

Title page

Type of manuscript: Full length article

Genetically predicted 17beta-estradiol and cardiovascular risk factors in women: a Mendelian randomization analysis using young women in Hong Kong and older women in the Guangzhou Biobank Cohort Study

Shiu Lun Au Yeung PhD^a, Kar Keung Cheng PhD^b, Jie Zhao PhD,^a Weisen Zhang MD^c, Chaoqiang Jiang MD^{c,†}, Tai Hing Lam MD^a, Gabriel Matthew Leung MD^a, C Mary Schooling PhD^{a,d}

^aSchool of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

^bDepartment of Public Health and Epidemiology, University of Birmingham, UK

^cGuangzhou Number 12 Hospital, Guangzhou, China

^dCity University of New York School of Public Health and Hunter College, New York, NY, USA

†Corresponding author:

Prof C Jiang

Molecular Epidemiological Research Centre

Guangzhou No. 12 Hospital

Guangzhou, China

Email: cqjiang@hkucc.hku.hk

Abstract: 198

Main text: 2,716

Tables: 3

Figure: 1

Running title: Estrogen and CVD risk factors using MR analysis

Keywords: estrogen; cardiovascular risk factors; Mendelian randomization analysis; Chinese; women

List of abbreviations

BMI: Body mass index

DNA: Deoxyribonucleic acid

HDL cholesterol: High density lipoprotein cholesterol

HEPA: Health-enhancing physical activity

HRT: Hormone replacement therapy

GBCS: Guangzhou Biobank Cohort Study

GHHARE: The Guangzhou Health and Happiness Association for the Respectable Elders

GWAS: Genome wide association studies

IPAQ: International Physical Activity Questionnaire

LDL cholesterol: Low density lipoprotein cholesterol

MET: Metabolic equivalent

RCT: Randomized controlled trial

SNP: Single nucleotide polymorphism

WHI: Women's Health Initiative

Abstract

Purpose: The role of estrogen in cardiovascular health remains contested with discrepancies between findings from randomized controlled trials and observational studies. Mendelian randomization, which assesses the effect of lifelong endogenous exposure, may help elucidate these discrepancies.

Methods: We used separate sample instrumental variable analysis to estimate the association of log 17β estradiol with factors related to cardiovascular disease risk (systolic and diastolic blood pressure, lipids, fasting glucose, body mass index, waist hip ratio, and waist circumference), and Framingham score, a predictor of 10-year risk of ischemic heart disease events, in older Chinese women from the Guangzhou Biobank Cohort Study (GBCS, $n=3,092$). The estimate was derived using the Wald estimator, i.e. the ratio of the association of genetic determinants (rs1008805 and rs2175898) of log 17β -estradiol with cardiovascular disease risk factors and Framingham score in GBCS and the association of these genetic determinants with log 17β -estradiol in a sample of young women from Hong Kong ($n=236$).

Results: Genetically higher 17β -estradiol was not associated with any cardiovascular disease related risk factor, or with Framingham score (-0.01 , 95% confidence interval -1.34 to 1.31).

Conclusions: Lifetime exposure to estrogen does not appear to be cardio-protective via the cardiovascular disease related risk factors examined.

Highlights

- Mendelian randomization assesses lifelong exposure and is less prone to confounding
- Mendelian randomization showed no effect of lifelong estrogen on CVD risk factors
- Our study does not support the timing hypothesis as explanation for trial results
- Discrepancies between observational studies and trials is likely due to confounding

Introduction

The role of estrogen in cardiovascular disease and its risk factors has been intensively investigated for many years. Initially it was thought that estrogen protects against cardiovascular disease because women have lower cardiovascular disease rates than men and the menopause precedes an increase in cardiovascular deaths. However, randomized controlled trials (RCTs) suggests that hormone replacement therapy (HRT) has little overall effect on cardiovascular disease or its risk factors but decreases LDL-cholesterol slightly (1), decreases diabetes and increases the risk of thrombosis and stroke (2). Nevertheless, findings from these RCTs have been contested (3), and the possibility that estrogens have some cardiovascular benefits remains a topic of active experimental investigation (4). First, some hormone replacement trials mainly used estrogen and progestin, whose effects may differ from estrogen alone (2). For example in the Women's Health Initiative (WHI) trial, estrogen and progestin together may increase the risk of ischemic heart disease whereas estrogen only had no effect (2). Second, the effect of estrogen on cardiovascular risk may depend on other factors, such as timing of HRT initiation (4, 5) or the presence of atherosclerosis, i.e., when atherosclerosis is absent estrogen protects but when atherosclerosis is present estrogen is detrimental (6). Given the risk of breast cancer and venous thrombosis from estrogen administration, further major RCTs of estrogen are unlikely, but the role of estrogen in cardiovascular disease remains an important topic relevant to the widespread use of oral contraceptives including estrogen. In this situation replication, or otherwise, using a different approach can make a key contribution. In this study, for the first time, we used a separate sample Mendelian randomization analysis to assess the lifelong effect of estrogen on cardiovascular disease risk factors, overall and in women without cardiovascular disease.

Mendelian randomization analysis, i.e., instrumental variable analysis with genetic instruments, is increasingly used to evaluate the causal role of risk factors in disease, particularly when RCTs are unavailable. Mendelian randomization may also enable assessment of the effects of endogenous rather than exogenous exposure, which may differ, because Mendelian randomization tests a causal pathway while an RCT tests an intervention which can have unknown pleiotropic effects (7). Genetic polymorphisms associated with the exposure are randomly allocated during conception, so this resembles the randomization process in RCTs and hence is less susceptible to confounding (8). Mendelian randomization analyses have clarified the role of many factors in cardiovascular disease etiology, such as C reactive protein (9). However, a conventional Mendelian randomization analysis is sensitive to measurement error of the exposure, which may lead to inflated estimates (10). Separate sample Mendelian randomization analysis, where a genetic prediction rule is generated in a sample less susceptible to measurement error, may alleviate this problem (11). We have successfully implemented this approach to examine the effect of testosterone on cardiovascular disease risk factors and of testosterone and estrogen on inflammation in older people (12-14). In this study, we examined the relation of lifelong exposure to estrogen with cardiovascular disease risk factors using a Mendelian randomization analysis among Southern Chinese women to clarify the role of lifelong estrogen exposure in health.

Materials and methods

Sources of data

Two groups of women of different ages from the same genetic background, that is, from Hong Kong and Guangzhou, the capital of Guangdong, in Southern China, were recruited. Most Hong Kong residents are first, second or third generation migrants from Guangdong (15). First, 237 young women (mean age 21.0 years) were recruited with restriction to those with both parents and at least three grandparents born in Hong Kong or Guangdong and not taking hormone-related medication. Morning blood samples were taken on the 4th day to 7th day of the menstrual cycle for 17 β -estradiol assessment, by immunoassay (Ortho Clinical Diagnostics Vitros Eci), and DNA extraction. Self-administered questionnaires were used to collect information, such as socioeconomic position and health status. Second, we used a sample of older women (50+ years) from GBCS, an ongoing collaboration of Guangzhou Number 12 Hospital, the Universities of Hong Kong and Birmingham, UK (16). Recruitment of participants was in 3 phases. All participants were permanent residents of Guangzhou and members of “The Guangzhou Health and Happiness Association for the Respectable Elders” (GHHARE), a community social and welfare association unofficially aligned with the municipal government. Membership is open to older people for a monthly fee of 4 Yuan (50 US cents). About 7% of permanent Guangzhou residents aged 50+ years are members of GHHARE, of whom 11% (about 10,000 participants) enrolled for each of phases one, two and three. The inclusion criteria were that they were capable of consenting, ambulatory, and not receiving treatment modalities which, if omitted, may result in immediate life-threatening risk, such as chemotherapy or radiotherapy for cancer, or dialysis for renal failure. The methods of measurement have previously been reported (16). Standing height was measured without shoes to the nearest 0.1 centimetre. Weight was measured in light clothing to the nearest 0.1 kilogram. Hip circumference was measured at the greatest circumference round the buttocks below the iliac crest. Waist circumference was measured

horizontally around the smallest circumference between the ribs and iliac crest, or at the level of the naval for obese participants. Body mass index (BMI) was calculated as the weight in kilograms divided by the square of height in meters. We recorded seated blood pressure as the average of the last two or three measurements, using the Omron 705CP sphygmomanometer (Omron Corp., Kyoto, Japan). Fasting low density lipoprotein (LDL) cholesterol, HDL cholesterol, triglyceride, and glucose levels were determined with a Shimadzu CL-8000 clinical chemical analyzer (Shimadzu Corp., Kyoto, Japan) in the hospital laboratory. Fasting blood samples were collected at recruitment in phase 3 or at follow up for participants recruited in other phases. Samples were stored, as whole blood or as buffy coat and sera, at -80°C for all apart from a subset of phase 3 participants whose DNA was extracted from fresh blood and stored at -80°C .(17) The University of Hong Kong-Hospital Authority Hong Kong West Cluster Joint Institutional Review Board approved the study. The Guangzhou Medical Ethics Committee of the Chinese Medical Association approved GBCS, including the use of genetic data. All participants gave written, informed consent prior to participation.

DNA extraction and SNP analysis

DNA was extracted using QIAamp DNA Blood Midi Kit (Catalog No.51185) for fresh blood in Hong Kong, phenol-chloroform extraction for fresh blood in GBCS and magnetic bead extraction for previously stored specimens in GBCS (17). SNPs were selected from genes (*ESR1*, *ESR2* and *CYP19A1*) (18-21) functionally relevant to estradiol or breast cancer, with minor allele frequency $>5\%$ in Chinese (22). Five SNPs (rs2175898 from *ESR1*, rs1256030 and rs1256031 from *ESR2*, and rs10046 and rs1008805 from *CYP19A1*) were analyzed at the Centre for

Genomic Sciences of the University of Hong Kong, for the Hong Kong sample, and a commercial company (Beijing Capital Bio Corporation) in Beijing, for the GBCS sample, using a Mass ARRAY system (Sequenom, San Diego, California). For DNA quality analysis we used spectrophotometry for most of the samples and gel electrophoresis for four duplicate check controls and six randomly selected samples in each DNA sample plate. The determined sample concentration and A260/280 ratios were 10-20ng/uL and 1.7-2.0, respectively. A call rate <80% was considered failure. All SNPs passed with a call rate >95%.

Exposure

The exposure was genetically predicted log 17β -estradiol (pmol/L).

Outcome

The outcomes were systolic and diastolic blood pressure, LDL-cholesterol, HDL-cholesterol, triglycerides, glucose, body mass index, waist hip ratio, waist circumference, and Framingham score. The Framingham score overestimates absolute risk of cardiovascular disease in Chinese populations (23), but provides a risk ranking. The Framingham score was calculated from age, LDL-cholesterol, HDL-cholesterol, systolic blood pressure, diastolic blood pressure, diabetes (fasting plasma glucose ≥ 7.0 mmol/L, previous diagnosis or use of anti-diabetic medication) but excluded smoking to assess cardiovascular disease risk reflected by biological factors which provides more mechanistic information and is assessed more precisely.

Statistical analysis

In the sample of young Chinese women from Hong Kong, we established genetic predictors of log 17 β -estradiol based on 2 SNPs from stepwise linear regression starting with 5 SNPs with replication in 1,000 bootstrapping samples, as described previously (13). In the GBCS sample, we tested for Hardy-Weinberg equilibrium at the SNP locus on a contingency table of observed-versus-predicted frequencies with an exact test. We used ANOVA to assess whether genetically predicted log 17 β -estradiol was associated with potential confounders. We used separate sample instrumental variable analysis to estimate the association of log 17 β -estradiol with cardiovascular risk factors and Framingham score in the sample of older Chinese women from GBCS by using *suest* (seemingly unrelated regression command in Stata) to generate the Wald estimates from the ratio of the association of genetic determinants of log 17 β -estradiol with cardiovascular risk factors and Framingham score in GBCS and the association of genetic determinants of log 17 β -estradiol with log 17 β -estradiol in the sample of young women (24). Figure 1 shows the flow chart of the study.

Sensitivity analyses

Given the hypothesis that estrogen may have a different effect depending on cardiovascular disease status, we repeated the analyses by restricting the samples without self-reported history of cardiovascular diseases (ischemic heart disease, stroke/ transient ischemic attack, angina, myocardial infarction, valvular heart disease or peripheral vascular disease). We did not include

rheumatic heart disease because it has an infectious origin different from other types of cardiovascular disease (25).

All statistical analyses were conducted using Stata 13.1 (StataCorp LP, College Station, Texas, USA).

Results

Among the young Chinese woman samples (n=236, one participant was excluded because of an invalid 17 β -estradiol value). The F statistic for the regression on log 17 β -estradiol on genetic score was 13.2, with an adjusted R² of 4.9%, suggesting that the analyses were unlikely to be susceptible to weak instrument bias. Among the 22,067 women in all 3 phases of GBCS, SNP testing was available for 3,316 women, with availability depending on the phase of recruitment and other logistical concerns, but not on cardiovascular risk factors. All the SNPs had a call rate >95%. Among these 3,316 women, 3,096 (93.4%) had all selected SNPs. The 2 SNPs used to generate the genetic score did not deviate from the Hardy Weinberg equilibrium (p=0.42 for rs1008805; and p=0.73 for rs2175898) in the GBCS sample. Table 1 shows genetically predicted estrogen was unrelated to age, smoking, alcohol use, physical activity, education and longest held occupation among older Southern Chinese women from GBCS.

Table 2 shows the associations of genetically predicted log 17 β -estradiol with cardiovascular disease risk factors. Genetically predicted log 17 β -estradiol was not associated with any of the cardiovascular disease risk factors among older Southern Chinese women. Genetically predicted

17 β -estradiol was also not associated with Framingham score with estimates very close to null (-0.01, 95% confidence interval -1.34 to 1.31). The analyses restricted to those without self-reported cardiovascular disease produced directionally similar estimates (Table 3).

Discussion

To date, this is the first study using Mendelian randomization analysis to examine the role of estrogen in traditional cardiovascular disease risk factors. Using a study design less susceptible to confounding than observational studies which also allows examination of the effect of lifelong estrogen exposure, our results are consistent with RCTs of HRT where estrogen did not affect systolic and diastolic blood pressure, HDL cholesterol or body mass index (1), suggesting exogenous and endogenous estrogen have similar effects, as well as partly validating the Mendelian randomization approach. However, we did not show that estrogen reduced LDL cholesterol although the direction of the estimate was the same as in RCTs of HRT.(1) The null findings for glucose are inconsistent with the improvement in glycemic metabolism and lower waist circumference seen in RCTs of HRT (26, 27). Lastly, we did not find any association of estrogen with predicted risk of future ischemic heart disease events, proxied by Framingham score, consistent with the WHI findings in the estrogen alone treatment arm (28).

Although we used a separate sample Mendelian randomization analysis which is less susceptible to confounding and allows estimation of the effect of lifelong estrogen exposure, limitations existed. First, cardiovascular events are not available, so estrogen could protect against cardiovascular events by pathways other than those examined here. Our previous study showed

estrogen reduced the inflammation implicated in cardiovascular disease (13). Estrogen might also reduce testosterone (29), which could be beneficial. However, estrogen also promotes thrombosis that can provoke cardiovascular events (30). Second, although a separate sample Mendelian randomization analysis may be less susceptible to measurement of the exposure compared to a conventional Mendelian randomization analysis, we are unable to assess if the instrument violates the exclusion restriction criteria, i.e., that the genetic polymorphisms only affected cardiovascular disease risk via estrogen, as we do not have estrogen polymorphisms, and estrogen and cardiovascular risk factors were all measured in the same dataset. However, the estrogen related genetic polymorphisms are in the sex-steroid pathway, and genetically predicted estrogen was unrelated to testosterone in the sample of young women (13). Third, we did not use estrogen polymorphisms from genome wide association studies (GWAS), because no such studies have reliably identified estrogen related genetic polymorphisms in Chinese women. Therefore, we have established our own genetic prediction rule in a young Chinese women sample using stepwise regressions and bootstrapping methods, as described previously.(13) Fourth, Mendelian randomization analysis may be susceptible to confounding by population stratification but the participants from this study had the same genetic origin. Fifth, although our study is largely in line with the results from RCTs of HRT but the confidence intervals are wide, so we cannot rule out small effects on glucose and LDL cholesterol as seen in RCTs (31). Our study was powered to detect a change of 0.23 standard deviation in cardiovascular disease risk factors, for example a blood pressure change of 5 mmHg per increase in one log unit change in estrogen. As such our study was underpowered to detect small effects of estrogen that could be important to population health, although less clinically relevant. Sixth, we used genetic determinants of estrogen in young women to avoid reverse causality, however effects of genetic

determinants could vary with age. Lastly, our study only focused on estrogen and hence our result could not directly infer the relation of combined treatment (i.e. estrogen and a progestin) on cardiovascular health. Furthermore, this makes comparison of our results with previous RCTs of HRT more difficult because not all trials had an estrogen only arm. However, this may also imply our study adds to our understanding of the effect of estrogen alone on cardiovascular risk factors.

Our study provides an example of a feasible method of implementing Mendelian randomization studies cost-effectively when GWAS of uncommonly measured exposures are not available. It also confirms a fairly minor role of estrogen in cardiovascular disease risk factors, consistent with the minor inflection in cardiovascular disease rates at the menopause (32). Our study does not suggest that the effect of estrogen on cardiovascular disease varies by disease status. As such, observations concerning differences by timing of HRT initiation (5) might be a false positive to which post-hoc analyses are susceptible, and would need confirmation from studies specifically designed to examine the timing hypothesis (4). Our findings are not consistent with observational studies of HRT, but these are known to be biased by HRT users being of higher socio-economic status and more health conscious than non-users (33-35). Our results might also be relevant to the much more common estrogen exposure from the oral contraceptive pill, whose effects on cardiovascular events have never been assessed in an RCT (36). Observational studies have found combined contraceptive pills (i.e. estrogen and progestin) associated with higher blood pressure, LDL cholesterol, higher risk of thrombosis, and lower HDL cholesterol, which could be due to the estrogen or progestin or residual confounding (36). However, we cannot assess this possibility because we do not have these events, or their biomarkers, such as thrombin.

Our study provides new evidence concerning the relation of estrogen with cardiovascular disease risk factors using a separate sample Mendelian randomization analysis. No evidence supporting potential cardio-protection by estrogen or evidence for the timing hypothesis was found.

However, given the limitations of this study, a larger Mendelian randomization analysis with cardiovascular outcomes may help clarify the role of estrogen in cardiovascular disease in a timely manner, with corresponding implications for public health, clinical practice and etiology concerning estrogen exposures such as HRT and oral contraceptives.

Conflict of Interest and Source of Funding:

The Guangzhou Biobank Cohort Study was funded by the University of Hong Kong Foundation for Development and Research (Hong Kong, China); the University of Hong Kong University Research Committee Strategic Research Theme of Public Health (Hong Kong, China); Guangzhou Public Health Bureau (Guangzhou, China; Key technology collaboration project, grant number 2012J5100041), Guangzhou Science and Technology Bureau (Guangzhou, China), Bureau of Guangzhou Science and Technology (Grant 2012J5100041; 2013J4100031), and the University of Birmingham (Birmingham, United Kingdom). This sub-study was funded by the Research Grant Council General Research Fund (grant number 769710), Research Grant Council of Hong Kong, Hong Kong SAR, People's Republic of China. In addition to Kar Keung Cheng's appointment at University of Birmingham, he is affiliated to Department of General Practice at Peking University Health Science Centre. The latter receives support from Pfizer China to support the training of family doctors (approximately US\$100,000 a year for 2014-16). The authors have no other conflict of interest, financial or otherwise.

The funders had no role in the study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

Acknowledgements

The Guangzhou Biobank Cohort Study investigators include: Guangzhou No. 12 Hospital--Dr. Zhang WS, Prof. Jiang CQ (Co-Principal Investigator (PI)); University of Hong Kong--Dr. C. M. Schooling, , Prof. R Fielding, Prof. GM Leung, Prof. TH Lam (Co-PI); University of Birmingham--Dr. G. N. Thomas, Prof. P Adab, Prof. KK Cheng (Co-PI). The authors thank Dr Sushma Kavikondala for convoking the student sample and Drs. SS Cherny, MM Garcia-Barcelo, SO Man, Xu L, Zhu T for facilitating the DNA extraction.

References

1. Casanova G, Bossardi Ramos R, Ziegelmann P, Spritzer PM. Effects of low-dose versus placebo or conventional-dose postmenopausal hormone therapy on variables related to cardiovascular risk: a systematic review and meta-analyses of randomized clinical trials. *J Clin Endocrinol Metab.* 2015;100(3):1028-37.
2. Manson JE, Chlebowski RT, Stefanick ML, Aragaki AK, Rossouw JE, Prentice RL, et al. Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. *JAMA.* 2013;310(13):1353-68.
3. Prentice RL, Langer R, Stefanick ML, Howard BV, Pettinger M, Anderson G, et al. Combined postmenopausal hormone therapy and cardiovascular disease: toward resolving the discrepancy between observational studies and the Women's Health Initiative clinical trial. *Am J Epidemiol.* 2005;162(5):404-14.
4. Hodis HN, Mack WJ, Shoupe D, Azen SP, Stanczyk FZ, Hwang-Levine J, et al. Methods and baseline cardiovascular data from the Early versus Late Intervention Trial with Estradiol testing the menopausal hormone timing hypothesis. *Menopause (New York, NY).* 2015;22(4):391-401.
5. Manson JE, Bassuk SS. Invited commentary: hormone therapy and risk of coronary heart disease why renew the focus on the early years of menopause? *Am J Epidemiol.* 2007;166(5):511-7.
6. Michels KB. Hormone replacement therapy in epidemiologic studies and randomized clinical trials - are we checkmate? *Epidemiology.* 2003;14(1):3-5.
7. Schooling CM, Freeman G, Cowling BJ. Mendelian randomization and estimation of treatment efficacy for chronic diseases. *Am J Epidemiol.* 2013;177(10):1128-33.
8. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey-Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med.* 2008;27(8):1133-63.
9. Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, Shah T, et al. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ.* 2011;342.
10. Schooling CM, Au Yeung SL, Freeman G. Mendelian Randomization Estimates May Be Inflated. *J Am Coll Cardiol.* 2013;61(18):1931-.
11. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol.* 2013;178(7):1177-84.
12. Zhao J, Jiang C, Lam TH, Liu B, Cheng KK, Xu L, et al. Genetically predicted testosterone and cardiovascular risk factors in men: a Mendelian randomization analysis in the Guangzhou Biobank Cohort Study. *Int J Epidemiol.* 2014;43(1):140-8.
13. Zhao J, Jiang CQ, Lam TH, Liu B, Cheng KK, Kavikondala S, et al. Genetically predicted 17beta-estradiol and systemic inflammation in women: a separate-sample Mendelian randomisation analysis in the Guangzhou Biobank Cohort Study. *J Epidemiol Community Health.* 2014;68(8):780-5.
14. Zhao J, Jiang C, Lam TH, Liu B, Cheng KK, Xu L, et al. Genetically predicted testosterone and systemic inflammation in men: a separate-sample Mendelian randomization analysis in older Chinese men. *PLoS One.* 2015;10(5):e0126442.
15. Tsang SY-S. A modern history of Hong Kong. London ; New York: I.B. Tauris; 2004. xi, 340 p. p.
16. Jiang C, Thomas GN, Lam TH, Schooling CM, Zhang W, Lao X, et al. Cohort profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration. *Int J Epidemiol.* 2006;35(4):844-52.
17. Au Yeung SL, Jiang CQ, Cheng KK, Liu B, Zhang WS, Lam TH, et al. Evaluation of Moderate Alcohol Use and Cognitive Function Among Men Using a Mendelian Randomization Design in the Guangzhou Biobank Cohort Study. *Am J Epidemiol.* 2012;175(10):1021-8.

18. Breast, Prostate Cancer Cohort C, Cox DG, Bretsky P, Kraft P, Pharoah P, et al. Haplotypes of the estrogen receptor beta gene and breast cancer risk. *Int J Cancer*. 2008;122(2):387-92.
19. Lundin E, Wirgin I, Lukanova A, Afanasyeva Y, Krogh V, Axelsson T, et al. Selected polymorphisms in sex hormone-related genes, circulating sex hormones and risk of endometrial cancer. *Cancer Epidemiol*. 2012;36(5):445-52.
20. Udler MS, Azzato EM, Healey CS, Ahmed S, Pooley KA, Greenberg D, et al. Common germline polymorphisms in COMT, CYP19A1, ESR1, PGR, SULT1E1 and STS and survival after a diagnosis of breast cancer. *Int J Cancer*. 2009;125(11):2687-96.
21. Talbott KE, Gammon MD, Kibriya MG, Chen Y, Teitelbaum SL, Long CM, et al. A CYP19 (aromatase) polymorphism is associated with increased premenopausal breast cancer risk. *Breast Cancer Res Treat*. 2008;111(3):481-7.
22. .National Centre for Biotechnology Information UNLoMNSNPd. Available from: <http://www.ncbi.nlm.nih.gov/snp>.
23. Liu J, Hong Y, D'Agostino RB, Sr., Wu Z, Wang W, Sun J, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. *JAMA*. 2004;291(21):2591-9.
24. Wooldridge JM. *Introductory econometrics : a modern approach*. 4th ed. Mason, OH: South Western, Cengage Learning; 2009. xx, 865 p. p.
25. Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic heart disease. *Lancet*. 2012;379(9819):953-64.
26. Salpeter SR, Walsh JM, Ormiston TM, Greyber E, Buckley NS, Salpeter EE. Meta-analysis: effect of hormone-replacement therapy on components of the metabolic syndrome in postmenopausal women. *Diabetes Obes Metab*. 2006;8(5):538-54.
27. Bonds DE, Lasser N, Qi L, Brzyski R, Caan B, Heiss G, et al. The effect of conjugated equine oestrogen on diabetes incidence: the Women's Health Initiative randomised trial. *Diabetologia*. 2006;49(3):459-68.
28. Prentice RL, Langer RD, Stefanick ML, Howard BV, Pettinger M, Anderson GL, et al. Combined analysis of Women's Health Initiative observational and clinical trial data on postmenopausal hormone treatment and cardiovascular disease. *Am J Epidemiol*. 2006;163(7):589-99.
29. Zimmerman Y, Eijkemans MJ, Coelingh Bennink HJ, Blankenstein MA, Fauser BC. The effect of combined oral contraception on testosterone levels in healthy women: a systematic review and meta-analysis. *Hum Reprod Update*. 2014;20(1):76-105.
30. Curb JD, Prentice RL, Bray PF, Langer RD, Van Horn L, Barnabei VM, et al. Venous thrombosis and conjugated equine estrogen in women without a uterus. *Arch Intern Med*. 2006;166(7):772-80.
31. Freeman G, Cowling BJ, Schooling CM. Power and sample size calculations for Mendelian randomization studies using one genetic instrument. *Int J Epidemiol*. 2013;42(4):1157-63.
32. Tunstall-Pedoe H. Myth and paradox of coronary risk and the menopause. *Lancet*. 1998;351(9113):1425-7.
33. Petitti D. Commentary: hormone replacement therapy and coronary heart disease: four lessons. *Int J Epidemiol*. 2004;33(3):461-3.
34. Lawlor DA, Davey-Smith G, Ebrahim S. Socioeconomic position and hormone replacement therapy use: Explaining the discrepancy in evidence from observational and randomized controlled trials. *Am J Public Health*. 2004;94(12):2149-54.
35. Petitti DB, Freedman DA. Invited commentary: how far can epidemiologists get with statistical adjustment? *Am J Epidemiol*. 2005;162(5):415-8; discussion 9-20.
36. Shufelt CL, Bairey Merz CN. Contraceptive hormone use and cardiovascular disease. *J Am Coll Cardiol*. 2009;53(3):221-31.

Table 1: Genetically predicted log 17 β -estradiol by lifestyle and socio-demographics among Southern Chinese women in the Guangzhou Biobank Cohort Study

Characteristic		Sample size	Predicted log 17 β -estradiol (pmol/L)	^a p value
Age group	50-54	379	4.73 (0.10)	0.24
	55-59	1061	4.74 (0.10)	
	60-64	696	4.73 (0.10)	
	65-69	538	4.73 (0.10)	
	70-74	289	4.73 (0.09)	
	75-79	94	4.73 (0.10)	
	80+	38	4.74 (0.09)	
Smoking status	Never smokers	2,992	4.7 (0.10)	0.44
	Former smokers	52	4.8 (0.08)	
	Current smokers	45	4.7 (0.10)	
Alcohol status	Never drinkers	2,157	4.7 (0.10)	0.70
	Former drinkers	101	4.7 (0.10)	
	Current drinkers	808	4.7 (0.10)	
^b Physical activity (IPAQ)	Inactive	194	4.7 (0.10)	0.75
	Minimally active	998	4.7 (0.10)	
	HEPA active	1,903	4.7 (0.10)	
Education	Less than primary	426	4.7 (0.10)	0.26
	Primary	1,004	4.7 (0.10)	
	Junior middle	746	4.7 (0.10)	
	Senior middle	688	4.7 (0.10)	
	Junior college	154	4.7 (0.09)	
	College	77	4.7 (0.10)	
^c Longest-held occupation	Manual	2,099	4.7 (0.10)	0.72
	Non-manual	638	4.7 (0.10)	
	Others	321	4.7 (0.10)	

^aP value obtained from ANOVA.

^bHEPA: Health-enhancing physical activity (i.e. vigorous activity at least 3 days a week achieving at least 1500 metabolic equivalent (MET) minutes per week or activity on 7 days of the week, achieving at least 3000 MET minutes per week (IPAQ: International Physical Activity Questionnaire).

^cManual occupations are agricultural worker, factory work or sales and service; non-manual are administrator/ manager, professional/technical, military/disciplined.

Table 2: Effect of genetically predicted log 17β -estradiol on cardiovascular disease risk and its risk factors among Southern Chinese women in the Guangzhou Biobank Cohort Study using a Mendelian randomization analysis

Cardiovascular risk factors	Sample size	Estimate	95% Confidence Interval
Systolic blood pressure (mmHg)	3,092	-3.83	-11.8 to 4.78
Diastolic blood pressure (mmHg)	3,092	-1.18	-5.11 to 2.76
HDL cholesterol (mmol/L)	3,079	-0.02	-0.17 to 0.12
LDL cholesterol (mmol/L)	3,074	0.16	-0.12 to 0.44
Triglyceride (mmol/L)	3,079	-0.08	-0.44 to 0.28
Fasting glucose (mmol/L)	3,072	0.08	-0.47 to 0.63
Body mass index (kg/m ²)	3,090	-0.13	-1.35 to 1.08
Waist hip ratio	3,085	-0.003	-0.03 to 0.02
Waist circumference (cm)	3,087	-0.59	-3.85 to 2.67
Framingham score	3,071	-0.01	-1.34 to 1.31

Table 3: Effect of genetically predicted log 17β -estradiol on cardiovascular disease risk and its risk factors among Southern Chinese women in the Guangzhou Biobank Cohort Study restricted to those without self reported cardiovascular disease status, using a Mendelian randomization analysis

	Absence of self reported cardiovascular diseases		
Cardiovascular risk factors	Sample size	Estimate	95% Confidence Interval
Systolic blood pressure (mmHg)	2,916	-1.84	-9.84 to 6.16
Diastolic blood pressure (mmHg)	2,916	-0.44	-4.44 to 3.56
HDL cholesterol (mmol/L)	2,904	-0.02	-0.17 to 0.12
LDL cholesterol (mmol/L)	2,899	0.17	-0.12 to 0.46
Triglyceride (mmol/L)	2,904	-0.10	-0.47 to 0.27
Fasting glucose (mmol/L)	2,900	-0.02	-0.58 to 0.55
Body mass index (kg/m ²)	2,915	-0.02	-1.25 to 1.22
Waist hip ratio	2,910	-0.003	-0.03 to 0.02
Waist circumference (cm)	2,911	-0.48	-3.82 to 2.86
Framingham score	2,896	0.03	-1.33 to 1.39

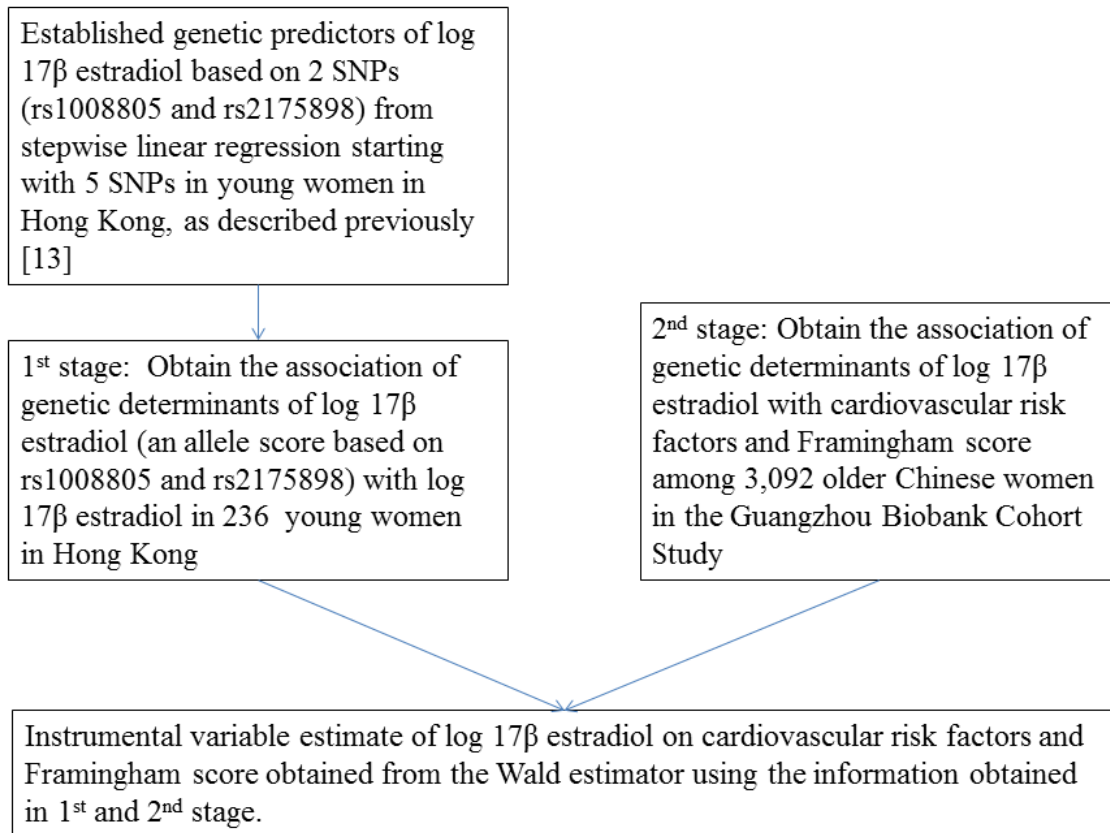


Figure 1: Flow chart of the study. A separate-sample instrumental variable analysis with genetic instruments, i.e. Mendelian randomization analysis, was used. See detailed in the Methods