**ARIC Manuscript Proposal #4127**

| | | |
|---|---|---|
| **PC Reviewed: 9/13/22** | **Status: _____** | **Priority: 2** |
| **SC Reviewed: _____** | **Status: _____** | **Priority: ____** |

**1.a. Full Title**: Does alcohol consumption modify the genetic risk of breast cancer in postmenopausal women: The Atherosclerosis Risk in Communities (ARIC) Study

**b. Abbreviated Title (Length 26 characters)**: Alcohol, PRS/SNP, and cancer

**2. Writing Group**:

Writing group members: Minghui Zhang, Ken Butler, Nilanjan Chatterjee, David Couper, Corinne E. Joshu (invited), Anna Prizment, Elizabeth A. Platz

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. ___MZ___ **[please confirm with your initials electronically or in writing]**


**First author**: **Minghui Zhang**
Address: Department of Epidemiology,
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street Baltimore, MD 21205

Phone: 443-570-6826 Fax: N.A.
E-mail: mzhan133@jhu.edu

**ARIC author** to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: **Dr. Elizabeth Platz**
Address: 615 N. Wolfe St., Rm E6132, Baltimore, MD 21205

Phone: 410-614-9674 Fax:
E-mail: Eplatz1@jhu.edu


**3. Timeline**: 1-2 years

**4. Rationale**:

Breast cancer is one of the most common cancers and the second leading cause of cancer death among women in the United States [1], and it's a complex disease resulting from a combination of genetic, environmental, and reproductive factors [2-4].

Studies have shown that women with a positive family history have a higher lifetime risk of breast cancer, in part, indicating an inherited risk of breast cancer [5-6]. The heritability of breast cancer is estimated to be about thirty percent from twins' studies [7-8]. High penetrance genes (e.g., *BRCA1, BRCA2, PALB2, ATM*, and *CHEK2*) contribute to a relatively high lifetime risk of breast

cancer [9]. However, these variants are rare and account for only a small proportion of breast cancer [10]. A recent approach to discovering low penetrance genes of breast cancer risk is to use genome-wide association studies (GWAS). Although each of the validated single nucleotide polymorphism (SNP) individually contributes only to a small risk, they are common in the general population and thus their overall contribution to breast cancer is estimated to be greater than known high penetrate genes [11-12]. In 2015, the first polygenic risk score (PRS) assessing the joint contribution of 77 common SNPs on breast cancer risk was developed [13]. In the study, women in the highest 1% of the PRS had a three times higher risk of breast cancer compared with women in the middle quintile. In 2019, an updated PRS of 313 SNPs was developed [11]. The results showed that women in the highest 1% of the PRS have a 4.04 times higher risk of breast cancer, compared to those in the middle quintile. This study suggested that the PRS is a discriminatory and calibrated method for predicting breast cancer risk.

Besides genetic factors and reproductive factors, various well-established environmental factors are associated with breast cancer risk [2-3], including alcohol consumption. Alcohol is a known human carcinogen that is associated with an elevated risk of several types of cancer [14]. Systematic reviews have found a positive association between alcohol drinking and breast cancer risk [15-16]. For every 10 g of ethanol consumed daily, the relative risk of breast cancer increases by approximately 10%. Potential mechanisms include the differences in people's ability to metabolize acetaldehyde and the impact of alcohol on estrogen levels, enzymatic activities, and oxidative stress [15,17-19].

However, the interaction between alcohol consumption and genetic factors on breast cancer development is poorly understood. To our knowledge, only four studies have attempted to assess the interaction between PRS and alcohol on breast cancer risk. A population-based, nested case-control study in Korea found a quantitative interaction between alcohol intake and a 4-SNP estrogen-related PRS on breast cancer risk [20]. The association between high PRS and breast cancer risk was more robust in women with moderate alcohol consumption than those with no or mild alcohol consumption. A nested case-control study among White women included in 8 prospective cohorts from the Breast and Prostate Cancer Cohort Consortium in the U.S. found no interaction between alcohol consumption and a 24-SNP PRS on breast cancer risk [21]. Another analysis that included participants of European ancestry in 20 case-control studies from the Breast Cancer Association Consortium (BCAC), both prospective and retrospective, found a significant sub-multiplicative interaction between a 77-SNP PRS and alcohol consumption on breast cancer risk (ORinteraction= 0.90 (0.82–0.98), P-value: 0.016) [22]. However, a further study using data from women of European ancestry from 46 studies included in BCAC showed no statistically significant interaction between a 313-SNP PRS and alcohol intake [23]. Additional support for the hypothesis that lifestyle factors may interact with PRS on breast cancer risk comes from a British study of women in the UK Biobank [2]. The study found an additive but not multiplicative interaction between overall environmental factors, as measured by the Healthy Lifestyle Index, and the genetic factors, as measured by a 304-SNP PRS, on breast cancer risk. However, another similar study examining overall lifestyle factors and a 305-SNP PRS in women in the UK Biobank found no significant interaction between the lifestyle index and genetic risk groups for breast cancer risk, though with limited study power [24]. The overall environmental factors in these latter two studies included alcohol, but they did not separately evaluate the alcohol-PRS association.

Given the inconsistent findings described above and the importance of determining whether the genetic risk of breast cancer can be modified, more research is needed to evaluate this interaction.

Therefore, in our study in ARIC, we will evaluate whether alcohol consumption modifies the association between the genetic risk of breast cancer as measured by the most up to date PRS (313 SNPs). Since alcohol affects the risk of cancer in multiple ways through a variety of pathways, we will assess genetic risk across the whole genome using the 313-SNP PRS, rather than restricting to SNPs in alcohol metabolism genes. We hypothesize that there's a multiplicative interaction between alcohol consumption and PRS for breast cancer risk. If our hypothesis is supported, this gene-environment interaction could better inform the implementation of lifestyle interventions to reduce breast cancer risk in high genetic risk subgroups.

## 5. Main Hypothesis/Study Questions:

Our study question is whether alcohol consumption modifies the genetic risk of breast cancer. We will first confirm that each of the characteristics, the 313-SNP PRS and alcohol consumption, are associated with breast cancer in ARIC, and then analyze our study question.

We hypothesize that there is a multiplicative interaction between alcohol consumption and 313-SNP PRS for breast cancer risk. That is, women with a higher PRS and high alcohol consumption will have a higher breast cancer risk than the expected value of the risk for women who have a high alcohol consumption but low genetic risk multiplied by the risk for women who have a high genetic risk but low alcohol consumption.

The previous four studies on this hypothesis have conflicting results. The Korean study supports our hypothesis of multiplicative interaction between alcohol drinking and PRS on breast cancer risk, but the other three studies of participants of European ancestry do not. One of these three studies found a sub-multiplicative interaction, and the remaining two did not find a statistically significant interaction. This inconsistency may be due to differences in the PRS they used and the differences in the study populations. The Korean study used an estrogen-related polygenic risk score, whereas the other studies used polygenic risk scores assessing genome-wide genetic risk, and only one study used the most up to date PRS, (4/24/77 vs 313). Besides, the ability to metabolize alcohol differs between populations of different ancestry [24], and these PRSs do not include those alcohol metabolic genes. Therefore, further studies with diverse populations, long-term prospective follow-up, the most updated PRS, along with detailed information on potential established and suspected confounders, are needed.

However, whether our finding suggests a multiplicative (RR11>RR01*RR10) or sub-multiplicative (RR11<RR01*RR10) interaction between PRS and alcohol consumption on breast cancer risk, both imply public health action. Since women with high PRS and high alcohol consumption have a higher breast cancer risk than women with only one of the two, this would convey an important message that women with high PRS should avoid excessive alcohol consumption.

## 6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

6.1 Study design:

Prospective cohort study

6.2 Inclusion/exclusion:

Women enrolled in ARIC from 1987-2015, at risk for 1st primary breast cancer, with genetic information and genetic consent, with consent to use their data other than cardiovascular outcomes

6.3 Variables:

Outcome variable: first postmenopausal primary incident breast cancer

Main variables: alcohol intake (g) per day (Visit 2), 313-SNP polygenic risk score

Covariates (For all the covariates we will use data from Visit 2 for all participants):

age at recruitment, race, education, age at menarche, age at menopause, parity, age at first live birth/pregnancy, hormone replacement therapy use, oral contraceptives use, history of mammograms, family history of cancer, BMI, height, physical activity (meet the guidelines or not), smoking in pack years, energy intake (Calorie), total folic acid intake

6.4 Statistical data analysis plan

6.4.1 Evaluate the baseline characteristic of participants (Visit 2).

Table 1a. Baseline characteristics by PRS (median as a cut point)

1b. Baseline characteristics by alcohol consumption (never vs. ever)

1c. Joint table of PRS and alcohol consumption under null hypothesis

6.4.2 Pre-analysis assumption check (Cox proportional hazards regression, adjusting for the same covariates as the main analysis)

1) Check whether the quintiles of the 313-SNP PRS are associated with breast cancer risk in the study population. If so, we will check whether the association between PRS and breast cancer is linear on the natural logarithm scale using the Likelihood Ratio test comparing continuous model and spline model.
2) Check whether alcohol consumption (grams of ethanol) is associated with breast cancer risk in the study population. If so, we will check whether the association between alcohol consumption and breast cancer is linear on the natural logarithm scale using the Likelihood Ratio test comparing continuous model and spline model.

6.4.3 Main analysis

1) Prevalent cancers at Visit 1 and breast cancer cases up to Visit 2 will be excluded. We will use Cox proportional hazards regression to estimate hazard ratios and 95% confidence intervals, adjusting for the covariates under 6.3, which are known or purported confounders for the association between alcohol drinking and breast cancer. Follow-up time will be estimated from Visit 2 until December 31st, 2015. To test for multiplicative interaction, we will include an interaction term (e.g., alcohol*PRS both continuous; or current alcohol drinker (vs not) *PRS (divided at the median) in the regression models and test its coefficient using the Likelihood Ratio test. Alcohol intake will be pre-categorized as No", "Low to moderate" (>0 to 1/day), and "High" (>1/day), based on conventional guidelines that recommend women limit their intake to 1 drink or less per day [29]. For breast cancer, no threshold has been identified (https://www.wcrf.org/wp-content/uploads/2021/02/Alcoholic-Drinks.pdf), thus we will use “No” as the reference group. We also will divide ethanol intake in grams/day (0 and tertiles) based on the distribution in the analytic cohort.

Table 2 Main result - Stratified table

| | Alcohol intake | | |
|---|---|---|---|
| | No | Low to moderate | High |
| Low-PRS | 1 (ref) | | |
| Medium-PRS | 1 (ref) | | |
| High-PRS | 1 (ref) | | |

Table 3 Main result - Joint table

| | Alcohol intake | | |
|---|---|---|---|
| | No | Low to moderate | High |
| Low-PRS | 1 (ref) | | |
| Medium-PRS | | | |
| High-PRS | | | |

2) Check proportional hazards assumption by Schoenfeld residuals.
3) If there is no multiplicative interaction, we will look at the additive scale by estimating relative excess risk due to interaction (RERI)

6.4.4 Further stratify analysis for other EMMs

1) Stratify by race
2) Stratify by history of mammograms
3) Stratify by family history of cancer

6.5 Minimal detectable size of the interaction

1) Total population: In ARIC Visit 2, there were 13162 participants, of which approximately 55% were female. Approximately 82.15% of them had genetic information, yielding a total study population estimate of 5883.
2) Total cases: There were 568 cases of primary breast cancer during the follow-up period, and assuming that the proportion of them with genetic information was the same as the total cohort, this means we will have roughly 462 cases.
3) Probability at baseline: We estimated that at 30 years of follow-up, the baseline incidence of participants aged 60 years or older was about 7.5% [26].
4) Alcohol drinking prevalence: Classifying former drinkers into non-drinker category, the prevalence of alcohol drinking among the ARIC female population was 52.96% [27].
5) Polygenic risk score was dichotomized by median.
6) The minimal detectable size of the interaction is estimated to be 1.88, using a two-sided test with an alpha level of 0.05 and 0.8 power [28]. While the size of this interaction is not small, if such an interaction of this size was observed, this would imply that alcohol is a meaningful target for breast cancer intervention in women with high PRS.

6.6 Limitations：

1) The sample size of our study is not large enough to detect the interaction between multiple categories of PRS and alcohol. It is unclear whether the interaction with the PRS would differ by former vs current drinking status.
2) The PRS was developed in European ancestry women and may be less relevant to African ancestry women.
3) We do not have complete information on all cases of breast cancer receptors (ER, PR, HER2/NEU.
4) There is insufficient power to detect effect measure modifiers of the interaction term between alcohol consumption and PRS.
5) It is possible that only excessive alcohol consumption, i.e., binge drinking, has a more pronounced effect on breast cancer risk, but we will not be able to assess this due to limited study power.
6) We will only analyze the total amount of alcohol consumption, without separating the effects of different types of alcohol because the hypothesis is about ethanol.
7) Our study population is restricted to postmenopausal women. Therefore, our findings may not be generalizable to premenopausal women.
8) Our study population consists mainly of women of European ancestry and women of African ancestry. Therefore, the study results may not be generalizable to women of other ancestries.

**7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __x__ No**

**b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No**
(The file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

**8.a. Will the DNA data be used in this manuscript? __x__ Yes ____ No**

**8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? __x__ Yes ____ No**

**9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.**
ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aricproposals/dtSearch.html

___x___ Yes _____ No

**10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?**

10.1 Enhancing the Infrastructure of the Atherosclerosis Risk in Communities (ARIC) Study for Cancer Epidemiology Research: ARIC Cancer #2975
10.2 All GWAS studies in ARIC cohort

**11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __x__ Yes ____ No**

**11.b. If yes, is the proposal**

**_x_ A. primarily the result of an ancillary study (list number*** 2011.07 Enhancing ARIC Infrastructure to Yield a New Cancer Epidemiology Cohort; 1995.04 Cancer Study**)**

**___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)**

*ancillary studies are listed by number https://sites.cscc.unc.edu/aric/approved-ancillary-studies

**12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.**

**12b. The NIH instituted a Public Access Policy in April, 2008** which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

**Reference:**


[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022 Jan;72(1):7-33. doi: 10.3322/caac.21708. Epub 2022 Jan 12. PMID: 35020204.

[2] Arthur RS, Wang T, Xue X, Kamensky V, Rohan TE. Genetic Factors, Adherence to Healthy Lifestyle Behavior, and Risk of Invasive Breast Cancer Among Women in the UK Biobank [published correction appears in J Natl Cancer Inst. 2020 Oct 1;112(10):1076]. J Natl Cancer Inst. 2020;112(9):893-901. doi:10.1093/jnci/djz241

[3] Rudolph A, Chang-Claude J, Schmidt MK. Gene-environment interaction and risk of breast cancer. Br J Cancer. 2016;114(2):125-133. doi:10.1038/bjc.2015.439

[4] Anderson KN, Schwab RB, Martinez ME. Reproductive risk factors and breast cancer subtypes: a review of the literature. Breast Cancer Res Treat. 2014;144(1):1-10. doi:10.1007/s10549-014-2852-7

[5] Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer. 1997;71(5):800-809. doi:10.1002/(sici)1097-0215(19970529)71:5<800::aid-ijc18>3.0.co;2-b

[6] Colditz GA, Kaphingst KA, Hankinson SE, et al.: Family history and risk of breast cancer: nurses' health study. Breast Cancer Res Treat 133 (3): 1097-104, 2012

[7] Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000;343(2):78-85. doi:10.1056/NEJM200007133430201

[8] Mucci LA, Hjelmborg JB, Harris JR, et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries [published correction appears in JAMA. 2016 Feb 23;315(8):822]. JAMA. 2016;315(1):68-76. doi:10.1001/jama.2015.17703

[9] Michailidou K, Beesley J, Lindstrom S, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat Genet. 2015;47(4):373-380. doi:10.1038/ng.3242

[10] Song M, Lee KM, Kang D. Breast cancer prevention based on gene-environment interaction. Mol Carcinog. 2011;50(4):280-290. doi:10.1002/mc.20639

[11] Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. Am J Hum Genet. 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002

[12] Pashayan N., Duffy S.W., Chowdhury S., Dent T., Burton H., Neal D.E., Easton D.F., Eeles R., Pharoah P. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. Br. J. Cancer. 2011;104:1656–1663.

[13] Mavaddat N, Pharoah PD, Michailidou K, et al.: Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 107 (5): , 2015. [PUBMED Abstract]

[14] IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Alcohol consumption and ethyl carbamateExit Disclaimer. IARC Monographs on the Evaluation of Carcinogenic Risks in Humans 2010; 96:3–1383.

[15] Shield KD, Soerjomataram I, Rehm J. Alcohol Use and Breast Cancer: A Critical Review. Alcohol Clin Exp Res. 2016;40(6):1166-1181. doi:10.1111/acer.13071

[16] Key J, Hodgson S, Omar RZ, et al. Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues. Cancer Causes Control. 2006;17(6):759-770. doi:10.1007/s10552-006-0011-0

[17] Varela-Rey M, Woodhoo A, Martinez-Chantar ML, Mato JM, Lu SC. Alcohol, DNA methylation, and cancer. Alcohol Res. 2013;35(1):25-35.

[18] Seitz HK, Becker P. Alcohol metabolism and cancer risk. Alcohol Res Health. 2007;30(1):38-47.

[19] Rumgay H, Murphy N, Ferrari P, Soerjomataram I. Alcohol and Cancer: Epidemiology and Biological Mechanisms. Nutrients. 2021;13(9):3173. Published 2021 Sep 11. doi:10.3390/nu13093173

[20] Song SS, Kang S, Park S. Association of Estrogen-Related Polygenetic Risk Scores with Breast Cancer and Interactions with Alcohol Intake, Early Menarche, and Nulligravida. Asian Pac J Cancer Prev. 2022;23(1):13-24. Published 2022 Jan 1. doi:10.31557/APJCP.2022.23.1.13

[21] Maas P, Barrdahl M, Joshi AD, et al. Breast Cancer Risk From Modifiable and Nonmodifiable Risk Factors Among White Women in the United States [published correction appears in JAMA Oncol. 2016 Oct 1;2(10 ):1374]. JAMA Oncol. 2016;2(10):1295-1302. doi:10.1001/jamaoncol.2016.1025

[22] Rudolph A, Song M, Brook MN, et al. Joint associations of a polygenic risk score and environmental risk factors for breast cancer in the Breast Cancer Association Consortium. Int J Epidemiol. 2018;47(2):526-536. doi:10.1093/ije/dyx242

[23] Kapoor PM, Mavaddat N, Choudhury PP, et al. Combined Associations of a Polygenic Risk Score and Classical Risk Factors With Breast Cancer Risk. J Natl Cancer Inst. 2021;113(3):329-337. doi:10.1093/jnci/djaa056

[24] Al Ajmi K, Lophatananon A, Mekli K, Ollier W, Muir KR. Association of Nongenetic Factors With Breast Cancer Risk in Genetically Predisposed Groups of Women in the UK Biobank Cohort. JAMA Netw Open. 2020;3(4):e203760. Published 2020 Apr 1. doi:10.1001/jamanetworkopen.2020.3760

[25] Wall TL, Luczak SE, Hiller-Sturmhöfel S. Biology, Genetics, and Environment: Underlying Factors Influencing Alcohol Metabolism. Alcohol Res. 2016;38(1):59-68.

[26] Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2017, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2017/, based on November 2019 SEER data submission, posted to the SEER web site, April 2020.

[27] He X, Rebholz CM, Daya N, Lazo M, Selvin E. Alcohol consumption and incident diabetes: The Atherosclerosis Risk in Communities (ARIC) study. Diabetologia. 2019;62(5):770-778. doi:10.1007/s00125-019-4833-1

[28] Garcia-Closas M, Lubin JH. Power and sample size calculations in case-control studies of gene-environmental interactions: Comments on different approaches. Am J Epidemiol 1999;149:689-93.

[29] US Department of Agriculture and US Department of Health and Human Services . Dietary Guidelines for Americans, 2020–2025. 9th ed. Washington, DC: US Government Publishing Office; 2020.