Sex hormones in the risk of breast cancer: a two-sample Mendelian randomization study

Bin Ke, Chunyu Li, Huifang Shang

Department of Neurology, Laboratory of Neurodegenerative Disorders, National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu, Sichuan, China

Received November 29, 2022; Accepted February 24, 2023; Epub March 15, 2023; Published March 30, 2023

Abstract: Multiple evidence has suggested the essential role of sex hormones in the susceptibility of breast cancer. However, whether there exists a causal association and the effect direction remains controversial. To examine the causative role of hormones in the risk of breast cancer, we first estimated their genetic correlation, and then conducted two-sample and multivariable Mendelian randomization analyses using summary statistics from genome-wide association studies of major sex hormones including testosterone (N=230,454), estradiol (N=163,985) and progesterone (N=1,261), together with breast cancer (N=228,951). We further performed subtype analysis focusing on estrogen receptor (ER)+ breast cancer (N=175,475) and ER- breast cancer (N=127,442), and conducted extensive sensitivity analyses. We identified significant positive genetic correlation between testosterone level and risk of breast cancer (genetic correlation: 0.09, P=1.10E-03). Genetically determined higher total testosterone level was associated with an increased risk of breast cancer (OR: 1.11, 95% CI: 1.06-1.16, P=4.55E-06). In the subtype analysis, higher total testosterone was associated with an increased risk of ER+ breast cancer (OR: 1.18, 95% CI: 1.11-1.26, P=6.00E-08). In contrast, no association was identified between estradiol, progesterone and the risk of breast cancer. These results elucidated the causal role of major sex hormones in the risk of breast cancer, especially in ER+ breast cancer. Future development of preventive or therapeutic interventions in clinical trials could attach importance to this.

Keywords: Sex hormone, breast cancer, Mendelian randomization

Introduction

Breast cancer is the most frequent malignancy in women worldwide, mainly affecting women over the age of 50 [1]. Epidemiological studies have identified several risk factors for breast cancer such as aging, family history and lifestyles like alcohol consumption and dietary fat intake [2]. However, the identified risk factors explain only a limited amount of variance in the disease risk. Exploring novel factors influencing breast cancer could help better understand the pathogenesis of the disease, and provide care and therapeutic strategies for the patients and clinicians.

Discrepant steroid hormones between sexes and different ages might be one determinant based on evidence from previous epidemiological and clinical studies. Sex hormones, mainly including estrogens, androgens and progestogens, are molecules produced by the endocrine system that send messages to various parts of the body, and help regulate the body's processes. Multiple evidence has suggested sex hormones were involved in the etiology of breast cancer [3]. From the epidemiological perspective, the rates of breast cancer increase rapidly in the premenopausal years, while the rate of increase slows at menopause when endogenous hormone levels decline. A retrospective study among 10,786 healthy women with a follow-up of 13.5 years found that total and free testosterone levels were directly associated with increased breast cancer risk, while higher estradiol was associated with increased risk of human epidermal growth factor receptor 2 ($HER2)^-$ cancer but reduced risk of $HER2^+$ cancer [4]. Similarly, a collaborative analysis of seven prospective studies found that circulating estrogens and androgens were positively associated with the risk of breast cancer in pre-

menopausal women [5]. In contrast, other functional and epidemiological studies also reported antiproliferative effects of estrogens [6] and androgens [7, 8]. These evidence suggested the essential role of sex hormones in the pathogenesis of breast cancer. However, the observational studies might be biased by unavoidable confounding factors and small sample size, and cannot determine causation. Therefore, the causal association between sex hormones and breast cancer is still elusive.

In this context, we performed a two-sample Mendelian randomization (MR) analysis to explore the causal role of major sex hormones including estrogens measured by estradiol, androgens measured by testosterone, and progestogens in the risk of breast cancer. The MR approach is less susceptible to reverse causation or confounding factors which may distort the interpretations of conventional observational studies. We found that higher total testosterone was causally associated with higher risk of breast cancer.

Methods

Datasets

We obtained summary statistics of total testosterone levels (N=230,454) and bioavailable testosterone levels (N=188,507) in females from a previous large genome-wide association study (GWAS) based on genotype and phenotype data from the UK Biobank [9]. Testosterone was measured by "one step competitive analysis on a Beckman Coulter Unicel Dxl 800". Summary statistics of estradiol in females (N=163,985) were from another GWAS based on data from the UK Biobank [10]. Estradiol was measured by "two step competitive analysis on a Beckman Coulter Unicel Dxl 800". Summary statistics of progesterone in females (N=1,261) were from GWAS on steroid hormone levels based on individuals of European ancestry [11]. Progesterone was measured by liquid chromatography-tandem mass spectrometry. Details of the summary data from all GWAS were listed in Supplementary Table 1. Single nucleotide polymorphisms (SNP) that passed the genome-wide significance threshold (P<5E-08) were chosen as instrumental variables, which were then clumped based on the 1,000 Genomes Project linkage disequilibrium (LD) structure. Index SNPs (R^2<0.001 with any other associated SNP within 10 Mb) with the minimum *P* value were kept.

We obtained GWAS summary statistics of breast cancer in females from a genome-wide association study (N_{case}=122,977, $N_{control}$=105,974) [12]. Summary statistics of estrogen receptor (ER)-negative (ER-) (N_{case}=21,468, $N_{control}$=105,974) and ER-positive (ER+) (N_{case}=69,501, $N_{control}$=105,974) breast cancer were also from this study. Harmonization was undertaken to rule out strand mismatches and ensure alignment of SNP effect sizes.

Genetic correlation

We estimated the genetic correlation between sex hormones and breast cancer using LDSC [13] and GNOVA [14]. The LDSC method uses GWAS summary data to regress association test statistics of SNPs on their LD scores, which is defined as the sum of LD r^2 measured with all other SNPs in the reference sample. Compared with LDSC, GNOVA provides greater statistical power and higher estimation accuracy, especially when the correlation is moderate [14]. We ran LDSC and GNOVA on SNPs in both traits together with reference data derived from the 1000 Genomes Project European population using default parameters. A *P* value below 4.17E-03 (0.05/12) was considered statistically significant after the Bonferroni correction.

Mendelian randomization analysis

We hypothesized that sex hormones as a risk factor could causally influence the risk of breast cancer, and the following assumptions were satisfied: the genetic variants used as instrumental variables are associated with sex hormone levels; the genetic variants are not associated with any confounders; the genetic variants are associated with risk of breast cancer through sex hormones (namely horizontal pleiotropy should not be present) (Supplementary Figure 1).

To evaluate the causative effect of sex hormones on the risk of breast cancer, we performed a two-sample MR analysis using the random effects inverse variance weighted (IVW) method, which is most widely used in MR studies and could provide robust causal estimates under the absence of directional pleiotropy. A *P* value below 4.17E-03 (0.05/12) was considered statistically significant after the Bonferroni correction. We further verified the results using the weighted median method, which generally has greater power with a positive causal effect, particularly as the proportion

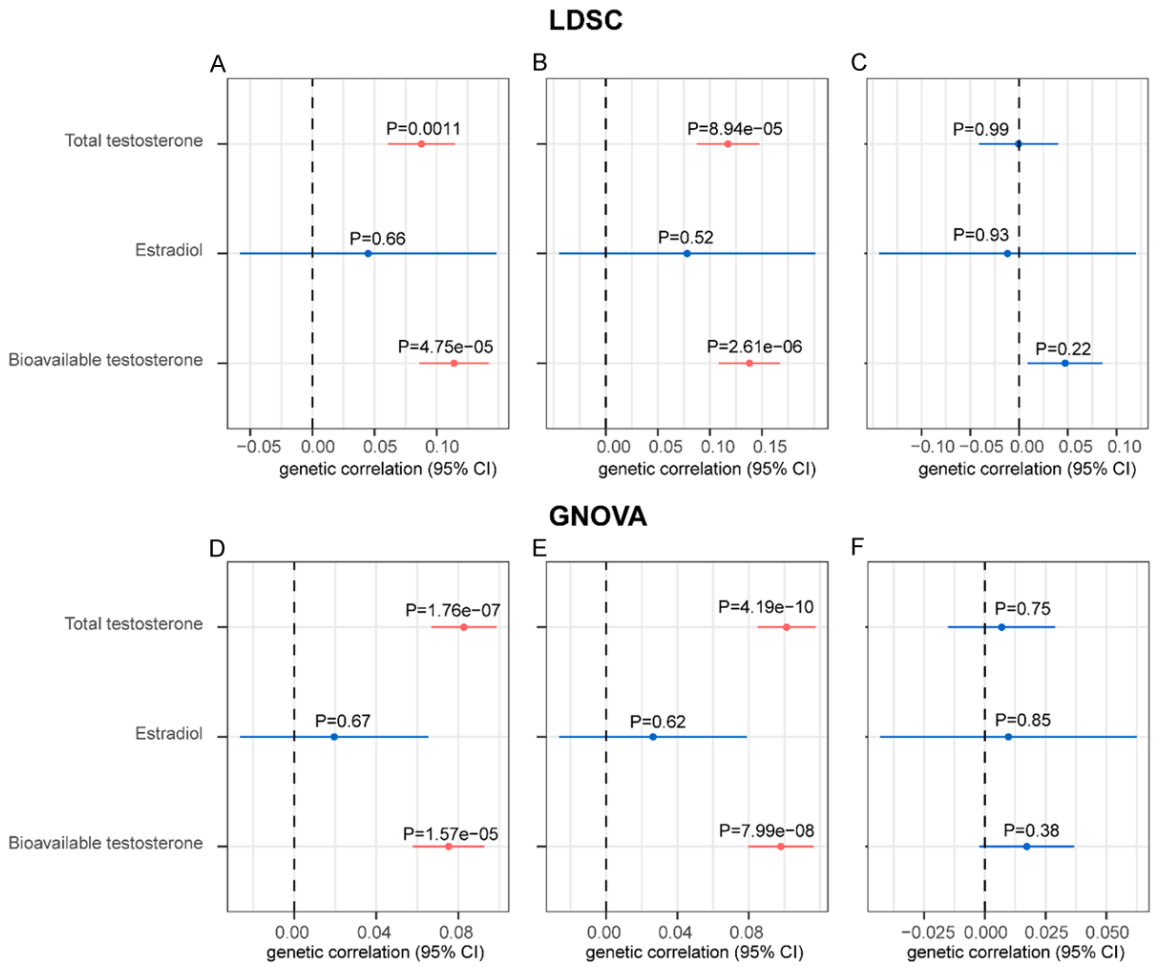


Figure 1. Genetic correlation between sex hormones and breast cancer. (A-C) Genetic correlation estimated using the LDSC method for (A) breast cancer, (B) ER+ breast cancer, and (C) ER- breast cancer. (D-F) Genetic correlation estimated using the GNOVA method for (D) breast cancer, (E) ER+ breast cancer, and (F) ER- breast cancer. Error bars indicate 95% confidence intervals. ER, Estrogen Receptor.

of invalid instrumental variables increases. Furthermore, we used the PhenoScanner v2 tool to check for variants associated with other phenotypes (P<5E-08) which might affect the risk of breast cancer independent of hormone levels [15].

In addition, we conducted comprehensive sensitivity analyses to estimate potential violations of the model assumptions in the MR analysis (Supplementary Figure 2). We conducted Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) analysis to detect outlier instrumental variables, which were removed step-by-step to reduce the effect of horizontal pleiotropy. Cochran's Q test was executed to check heterogeneity across the individual causal effects. MR-Egger regression was performed to evaluate the directional pleiotropy of instrumental variables. To evaluate the strength of each instrumental variable, we computed the F-statistic of each SNP. The statistical power was calculated using an online tool at http://cnsgenomics.com/shiny/mRnd/ [16]. The statistical analyses were conducted using the R package TwoSampleMR 0.5.5 [17].

Results

We first estimated the genetic correlation between each sex hormone and the risk of breast cancer. We detected a significant positive genetic correlation between breast cancer and bioavailable testosterone (genetic correlation: 0.11, P=4.75E-05), total testosterone (genetic correlation: 0.09, P=1.10E-03) (**Figure 1A**, **1D**). In the subtype analysis, similar results were identified between ER+ breast cancer and

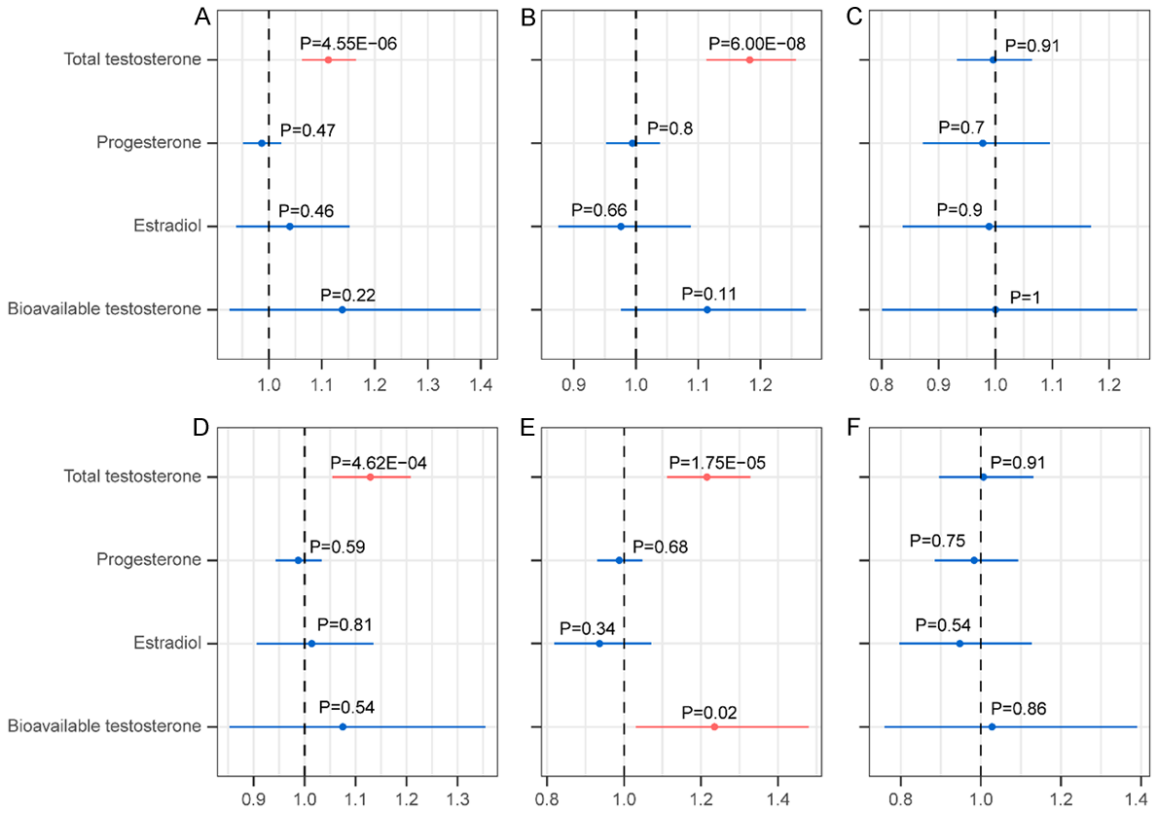


Figure 2. Forest plot showing results from the Mendelian randomization analysis. (A-C) Results from the Mendelian randomization analysis to evaluate causal role of sex hormones in (A) breast cancer, (B) ER+ breast cancer, (C) and ER- breast cancer using the inverse variance weighted method. (D-F) Results from the Mendelian randomization analysis to evaluate causal role of sex hormones in (D) breast cancer, (E) ER+ breast cancer, and (F) ER- breast cancer using the weighted median method. Estimates are per 1 standard deviation (SD) increase in the trait. ER, Estrogen Receptor.

bioavailable testosterone (genetic correlation: 0.14, P=2.61E-06), total testosterone (genetic correlation: 0.12, P=8.94E-05) (**Figure 1B**, **1E**), while for ER- breast cancer no significant association was identified (**Figure 1C**, **1F**).

We further analyzed the role of each hormone in the risk of breast cancer via the two-sample MR approach. Results showed that higher total testosterone level was associated with a higher risk of breast cancer (OR: 1.11, 95% CI: 1.06-1.16, P=4.55E-06) (**Figure 2A**). In the subtype analysis, similar association was detected between total testosterone level and ER+ breast cancer (OR: 1.18, 95% CI: 1.11-1.26, P=6.00E-08), while for ER- breast cancer no association was identified (**Figure 2B**, **2C**). Such results were further verified using the weighted median method (**Figure 2D-F**). The funnel plot displays a symmetric pattern of effect size variation around the point estimate (Supplementary Figures 3, 4, 5, 6, 7, 8, 9, 10, 11).

Furthermore, we performed extensive sensitivity analyses to validate the causal association between sex hormones and breast cancer. The Cochran's Q test did not detect the heterogeneity of effects across the instrumental variables (Supplementary Table 2). The F statistics of all the instrumental variables were above 10 (ranging from 24 to 1656), indicating the absence of weakness in the selected instrumental variables. No apparent horizontal pleiotropy was observed as the intercept of MR-Egger was not significantly deviated from zero. Meanwhile, no potential instrumental outlier was detected by the MR-PRESSO analysis. The leave-one-out results suggested that the causal effect was not driven by a single instrumental variable (Supplementary Figures 3, 4, 5, 6, 7, 8, 9, 10, 11).

Lastly, we used the PhenoScanner tool to check if the SNPs used in the MR analysis were associated with other phenotypes. As a result, several instrumental variables such as rs454-46698, rs112635299 and rs4453027 were associated with body mass index (BMI), which was suggested to affect the risk of breast cancer [18]. Therefore, we further performed multivariable MR analysis to elucidate the causal relationship between sex hormones and the risk of breast cancer adjusting potential pleiotropy due to BMI. The summary data of BMI was obtained from GWAS published by the Genetic Investigation of ANthropometric Traits (GIANT) consortium [19]. As a result, significant association was still identified between testosterone level and breast cancer in the multivariable MR analyses adjusting from BMI (Supplementary Table 3).

Discussion

In the current study, we investigated the causative role of three major hormones in the risk of breast cancer using the MR approach. The results showed that total testosterone level was positively associated with the risk of breast cancer, especially ER+ breast cancer. These findings provided a better understanding of the role of sex hormones in the risk of breast cancer, and had clinical implications.

Testosterone is a male sex hormone, mainly produced in a woman's ovaries in small amounts. Previous epidemiological studies have identified that higher testosterone level was associated with increased risk of breast cancer for women both before and after menopause [5, 20]. Similarly, another prospective cohort study found that estrogen plus testosterone therapies increased risk of invasive breast cancer compared with estrogen-only therapy [21]. These evidence suggested the close correlation between testosterone and breast cancer. Consistent with these findings, we identified that higher testosterone was associated with an increased risk of breast cancer from the genetic perspective using the MR approach. The mechanism of how testosterone increased the risk of breast cancer is still unknown. One explanation is that testosterone can be aromatized to estradiol, which increases proliferation and hence breast cancer risk. Notably, it was also reported that long term therapy with subcutaneous testosterone in women presenting with symptoms of androgen deficiency did not increase the risk of invasive breast cancer [22], and might even reduce the risk of breast cancer [23]. Therefore, moderate levels of testosterone might also be beneficial in breast cancer. However, based on the current dataset, we could not evaluate whether there exists a U-shaped effect of testosterone on the risk of breast cancer. Therefore, further studies investigating testosterone in breast cancer could pay attention to the effect of extreme levels of testosterone. In contrast, we did not identify significant association between bioavailable testosterone and breast cancer, though the effect direction was the same. Simiarly, one previous prospective study identified that total but not free testosterone was positively associatied with the risk of breast cancer [5], though such result was not consistent across studies [24]. Compared with total testosterone which has sex hormone binding globulin or albumin chemical receptors bound to it, unbound testosterone can act as receptors to any cell in the body. Therefore, the bound testosterone might play an important role in the pathogenesis of breast cancer. Nevertheless, we could not rule out the possibility that the failure to detect association might be due to the limited statistical power since the variance explained by the instrumental variables was relatively small (Supplementary Table 2). Future exploration based on summary data from GWAS with larger sample size was warranted to provide a more accurate estimate. In the subtype analysis, testosterone was associated with higher risk of ER+ breast cancer, but not ER- breast cancer. This result suggested the effect of testosterone was lower in ER- breast cancer, which does not have hormone receptors and won't be affected by endocrine treatments aimed at blocking hormones in the body.

Estradiol is a major regulator of growth for the subset of breast cancers that express the estrogen receptor. Previous prospective studies found that estrogens were positively associated with the risk of breast cancer in premenopausal women [5]. However, another cohort study also reported that estrogen was associated with lower incidence of invasive breast cancer among 10,739 postmenopausal women in a median follow-up of 11.8 years [25]. Therefore, the role of estradiol in breast cancer

was still elusive. Biologically, previous clinical findings suggested that after long-term oestrogen deprivation, adaptive changes in mammary tumor gene expression profiles render tumors paradoxically susceptible to oestrogen-induced apoptosis [26, 27]. However, as estrogen is a recognized mitogen that usually stimulates mammary cell proliferation through activation of the oestrogen receptor, too high levels of estrogen might be harmful as well. In the current study, we did not identify association between estradiol level and risk of breast cancer. However, only a few instrumental variables were available for estrodiol in the MR analysis, which limited the statistical power. Therefore, further replication with larger sample size was still necessary.

Progesterone is essential for normal breast development during puberty and in preparation for lactation and breastfeeding. A previous observational study found that estrogen plus progesterone use was associated with increased incidence of breast cancer among 41,449 postmenopausal women (HR=1.55, 95% CI=1.41-1.70, P<0.001) [28]. In the current study, we did not identify causal association between progesterone and breast cancer. This might be due to the limited effect of progesterone on breast cancer. However, we cannot exclude the possibility that we failed to detect association due to the insufficiency of current sample sizes as the effect might be relatively modest. In addition, breast cancer mainly affected women over the age of 50. Subgroup analysis on individuals of different ages might provide additional insight.

In conclusion, our results demonstrated that higher total testosterone level was associated with increased risk of breast cancer, particularly ER+ breast cancer. These findings help better understand the role of hormones in breast cancer, and will facilitate therapeutic management and drug discovery in future clinical trials.

Acknowledgements

We thank the investigators who made the GWAS summary data publicly available and the participants who contributed to those studies. This research was supported by the funding of the Sichuan Science and Technology Program (Grant No. 2021YJ0415). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript.

Disclosure of conflict of interest

None.

Abbreviations

ER, Estrogen Receptor; HER2, Human Epidermal growth factor Receptor 2; GWAS, Genome-Wide Association Study; IVW, Inverse Variance Weighted; LD, Linkage Disequilibrium; MR, Mendelian Randomization; SNP, Single Nucleotide Polymorphism.

Address correspondence to: Huifang Shang and Chunyu Li, Department of Neurology, Laboratory of Neurodegenerative Disorders, National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, No. 37, Guoxue Lane, Chengdu 610041, Sichuan, China. Tel: +86-18980602127; E-mail: hfshang2002@126.com (HFS); Tel: 0086-18980600976; E-mail: lichunyu@scu.edu.cn (CYL)

References

[1] Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, Ruddy K, Tsang J and Cardoso F. Breast cancer. Nat Rev Dis Primers 2019; 5: 66.

[2] Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, Shi W, Jiang J, Yao PP and Zhu HP. Risk factors and preventions of breast cancer. Int J Biol Sci 2017; 13: 1387-1397.

[3] Key T, Appleby P, Barnes I and Reeves G; Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002; 94: 606-616.

[4] Sieri S, Krogh V, Bolelli G, Abagnato CA, Grioni S, Pala V, Evangelista A, Allemani C, Micheli A, Tagliabue G, Schunemann HJ, Menard S, Berrino F and Muti P. Sex hormone levels, breast cancer risk, and cancer receptor status in postmenopausal women: the ORDET cohort. Cancer Epidemiol Biomarkers Prev 2009; 18: 169-176.

[5] Endogenous Hormones and Breast Cancer Collaborative Group, Key TJ, Appleby PN, Reeves GK, Travis RC, Alberg AJ, Barricarte A, Berrino F, Krogh V, Sieri S, Brinton LA, Dorgan JF, Dossus L, Dowsett M, Eliassen AH, Fortner RT, Hankinson SE, Helzlsouer KJ, Hoff man-

Bolton J, Comstock GW, Kaaks R, Kahle LL, Muti P, Overvad K, Peeters PH, Riboli E, Rinaldi S, Rollison DE, Stanczyk FZ, Trichopoulos D, Tworoger SS and Vineis P. Sex hormones and risk of breast cancer in premenopausal women: a collaborative reanalysis of individual participant data from seven prospective studies. Lancet Oncol 2013; 14: 1009-1019.

[6] Moggs JG, Murphy TC, Lim FL, Moore DJ, Stuckey R, Antrobus K, Kimber I and Orphanides G. Anti-proliferative effect of estrogen in breast cancer cells that re-express ERalpha is mediated by aberrant regulation of cell cycle genes. J Mol Endocrinol 2005; 34: 535-551.

[7] Glaser R and Dimitrakakis C. Testosterone and breast cancer prevention. Maturitas 2015; 82: 291-295.

[8] Dimitrakakis C, Zava D, Marinopoulos S, Tsigginou A, Antsaklis A and Glaser R. Low salivary testosterone levels in patients with breast cancer. BMC Cancer 2010; 10: 547.

[9] Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, Beaumont RN, Wittemans L, Martin S, Busch AS, Erzurumluoglu AM, Hollis B, O'Mara TA; Endometrial Cancer Association Consortium; McCarthy MI, Langenberg C, Easton DF, Wareham NJ, Burgess S, Murray A, Ong KK, Frayling TM and Perry JRB. Using human genetics to understand the disease impacts of testosterone in men and women. Nat Med 2020; 26: 252-258.

[10] Schmitz D, Ek WE, Berggren E, Höglund J, Karlsson T and Johansson Å. Genome-wide association study of estradiol levels and the causal effect of estradiol on bone mineral density. J Clin Endocrinol Metab 2021; 106: e4471-e4486.

[11] Pott J, Bae YJ, Horn K, Teren A, Kühnapfel A, Kirsten H, Ceglarek U, Loeffler M, Thiery J, Kratzsch J and Scholz M. Genetic association study of eight steroid hormones and implications for sexual dimorphism of coronary artery disease. J Clin Endocrinol Metab 2019; 104: 5008-5023.

[12] Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, Lemaçon A, Soucy P, Glubb D, Rostamianfar A, Bolla MK, Wang Q, Tyrer J, Dicks E, Lee A, Wang Z, Allen J, Keeman R, Eilber U, French JD, Qing Chen X, Fachal L, McCue K, McCart Reed AE, Ghoussaini M, Carroll JS, Jiang X, Finucane H, Adams M, Adank MA, Ahsan H, Aittomäki K, Anton-Culver H, Antonenkova NN, Arndt V, Aronson KJ, Arun B, Auer PL, Bacot F, Barrdahl M, Baynes C, Beckmann MW, Behrens S, Benitez J, Bermisheva M, Bernstein L, Blomqvist C, Bogdanova NV, Bojesen SE, Bonanni B, Børresen-Dale AL, Brand JS, Brauch H, Brennan P, Brenner H, Brinton L, Broberg P, Brock IW, Broeks A, Brooks-Wilson A, Brucker SY, Brüning T, Burwinkel B, Butterbach K, Cai Q, Cai H, Caldés T, Canzian F, Carracedo A, Carter BD, Castelao JE, Chan TL, David Cheng TY, Seng Chia K, Choi JY, Christiansen H, Clarke CL; NBCS Collaborators; Collée M, Conroy DM, Cordina-Duverger E, Cornelissen S, Cox DG, Cox A, Cross SS, Cunningham JM, Czene K, Daly MB, Devilee P, Doheny KF, Dörk T, Dos-Santos-Silva I, Dumont M, Durcan L, Dwek M, Eccles DM, Ekici AB, Eliassen AH, Ellberg C, Elvira M, Engel C, Eriksson M, Fasching PA, Figueroa J, Flesch-Janys D, Fletcher O, Flyger H, Fritschi L, Gaborieau V, Gabrielson M, Gago-Dominguez M, Gao YT, Gapstur SM, García-Sáenz JA, Gaudet MM, Georgoulias V, Giles GG, Glendon G, Goldberg MS, Goldgar DE, González-Neira A, Grenaker Alnæs GI, Grip M, Gronwald J, Grundy A, Guénel P, Haeberle L, Hahnen E, Haiman CA, Håkansson N, Hamann U, Hamel N, Hankinson S, Harrington P, Hart SN, Hartikainen JM, Hartman M, Hein A, Heyworth J, Hicks B, Hillemanns P, Ho DN, Hollestelle A, Hooning MJ, Hoover RN, Hopper JL, Hou MF, Hsiung CN, Huang G, Humphreys K, Ishiguro J, Ito H, Iwasaki M, Iwata H, Jakubowska A, Janni W, John EM, Johnson N, Jones K, Jones M, Jukkola-Vuorinen A, Kaaks R, Kabisch M, Kaczmarek K, Kang D, Kasuga Y, Kerin MJ, Khan S, Khusnutdinova E, Kiiski JI, Kim SW, Knight JA, Kosma VM, Kristensen VN, Krüger U, Kwong A, Lambrechts D, Le Marchand L, Lee E, Lee MH, Lee JW, Neng Lee C, Lejbkowicz F, Li J, Lilyquist J, Lindblom A, Lissowska J, Lo WY, Loibl S, Long J, Lophatananon A, Lubinski J, Luccarini C, Lux MP, Ma ESK, MacInnis RJ, Maishman T, Makalic E, Malone KE, Kostovska IM, Mannermaa A, Manoukian S, Manson JE, Margolin S, Mariapun S, Martinez ME, Matsuo K, Mavroudis D, McKay J, McLean C, Meijers-Heijboer H, Meindl A, Menéndez P, Menon U, Meyer J, Miao H, Miller N, Taib NAM, Muir K, Mulligan AM, Mulot C, Neuhausen SL, Nevanlinna H, Neven P, Nielsen SF, Noh DY, Nordestgaard BG, Norman A, Olopade OI, Olson JE, Olsson H, Olswold C, Orr N, Pankratz VS, Park SK, Park-Simon TW, Lloyd R, Perez JIA, Peterlongo P, Peto J, Phillips KA, Pinchev M, Plaseska-Karanfilska D, Prentice R, Presneau N, Prokofyeva D, Pugh E, Pylkäs K, Rack B, Radice P, Rahman N, Rennert G, Rennert HS, Rhenius V, Romero A, Romm J, Ruddy KJ, Rüdiger T, Rudolph A, Ruebner M, Rutgers EJT, Saloustros E, Sandler DP, Sangrajrang S, Sawyer EJ, Schmidt DF, Schmutzler RK, Schneeweiss A, Schoemaker MJ, Schumacher F, Schürmann P, Scott RJ, Scott C, Seal S, Seynaeve C, Shah M, Sharma P, Shen CY, Sheng G, Sherman ME, Shrubsole MJ, Shu XO, Smeets A, Sohn C, Southey MC, Spinelli JJ, Stegmaier C, Stewart-Brown S, Stone J, Stram DO, Surowy H, Swerdlow A, Tamimi R, Taylor JA, Tengström M, Teo SH, Beth Terry M, Tessier DC, Thanasitthichai S, Thöne K, Tollenaar

RAEM, Tomlinson I, Tong L, Torres D, Truong T, Tseng CC, Tsugane S, Ulmer HU, Ursin G, Untch M, Vachon C, van Asperen CJ, Van Den Berg D, van den Ouweland AMW, van der Kolk L, van der Luijt RB, Vincent D, Vollenweider J, Waisfisz Q, Wang-Gohrke S, Weinberg CR, Wendt C, Whittemore AS, Wildiers H, Willett W, Winqvist R, Wolk A, Wu AH, Xia L, Yamaji T, Yang XR, Har Yip C, Yoo KY, Yu JC, Zheng W, Zheng Y, Zhu B, Ziogas A, Ziv E; ABCTB Investigators; ConFab/AOCS Investigators; Lakhani SR, Antoniou AC, Droit A, Andrulis IL, Amos CI, Couch FJ, Pharoah PDP, Chang-Claude J, Hall P, Hunter DJ, Milne RL, García-Closas M, Schmidt MK, Chanock SJ, Dunning AM, Edwards SL, Bader GD, Chenevix-Trench G, Simard J, Kraft P and Easton DF. Association analysis identifies 65 new breast cancer risk loci. Nature 2017; 551: 92-94.

[13] Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, Duncan L, Perry JR, Patterson N, Robinson EB, Daly MJ, Price AL and Neale BM. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015; 47: 1236-1241.

[14] Lu Q, Li B, Ou D, Erlendsdottir M, Powles RL, Jiang T, Hu Y, Chang D, Jin C, Dai W, He Q, Liu Z, Mukherjee S, Crane PK and Zhao H. A powerful approach to estimating annotation-stratified genetic covariance via GWAS summary statistics. Am J Hum Genet 2017; 101: 939-964.

[15] Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS and Staley JR. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019; 35: 4851-4853.

[16] Brion MJ, Shakhbazov K and Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol 2013; 42: 1497-1501.

[17] Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G, Gaunt TR and Haycock PC. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018; 7: e34408.

[18] Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER and Slingerland JM. Obesity and adverse breast cancer risk and outcome: mechanistic insights and strategies for intervention. CA Cancer J Clin 2017; 67: 378-397.

[19] Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, Croteau-Chonka DC, Esko T, Fall T, Ferreira T, Gustafsson S, Kutalik Z, Luan J, Mägi R, Randall JC, Winkler TW, Wood AR, Workalemahu T, Faul JD, Smith JA, Zhao JH, Zhao W, Chen J, Fehrmann R, Hedman ÅK, Karjalainen J, Schmidt EM, Absher D, Amin N, Anderson D, Beekman M, Bolton JL, Bragg-Gresham JL, Buyske S, Demirkan A, Deng G, Ehret GB, Feenstra B, Feitosa MF, Fischer K, Goel A, Gong J, Jackson AU, Kanoni S, Kleber ME, Kristiansson K, Lim U, Lotay V, Mangino M, Leach IM, Medina-Gomez C, Medland SE, Nalls MA, Palmer CD, Pasko D, Pechlivanis S, Peters MJ, Prokopenko I, Shungin D, Stančáková A, Strawbridge RJ, Sung YJ, Tanaka T, Teumer A, Trompet S, van der Laan SW, van Setten J, Van Vliet-Ostaptchouk JV, Wang Z, Yengo L, Zhang W, Isaacs A, Albrecht E, Ärnlöv J, Arscott GM, Attwood AP, Bandinelli S, Barrett A, Bas IN, Bellis C, Bennett AJ, Berne C, Blagieva R, Blüher M, Böhringer S, Bonnycastle LL, Böttcher Y, Boyd HA, Bruinenberg M, Caspersen IH, Chen YI, Clarke R, Daw EW, de Craen AJM, Delgado G, Dimitriou M, Doney ASF, Eklund N, Estrada K, Eury E, Folkersen L, Fraser RM, Garcia ME, Geller F, Giedraitis V, Gigante B, Go AS, Golay A, Goodall AH, Gordon SD, Gorski M, Grabe HJ, Grallert H, Grammer TB, Gräßler J, Grönberg H, Groves CJ, Gusto G, Haessler J, Hall P, Haller T, Hallmans G, Hartman CA, Hassinen M, Hayward C, Heard-Costa NL, Helmer Q, Hengstenberg C, Holmen O, Hottenga JJ, James AL, Jeff JM, Johansson Å, Jolley J, Juliusdottir T, Kinnunen L, Koenig W, Koskenvuo M, Kratzer W, Laitinen J, Lamina C, Leander K, Lee NR, Lichtner P, Lind L, Lindström J, Lo KS, Lobbens S, Lorbeer R, Lu Y, Mach F, Magnusson PKE, Mahajan A, McArdle WL, McLachlan S, Menni C, Merger S, Mihailov E, Milani L, Moayyeri A, Monda KL, Morken MA, Mulas A, Müller G, Müller-Nurasyid M, Musk AW, Nagaraja R, Nöthen MM, Nolte IM, Pilz S, Rayner NW, Renstrom F, Rettig R, Ried JS, Ripke S, Robertson NR, Rose LM, Sanna S, Scharnagl H, Scholtens S, Schumacher FR, Scott WR, Seufferlein T, Shi J, Smith AV, Smolonska J, Stanton AV, Steinthorsdottir V, Stirrups K, Stringham HM, Sundström J, Swertz MA, Swift AJ, Syvänen AC, Tan ST, Tayo BO, Thorand B, Thorleifsson G, Tyrer JP, Uh HW, Vandenput L, Verhulst FC, Vermeulen SH, Verweij N, Vonk JM, Waite LL, Warren HR, Waterworth D, Weedon MN, Wilkens LR, Willenborg C, Wilsgaard T, Wojczynski MK, Wong A, Wright AF, Zhang Q; LifeLines Cohort Study; Brennan EP, Choi M, Dastani Z, Drong AW, Eriksson P, Franco-Cereceda A, Gådin JR, Gharavi AG, Goddard ME, Handsaker RE, Huang J, Karpe F, Kathiresan S, Keildson S, Kiryluk K, Kubo M, Lee JY, Liang L, Lifton RP, Ma B, McCarroll SA, McKnight AJ, Min JL, Moffatt MF, Montgomery GW, Murabito JM, Nicholson G, Nyholt DR, Okada Y, Perry JRB, Dorajoo R, Reinmaa E, Salem RM, Sandholm N, Scott RA, Stolk L, Takahashi A, Tanaka T, van 't Hooft FM, Vinkhuyzen AAE, Westra HJ, Zheng W,

Zondervan KT; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium; Heath AC, Arveiler D, Bakker SJL, Beilby J, Bergman RN, Blangero J, Bovet P, Campbell H, Caulfield MJ, Cesana G, Chakravarti A, Chasman DI, Chines PS, Collins FS, Crawford DC, Cupples LA, Cusi D, Danesh J, de Faire U, den Ruijter HM, Dominiczak AF, Erbel R, Erdmann J, Eriksson JG, Farrall M, Felix SB, Ferrannini E, Ferrières J, Ford I, Forouhi NG, Forrester T, Franco OH, Gansevoort RT, Gejman PV, Gieger C, Gottesman O, Gudnason V, Gyllensten U, Hall AS, Harris TB, Hattersley AT, Hicks AA, Hindorff LA, Hingorani AD, Hofman A, Homuth G, Hovingh GK, Humphries SE, Hunt SC, Hyppönen E, Illig T, Jacobs KB, Jarvelin MR, Jöckel KH, Johansen B, Jousilahti P, Jukema JW, Jula AM, Kaprio J, Kastelein JJP, Keinanen-Kiukaanniemi SM, Kiemeney LA, Knekt P, Kooner JS, Kooperberg C, Kovacs P, Kraja AT, Kumari M, Kuusisto J, Lakka TA, Langenberg C, Marchand LL, Lehtimäki T, Lyssenko V, Männistö S, Marette A, Matise TC, McKenzie CA, McKnight B, Moll FL, Morris AD, Morris AP, Murray JC, Nelis M, Ohlsson C, Oldehinkel AJ, Ong KK, Madden PAF, Pasterkamp G, Peden JF, Peters A, Postma DS, Pramstaller PP, Price JF, Qi L, Raitakari OT, Rankinen T, Rao DC, Rice TK, Ridker PM, Rioux JD, Ritchie MD, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA, Schunkert H, Schwarz PEH, Sever P, Shuldiner AR, Sinisalo J, Stolk RP, Strauch K, Tönjes A, Trégouët DA, Tremblay A, Tremoli E, Virtamo J, Vohl MC, Völker U, Waeber G, Willemsen G, Witteman JC, Zillikens MC, Adair LS, Amouyel P, Asselbergs FW, Assimes TL, Bochud M, Boehm BO, Boerwinkle E, Bornstein SR, Bottinger EP, Bouchard C, Cauchi S, Chambers JC, Chanock SJ, Cooper RS, de Bakker PIW, Dedoussis G, Ferrucci L, Franks PW, Froguel P, Groop LC, Haiman CA, Hamsten A, Hui J, Hunter DJ, Hveem K, Kaplan RC, Kivimaki M, Kuh D, Laakso M, Liu Y, Martin NG, März W, Melbye M, Metspalu A, Moebus S, Munroe PB, Njølstad I, Oostra BA, Palmer CNA, Pedersen NL, Perola M, Pérusse L, Peters U, Power C, Quertermous T, Rauramaa R, Rivadeneira F, Saaristo TE, Saleheen D, Sattar N, Schadt EE, Schlessinger D, Slagboom PE, Snieder H, Spector TD, Thorsteinsdottir U, Stumvoll M, Tuomilehto J, Uitterlinden AG, Uusitupa M, van der Harst P, Walker M, Wallaschofski H, Wareham NJ, Watkins H, Weir DR, Wichmann HE, Wilson JF, Zanen P, Borecki IB, Deloukas P, Fox CS, Heid IM, O'Connell JR, Strachan DP, Stefansson K, van Duijn CM, Abecasis GR, Franke L, Frayling TM, McCarthy MI, Visscher PM, Scherag A, Willer CJ, Boehnke M, Mohlke KL, Lindgren CM, Beckmann JS, Barroso I, North KE, Ingelsson E, Hirschhorn JN, Loos RJF and Speliotes EK. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015; 518: 197-206.

[20] Key T, Appleby P, Barnes I and Reeves G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002; 94: 606-616.

[21] Tamimi RM, Hankinson SE, Chen WY, Rosner B and Colditz GA. Combined estrogen and testosterone use and risk of breast cancer in postmenopausal women. Arch Intern Med 2006; 166: 1483-1489.

[22] Glaser RL, York AE and Dimitrakakis C. Incidence of invasive breast cancer in women treated with testosterone implants: a prospective 10-year cohort study. BMC Cancer 2019; 19: 1271.

[23] Donovitz G and Cotten M. Breast cancer incidence reduction in women treated with subcutaneous testosterone: testosterone therapy and breast cancer incidence study. Eur J Breast Health 2021; 17: 150-156.

[24] Schernhammer ES, Sperati F, Razavi P, Agnoli C, Sieri S, Berrino F, Krogh V, Abbagnato C, Grioni S, Blandino G, Schunemann HJ and Muti P. Endogenous sex steroids in premenopausal women and risk of breast cancer: the ORDET cohort. Breast Cancer Res 2013; 15: R46.

[25] Anderson GL, Chlebowski RT, Aragaki AK, Kuller LH, Manson JE, Gass M, Bluhm E, Connelly S, Hubbell FA, Lane D, Martin L, Ockene J, Rohan T, Schenken R and Wactawski-Wende J. Conjugated equine oestrogen and breast cancer incidence and mortality in postmenopausal women with hysterectomy: extended follow-up of the Women's Health Initiative randomised placebo-controlled trial. Lancet Oncol 2012; 13: 476-486.

[26] Jordan VC. The 38th David A. Karnofsky lecture: the paradoxical actions of estrogen in breast cancer-survival or death? J Clin Oncol 2008; 26: 3073-3082.

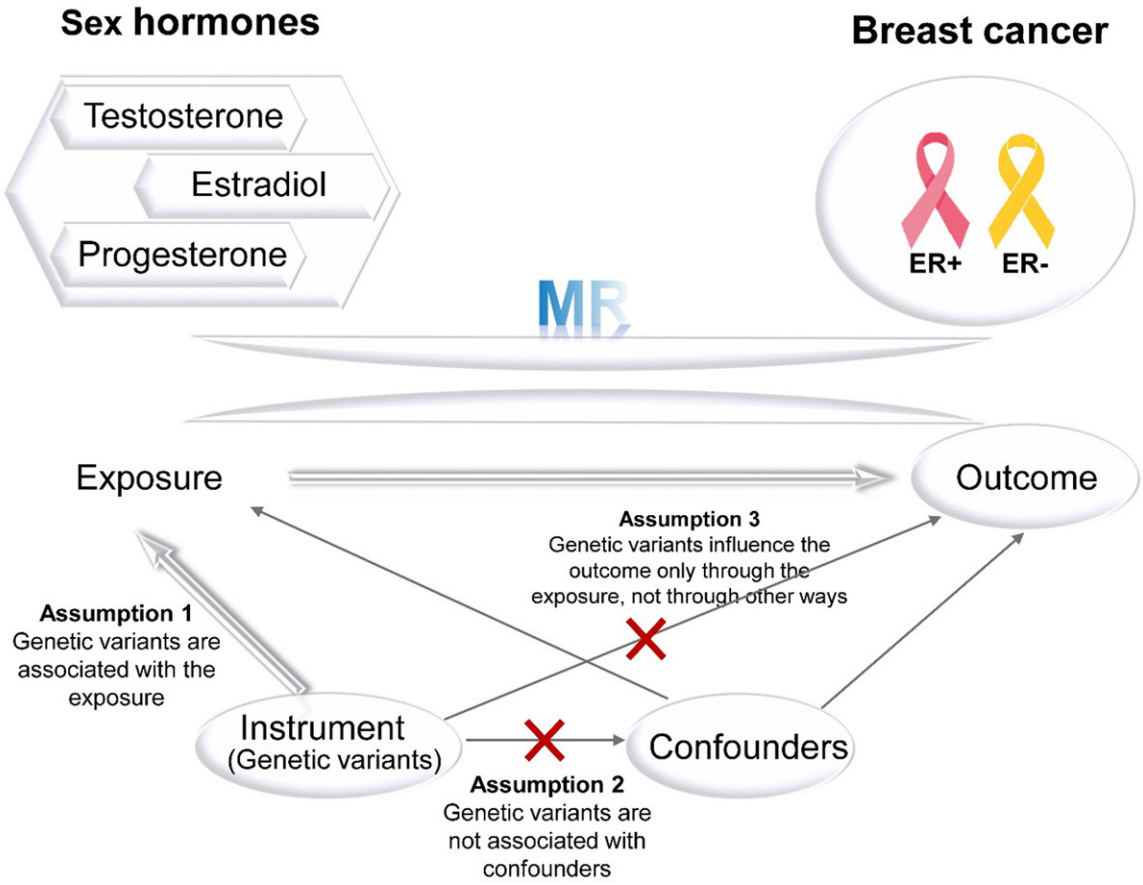
[27] Lewis-Wambi JS and Jordan VC. Estrogen regulation of apoptosis: how can one hormone stimulate and inhibit? Breast Cancer Res 2009; 11: 206.

[28] Chlebowski RT, Manson JE, Anderson GL, Cauley JA, Aragaki AK, Stefanick ML, Lane DS, Johnson KC, Wactawski-Wende J, Chen C, Qi L, Yasmeen S, Newcomb PA and Prentice RL. Estrogen plus progestin and breast cancer incidence and mortality in the Women's Health Initiative Observational Study. J Natl Cancer Inst 2013; 105: 526-535.

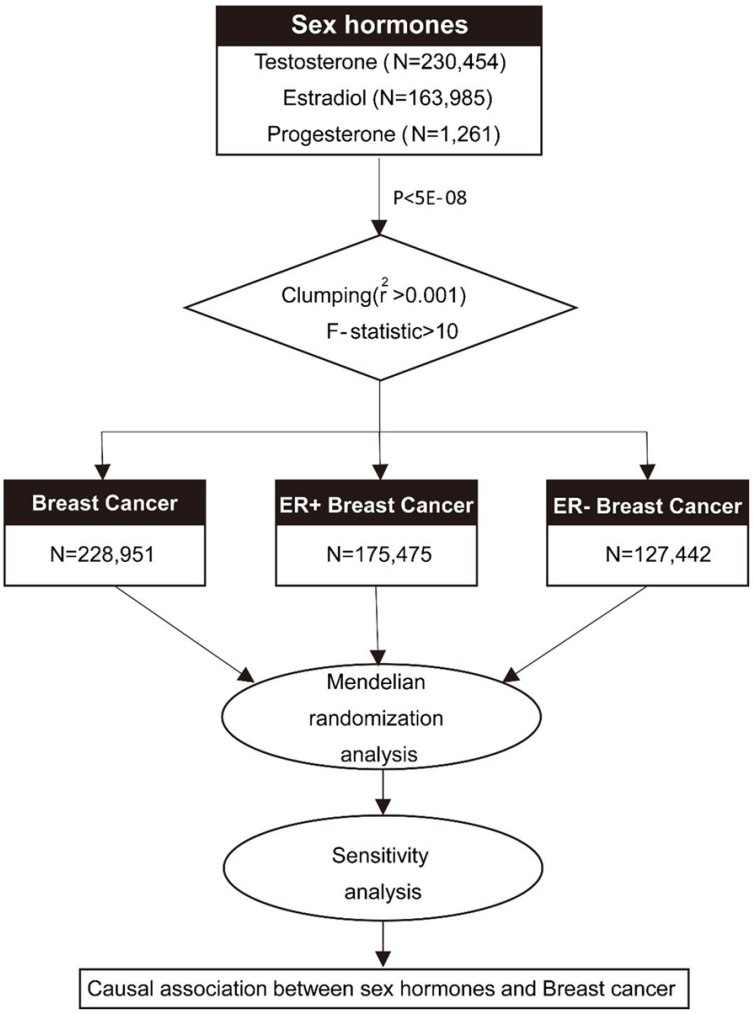
Supplementary Table 1. Summary data from all GWAS used in current study

Phenotype	Cases	Controls	Number of SNPs	PMID
bioavailable testosterone	188,507	-	16,585,745	32042192
total testosterone	230,454	-	16,580,850	32042192
estradiol	37,461	126,524	7,870,546	34255042
progesterone	1,261	-	n.a.	31169883
breast cancer	122,977	105,974	11,792,542	29059683
ER+ breast cancer	69,501	105,974	10,643,737	29059683
ER- breast cancer	21,468	105,974	10,643,737	29059683

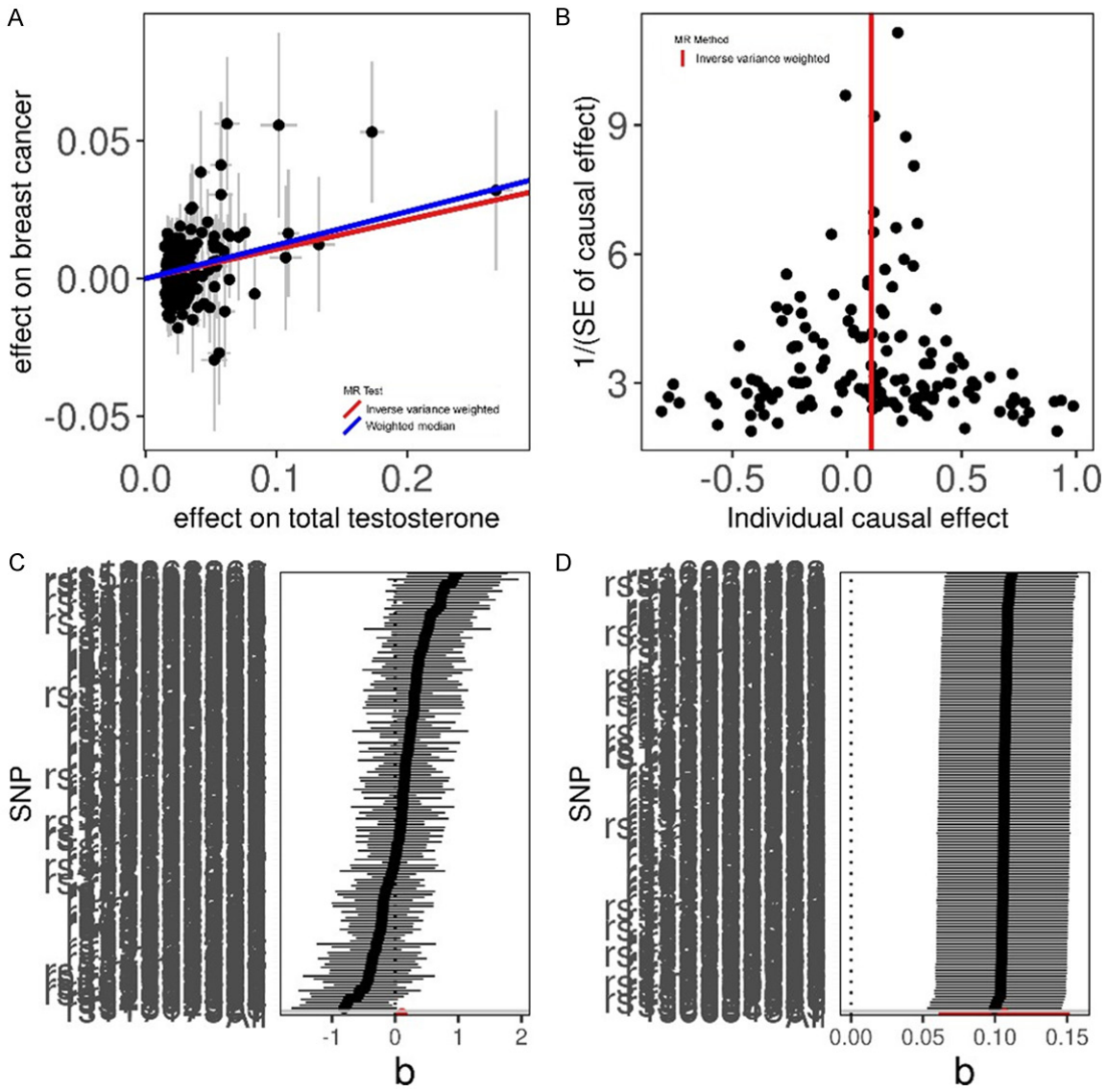
SNP, single nucleotide polymorphism; GWAS, genome-wide association study; PMID, PubMed ID; ER+, estrogen receptor positive; ER-, estrogen receptor negative; n.a., not available.



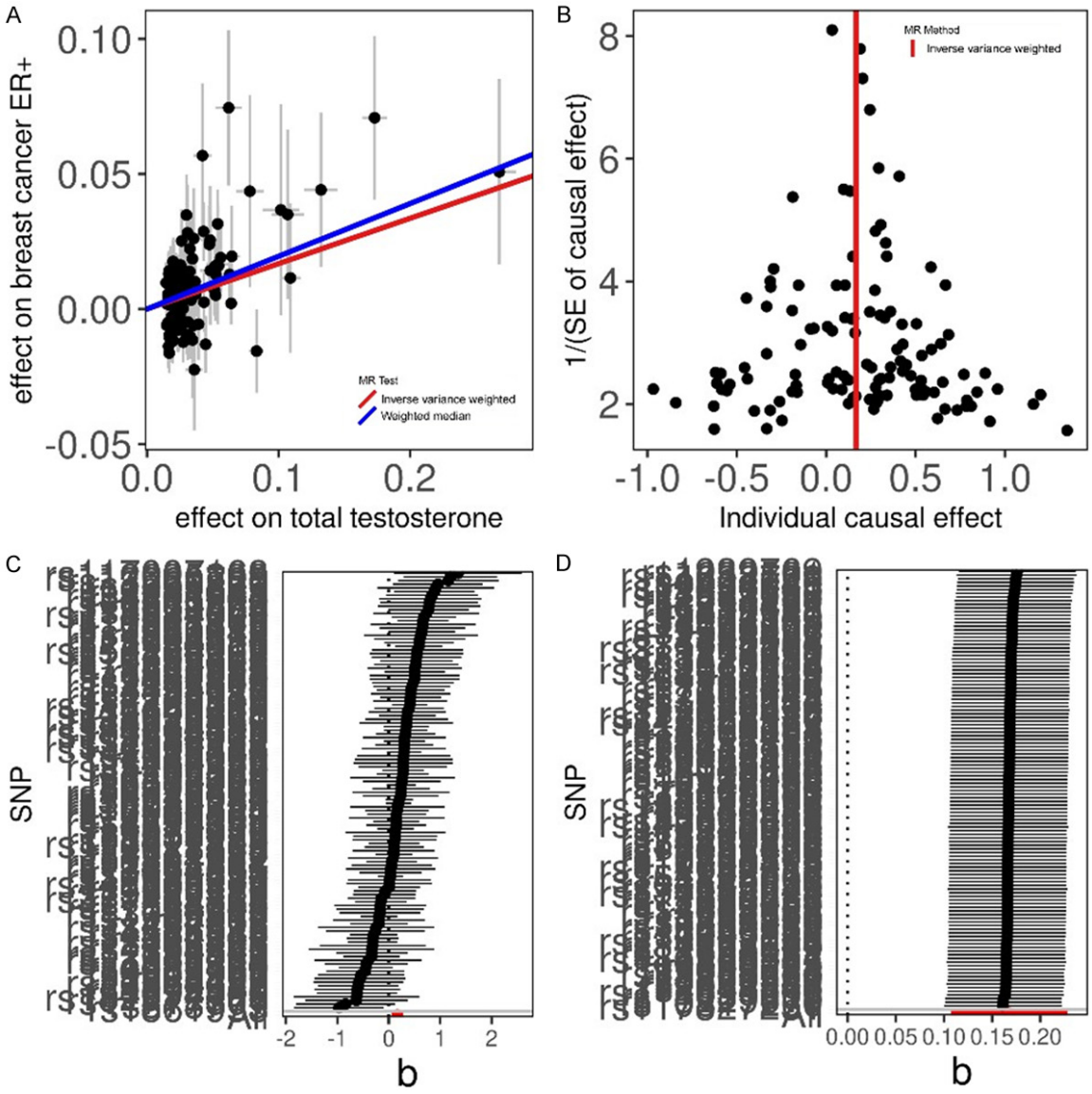
Supplementary Figure 1. Assumptions in Mendelian randomization analysis. Broken lines represent potential pleiotropic or direct causal effects between variables that would violate Mendelian randomization assumptions.



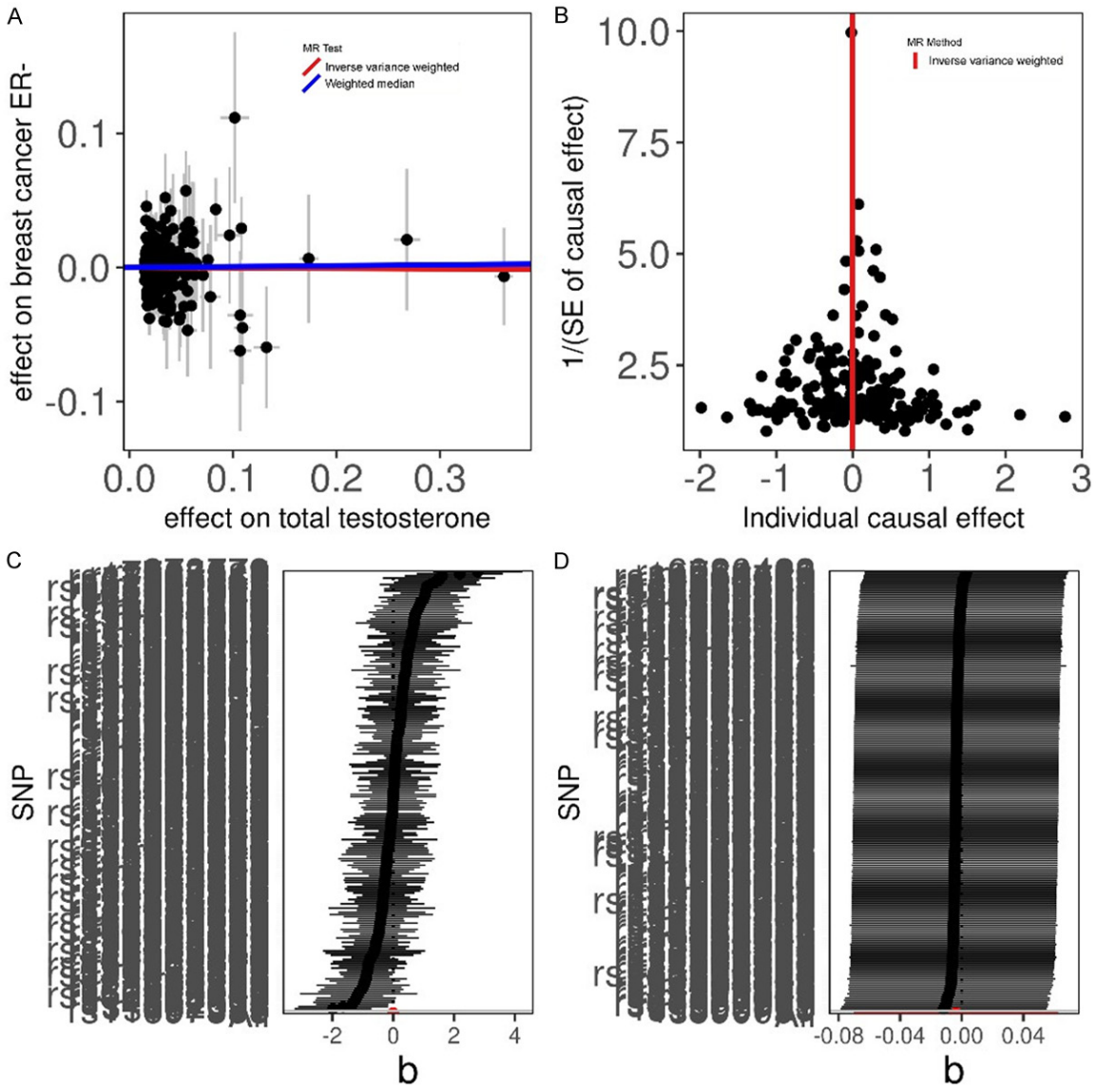
Supplementary Figure 2. Schematic analysis workflow.



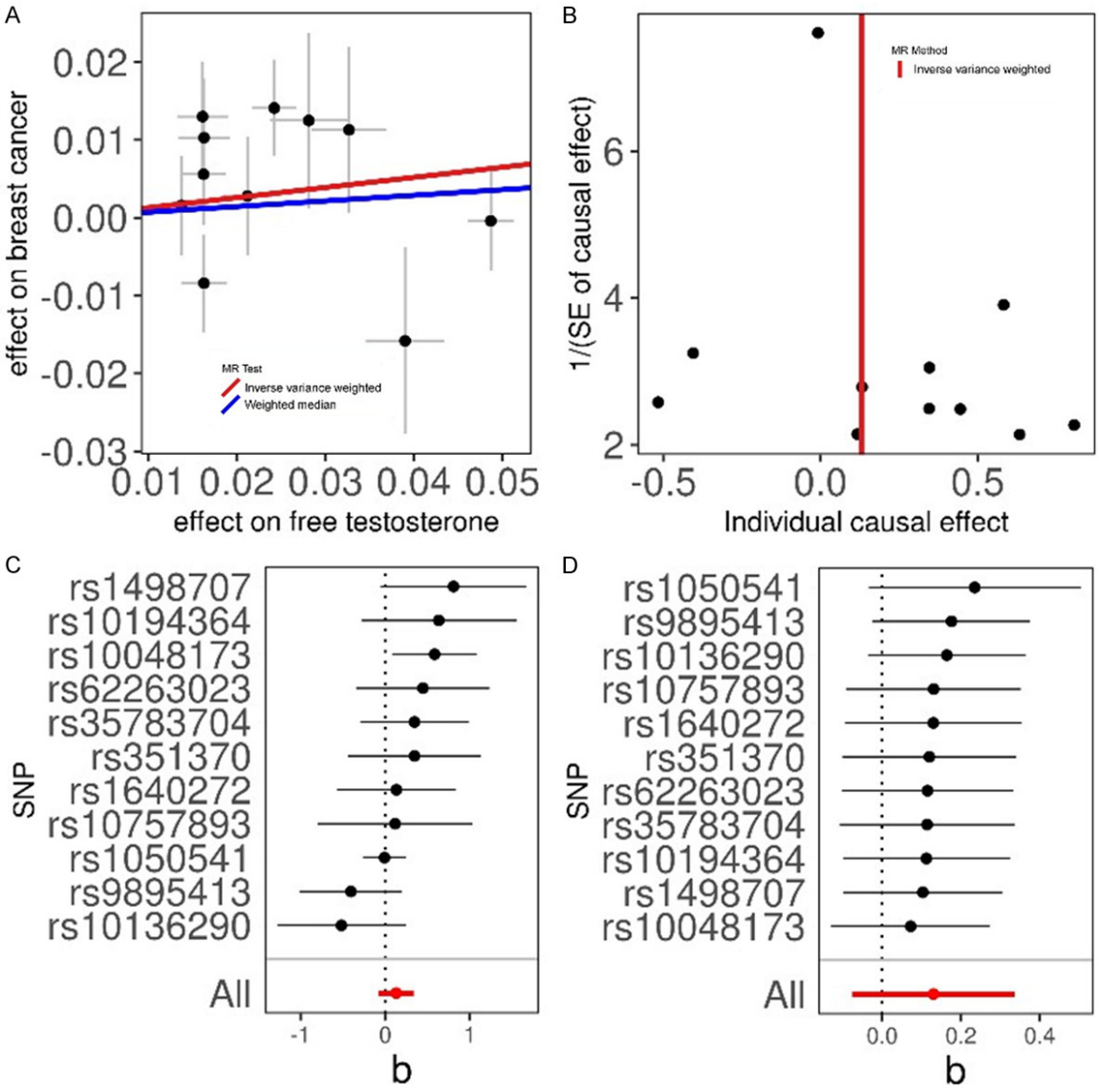
Supplementary Figure 3. Mendelian randomization analysis results for total testosterone on risk of breast cancer. A. Scatter plot of SNP effects on total testosterone and breast cancer. The 95% CI for the effect size on breast cancer is shown as vertical lines, while the 95% CI for the effect size on total testosterone is shown as horizontal lines. The slope of fitted lines represents the estimated MR effect per method. B. Funnel plot showing the estimation using the inverse of the standard error of the causal estimate with each individual SNP as a tool. The vertical line represents the estimated causal effect. C. Forest plot of the association of individual SNPs with total testosterone and breast cancer, together with pooled estimates. D. Forest plot of the results of the leave-one-out sensitivity analysis, where each SNP was iteratively removed from the instrumental variables. SNP, single nucleotide polymorphism.



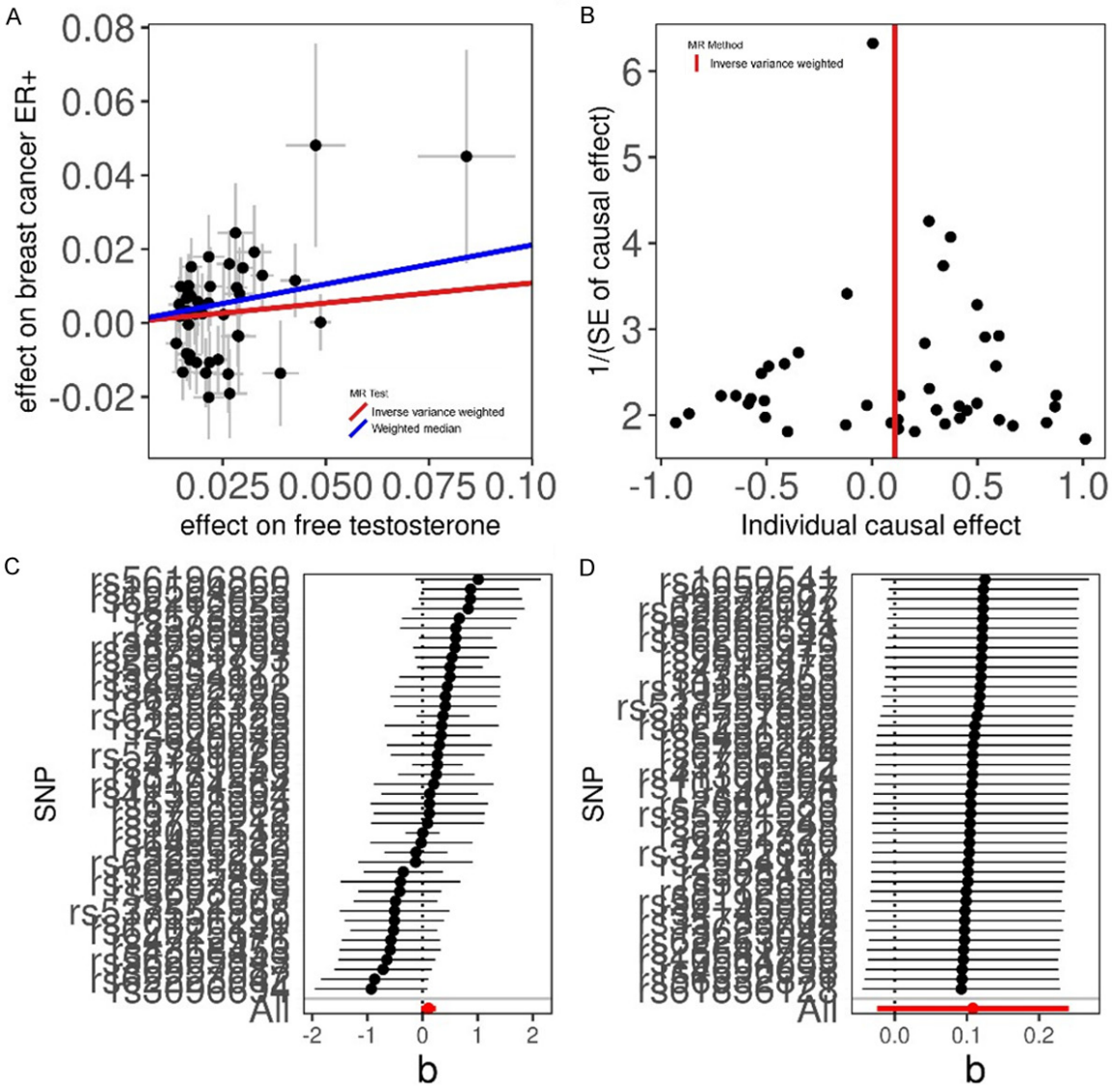
Supplementary Figure 4. Mendelian randomization analysis results for total testosterone on risk of ER+ breast cancer.



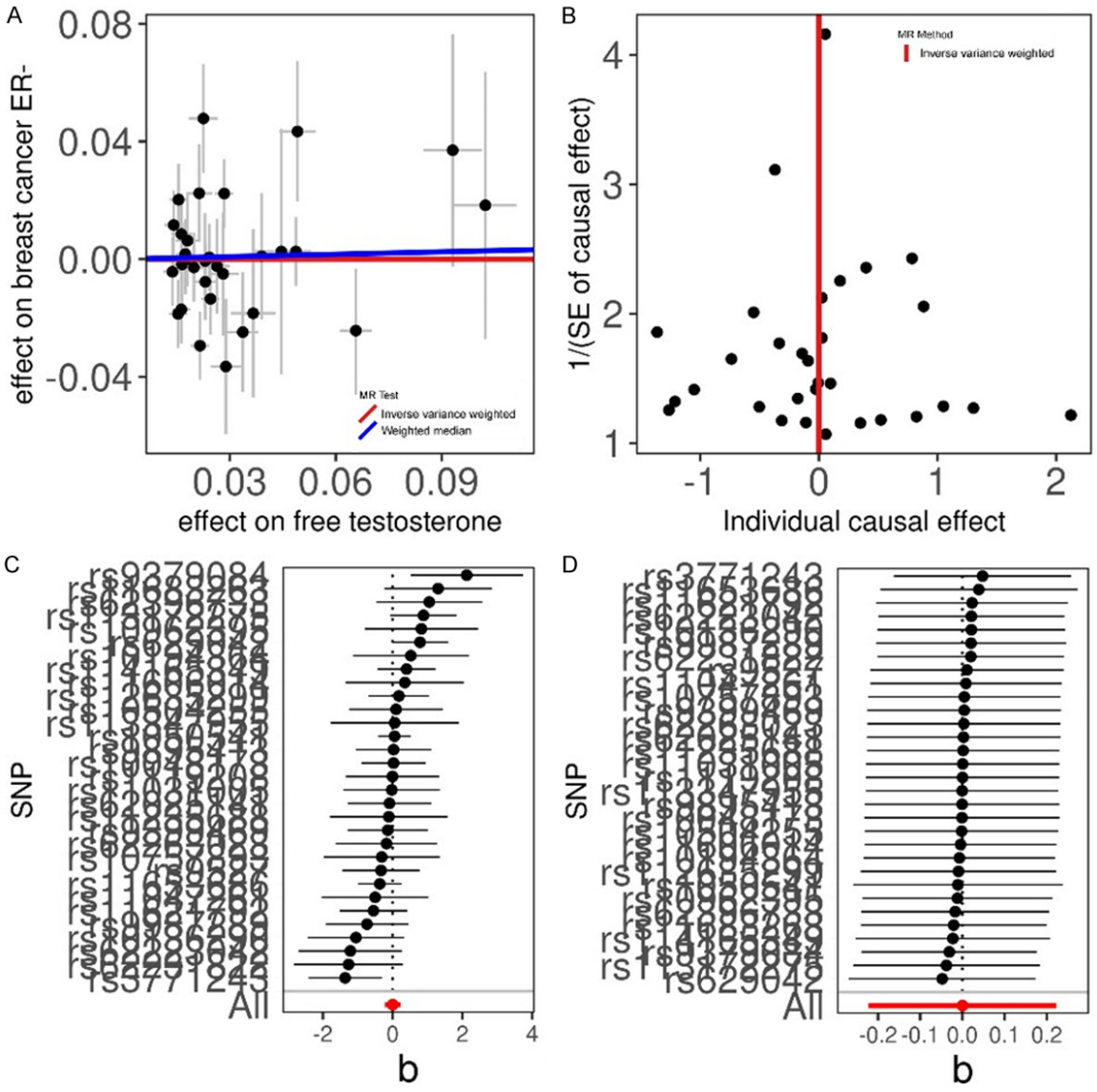
Supplementary Figure 5. Mendelian randomization analysis results for total testosterone on risk of ER- breast cancer.



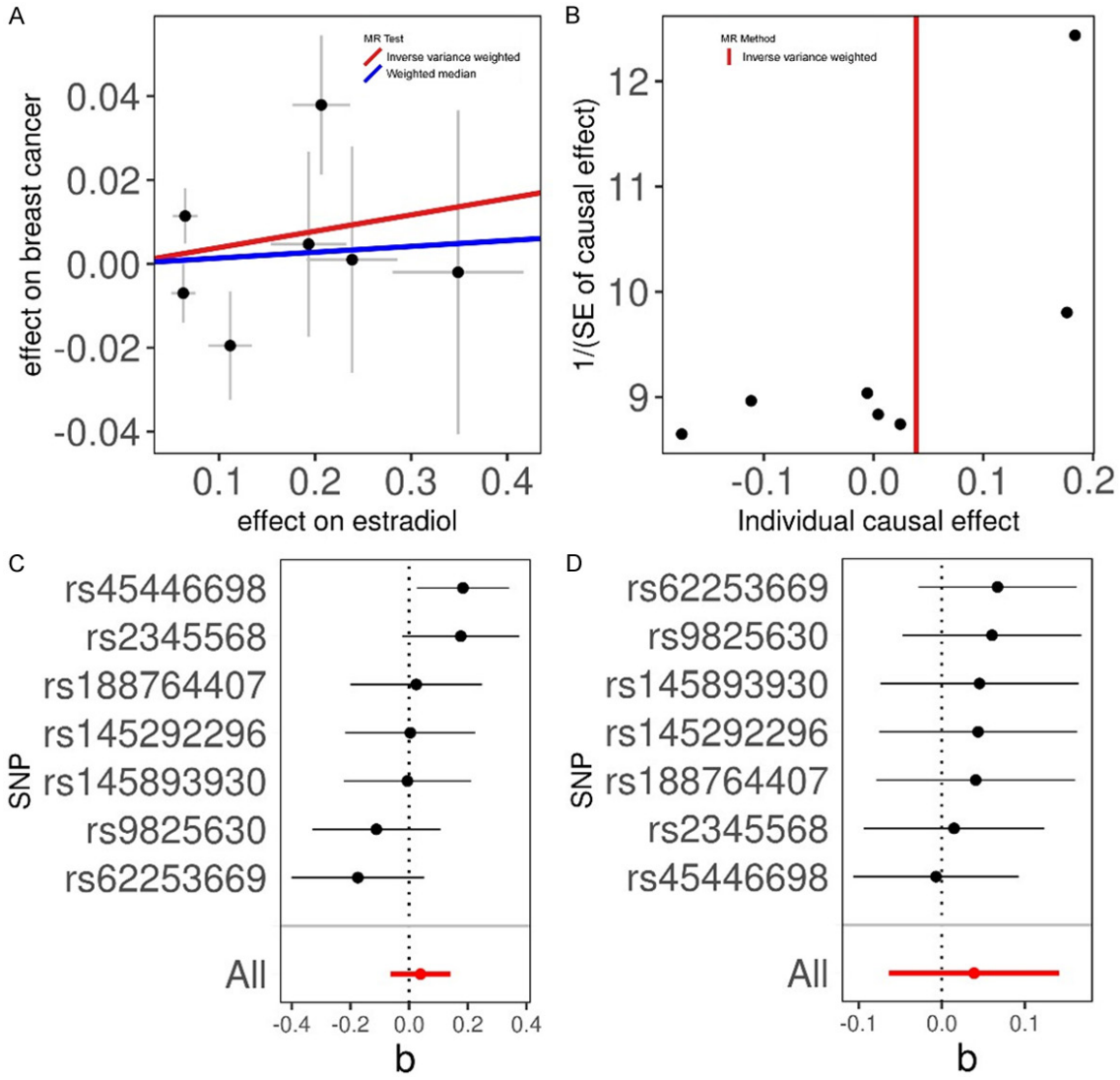
Supplementary Figure 6. Mendelian randomization analysis results for free testosterone on risk of breast cancer.



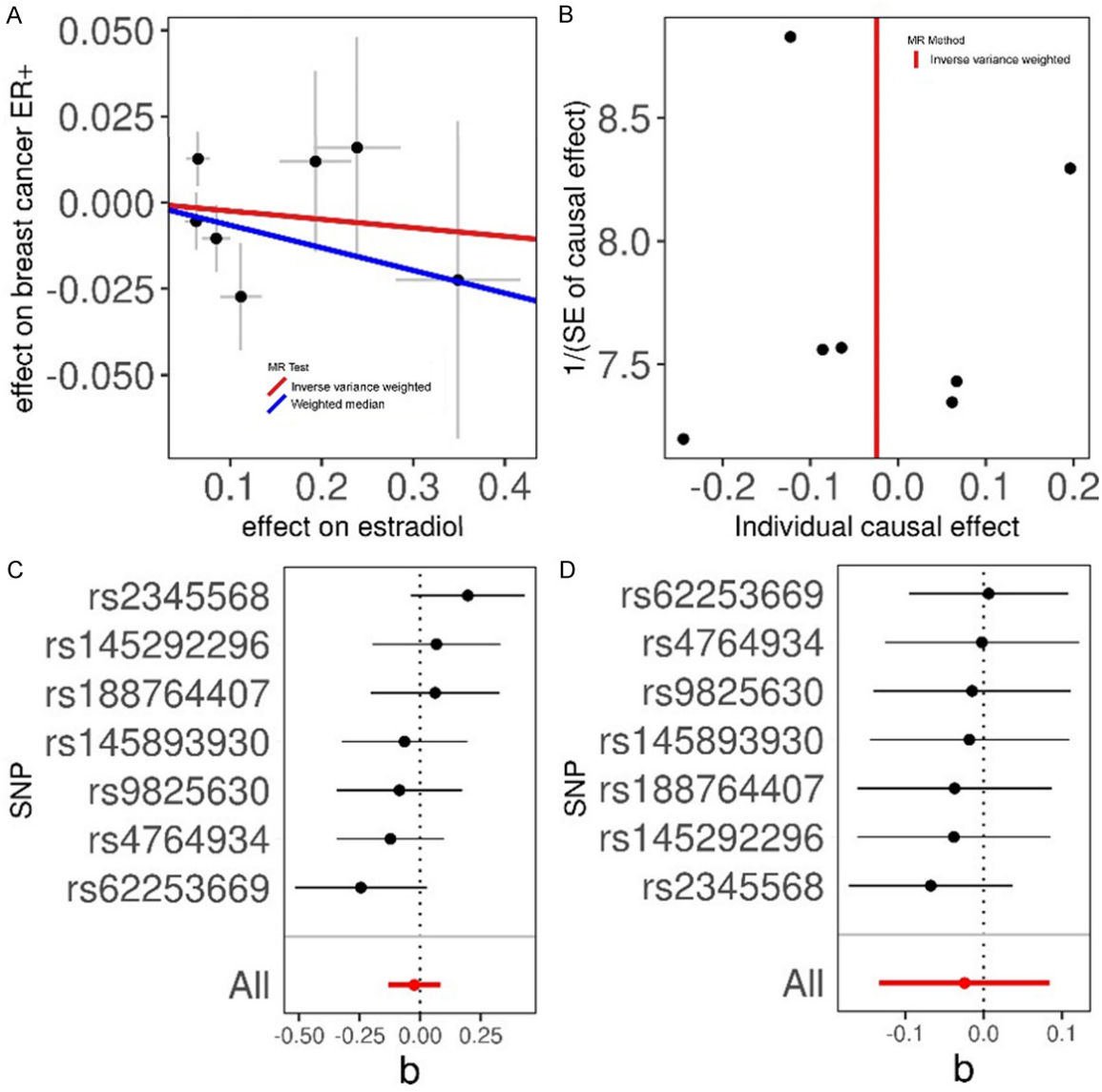
Supplementary Figure 7. Mendelian randomization analysis results for free testosterone on risk of ER+ breast cancer.



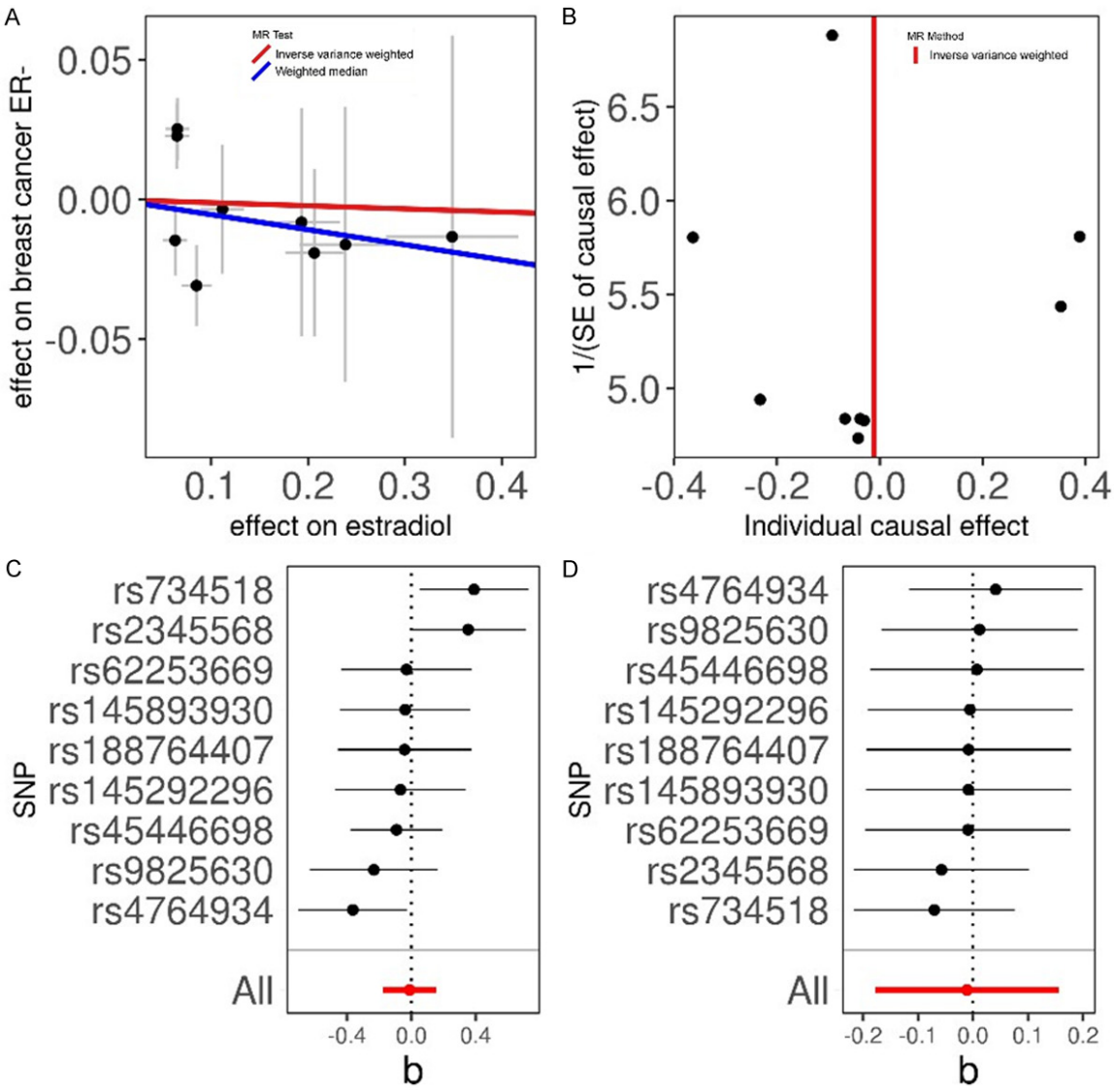
Supplementary Figure 8. Mendelian randomization analysis results for free testosterone on risk of ER- breast cancer.



Supplementary Figure 9. Mendelian randomization analysis results for estradiol on risk of breast cancer.



Supplementary Figure 10. Mendelian randomization analysis results for estradiol on risk of ER+ breast cancer.



Supplementary Figure 11. Mendelian randomization analysis results for estradiol on risk of ER- breast cancer.

Supplementary Table 2. Heterogeneity and horizontal pleiotropy analyses between hormone traits and breast cancer

hormone trait	Heterogeneity			Horizontal pleiotropy			MR-PRESSO	
	IVW Q	IVW Q df	IVW P	Egger intercept	SE	*P* value	*P* value	Beta
breast cancer as outcome								
bioavailable								
testosterone	14.89	10	0.14	7.41E-03	6.15E-03	0.26	0.12	0.17
total testosterone	178.32	149	0.05	-2.17E-03	1.46E-03	0.14	0.05	0.05
oestradiol	10.56	6	0.10	-3.23E-03	1.14E-02	0.79	0.09	0.34
progesterone	1.46	3	0.69	-6.11E-03	2.23E-02	0.81	0.71	0.07
ER+ breast cancer as outcome								
bioavailable								
testosterone	55.72	42	0.08	-4.69E-03	4.37E-03	0.29	0.09	0.11
total testosterone	143.26	119	0.06	-3.01E-03	1.92E-03	0.12	0.05	0.06
oestradiol	7.79	6	0.25	2.40E-03	1.22E-02	0.85	0.25	0.39
progesterone	0.88	3	0.83	-4.80E-03	2.66E-02	0.87	0.84	0.08
ER- breast cancer as outcome								
bioavailable								
testosterone	39.62	30	0.11	-2.99E-03	6.70E-03	0.66	0.12	0.18
total testosterone	240.04	204	0.04	-7.80E-04	2.06E-03	0.71	0.05	0.08
oestradiol	15.12	8	0.06	1.73E-02	1.69E-02	0.34	0.04	0.45
progesterone	8.69	3	0.03	3.80E-03	8.52E-02	0.97	0.05	0.12

IVW, Inverse variance weighted; Q, Cochran's Q test estimate; df, Cochran's Q test degrees of freedom; SE, standard error. Beta denotes the effect sizes can be detected with the power of 0.8 given the sample size, proportion of cases and variance explained by the instrumental variables.

Supplementary Table 3. Mendelian randomization estimates between sex hormone level and risk of breast cancer adjusting for body mass index

outcome	exposure	beta	SE	*P* value
Breast cancer	Bioavailable testosterone	0.13	0.06	0.02
	Total testosterone	0.13	0.03	1.24E-05
	Estradiol	0.05	0.50	0.93
ER+ breast cancer	Bioavailable testosterone	0.19	0.06	8.23E-04
	Total testosterone	0.17	0.03	1.65E-07
	Estradiol	0.18	0.60	0.76
ER- breast cancer	Bioavailable testosterone	-0.06	0.08	0.41
	Total testosterone	-0.02	0.4	0.61
	Estradiol	-0.49	0.61	0.42

ER+, estrogen receptor positive; ER-, estrogen receptor negative.