Analysis of Alcohol Consumption on the Risk of Breast Cancer Among U.S. Women

Anna Rothenberg

## Abstract

*Objective:* Researchers have identified hormonal, lifestyle, and environmental factors that may increase the risk of breast cancer. This study aims to investigate differential effects of alcohol consumption on the risk of breast cancer during adulthood.

*Approach and Results:* A cross sectional study was conducted to determine the effects of alcohol consumption at three different levels – low/none, medium, or high – on the risk of breast cancer among U.S. women. The data was collected via the Center for Disease Control’s National Health and Nutrition Examination Survey (NHANES 2015-2016). Participants in the study were 1,560 women, aged 20-80 years. Odds ratios were calculated and then adjusted for smoking, body mass index (BMI), history of pregnancy, use of birth control pills, use of female hormones, and use of estrogen/progestin combination pills. Using this adjusted model, the relationship between alcohol consumption and invasive breast cancer was assessed. No level of alcohol consumption was found to be related to risk of breast cancer in a significant manner.

*Conclusions:* The fact that the findings were insignificant in all categories highlights the need for a larger sample size and a better survey design for any future studies.

## Introduction

After skin cancer, breast cancer is the most commonly diagnosed cancer among women in the United States. While breast cancer can occur in both men and women, it's far more common in women and for this reason this study focused on women. For women in the U.S., breast cancer death rates are higher than those for any other cancer, besides lung cancer and while in recent years the number of deaths associated with this disease have decreased, approximately 1 in 8 women will be diagnosed with breast cancer during their lifetime. Higher prevention and survival rates are largely due to factors such as earlier detection, a new personalized approach to treatment, and a better understanding of the disease as researchers have begun to identify hormonal, lifestyle, and environmental factors that may increase the risk of breast cancer. The most significant [risk factors](https://www.breastcancer.org/risk/factors) for breast cancer are gender (being a woman) and age (growing older) however other known risk factors include family history, radiation exposure, obesity, and hormone therapy (see about adding more). Additionally, doctors estimate that about 5 to 10 percent of breast cancers are linked to gene mutations passed through generations of a family (the most well-known being the BRCA gene)

Research has also shown that people who drink excessively have a greater risk of liver disease, heart disease, depression, stroke, and stomach bleeding, as well as cancers of the oral cavity, esophagus, larynx, pharynx, liver, colon, and rectum.

## Methodology

*Study Population*

The data for the participants in this study were collected from the Center for Disease Control’s 2015-2016 National Health and Nutrition Examination Survey (NHANES) questionnaire data set. The population includes 1,560 female subjects, aged 20-80, who responded in some manner (yes/no) to both consuming alcohol and having cancer. All study participants provided informed consent to be included in the NHANES study prior to sharing health information.

*Exposure and Outcomes*

Both exposure and outcome information were self-reported by participants through the NHANES questionnaire. The primary predictor (exposure) in this study was alcohol consumption. Frequency, quantity, and age at consumption were ascertained through the question “Average number of alcoholic drinks/day in the past 12 months”**1**. One drink was classified as 12 ounces of beer, 5 ounces of wine, or 1.5 ounces of liquor. Women were classified based on drinking levels as defined by the National Institute for Alcohol Abuse and Alcoholism. 0-1 drinks per day indicated light/no drinking, 2-3 drinks per day indicated moderate drinking, and 4-15 drinks per day indicated heavy drinking.

The primary outcome in this study was the diagnosis of breast cancer, determined by two questions; if the participant had ever been told by a doctor or other health professional that he/she had cancer or a malignancy of any kind, and a follow up question about what kind of cancer it was**1**.

Other variables taken into account for confounding were age, race, sex, smoking, body mass index (BMI), history of pregnancy, use of birth control pills, use of female hormones, and use of estrogen/progestin combination pills.

*Statistical analysis*

All of the data was analyzed using SAS 9.3 programming. Prior to analyzing data, all participants who had missing data for either the exposure or outcome were excluded. A covariate analysis was done using logistic regression in order assess for confounding using variables that had been previously studied and found to have a relationship with breast cancer, and the crude odds ratio was compared to the adjusted odds ratio for each potential confounder.

## Results

Prior to adjusting for potential covariates, there were no significant associations between alcohol consumption and breast cancer at any level (Table 3). This was determined by the fact that the 95% confidence intervals for the crude odds ratios both include the value of 1 (Table 3). When considering the variables that entered the final model, no variables were shown to be statistically significant (Table 2).

## Discussion

It has been shown that a woman’s risk of breast cancer nearly doubles if she has a first-degree relative (mother, sister, daughter) who has been diagnosed with breast cancer **2**. Unfortunately, family history data was not available in the NHANES data set and as a result this was a limitation of the study.

Additionally, estrogen/progestin pill usage also served as a limitation as while this was a question asked in the NHANES questionnaire, only a small percentage of women answered it and when merged in the data set there were 1,415 missing answers for the 1,560 observations. Due to the small number of responses it was determined that it would not be statistically accurate to use this variable.

## References

Here is what I used to write my paper

1. “NHANES 2015-2016 Questionnaire Data.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Questionnaire&CycleBeginYear=2015.
2. “U.S. Breast Cancer Statistics.” *Breastcancer.org*, www.breastcancer.org/symptoms/understand\_bc/statistics.
3. “Drinking Levels Defined.” *National Institute on Alcohol Abuse and Alcoholism*, U.S. Department of Health and Human Services, www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking.

## Tables

|  |  |  |
| --- | --- | --- |
|  | No Breast Cancer  N=1,474 | History of Breast Cancer  N=86 |
| Median Age, years |  |  |
|  | 45±17.5 | |
| Alcohol Consumption, n (%) |  |  |
| Low/none                      Medium      Heavy | 654, (42.0%)  642 (41.2%)  178 (11.4%) | 26 (1.67%)  20 (1.28%)  2 (0.128%) |
| Race, n (%) |  |  |
| White                      Non-White | 558 (35.8%)  243 (15.6%) | 40 (2.56%)  11 (0.705%) |
| History of Pregnancy, n (%) |  |  |
| Yes                      No | 1,198 (76.8%)  276 (17.7%) | 74 (4.74%)  12 (0.769%) |
| Menopausal Status, n (%) |  |  |
| Pre                     Current/post | 745 (47.8%)  729 (46.7%) | 2 (0.128%)  84 (5.38%) |
| BMI, n (%) |  |  |
| Underweight                     Normal                     Overweight    Obese | 35 (2.24%)  408 (26.2%)  386 (24.7%)  645 (41.4%) | 5 (0.321%)  20 (1.28%)  26 (1.67%)  35 (2.24%) |
| Smoking Status, n (%) |  |  |
| Non-smoker                     Smoker | 22 (1.41%)  552 (35.4%) | 0 (0.00%)  40 (2.56%) |
| Birth Control Usage, n (%) |  |  |
| Yes                      No | 1,085 (69.6%)  389 (24.9%) | 53 (3.40%)  33 (2.12%) |
| Female Hormone Usage, n (%) |  |  |
| Yes                      No | 231 (14.8%)  1,243 (79.7%) | 23 (1.47%)  63 (4.04%) |

|  |  |  |  |
| --- | --- | --- | --- |
| Covariate | Estimate | P-Value | R2 |
| Age | 0.0552 (± 0.2456) | 0.822 | 0.0376 |
| Race | -0.4271 (± 0.2379) | 0.0727 | 0.0022 |
| Smoking | -0.2876 (± 0.3041) | 0.3443 | 0.0046 |
| BMI | -0.4185 (± 0.2372) | 0.0776 | 0.0027 |
| History of Pregnancy | -0.3959 (± 0.2363) | 0.0938 | 0.0036 |
| Menopausal Status | -0.0986 (± 0.2403) | 0.6814 | 0.0322 |
| Birth Control Pill Usage | -0.4109 (± 0.2359) | 0.0815 | 0.0028 |
| Female Hormone Usage | -0.3623 (± 0.2388) | 0.1291 | 0.0042 |

Table 1. Summary of Demographics (N=1,560).

Table 2. Risk factors contributing to the relation between alcohol consumption and breast cancer.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Odds Ratio | 95% Confidence Interval | P-Value | R2 |
| Moderate Alcohol Consumption |  |  |  |  |
| Crude                    Adjusted  \*models adjusted for smoking, BMI, history of pregnancy, menopausal status, birth control usage, female hormone usage, and estrogen/progestin combination pill usage. | 0.784  1.388 | 0.433-1.418  0.564-3.419 | 0.3709  0.2769 | 0.0028  0.0421 |
| Heavy Alcohol Consumption |  |  |  |  |
| Crude                    Adjusted  \*models adjusted for smoking, BMI, history of pregnancy, menopausal status, birth control usage, female hormone usage, and estrogen/progestin combination pill usage. | 0.283  0.640 | 0.066-1.202  0.133-3.070 | 0.1163  0.4221 | 0.0028  0.0421 |

Table 3. Relation between alcohol consumption and odds of breast cancer. Low/no alcohol consumption was used as the reference group.

## Syntax

**data** epigroup5;

set epigroup4 (where=(RIAGENDR=**2**));

run;

**data** epigroup5;

set epigroup5;

if MCQ220 = **1** and MCQ230a = **14** or MCQ230b = **14** then breastcancer=**1**;

else breastcancer=**0**;

if MCQ220 = **.** or MCQ230a = **99** then breastcancer = **.**;

run;

**proc** **freq** data = epigroup5;

table breastcancer;

run;

**data** epigroup5;

set epigroup5;

if ALQ130 = **1** then alcohol = **0**;

else if **1** < ALQ130 <= **3** then alcohol = **1**;

else if **3** < ALQ130 <= **15** then alcohol = **2**;

if ALQ130 = **.** or ALQ130 = **999** then alcohol = **.**;

run;

**proc** **freq** data = epigroup5;

table alcohol;

run;

**proc** **freq** data = epigroup5;

table alcohol\*breastcancer;

run;

**proc** **freq** data = epigroup5;

where alcohol=**0** or alcohol=**1**;

table alcohol\*breastcancer / OR;

run;

**proc** **freq** data = epigroup5;

where alcohol=**0** or alcohol=**2**;

table alcohol\*breastcancer / OR;

run;

**data** epigroup6;

set epigroup5 (where=(alcohol=**0** or alcohol=**1** or alcohol=**2** and breastcancer=**0** or breastcancer=**1**));

run;

**proc** **contents** data = epigroup6;

run;

**proc** **univariate** data = epigroup6;

var RIDAGEYR;

run;

**proc** **univariate** data=epigroup6;

var RIDAGEYR;

histogram;

run;

proc freq;

table RIDAGEYR;

run;

**data** epigroup6;

set epigroup6;

if RIDAGEYR <= **35** then agegroup = **0**;

else if **35** < RIDAGEYR <= **54** then agegroup = **1**;

else if **55** < RIDAGEYR <= **80** then agegroup = **2**;

run;

**proc** **freq** data = epigroup6;

table agegroup;

run;

\*////////Crude analysis/////////;

proc logistic;

class alcohol (ref='0');

model breastcancer(event='1')= alcohol/rsquare;

run;

proc logistic;

model breastcancer(event='1')= alcohol agegroup/rsquare;

run;

**data** epigroup6;

set epigroup6;

if RIDRETH3 = **1** and **2** and **4** and **6** and **7** then race = **0**;

else if RIDRETH3 = **3** then race = **1**;

run;

**proc** **freq** data = epigroup6;

table race;

run;

proc logistic;

model breastcancer(event='1')= alcohol race/rsquare;

run;

**data** epigroup6;

set epigroup6;

if SMD030 = **0** then smoke = **0**;

else if SMD030 = **.** and **999** and **777** then smoke = **.**;

else smoke = **1**;

run;

**proc** **freq** data = epigroup6;

table smoke;

run;

proc logistic;

model breastcancer(event='1')= alcohol smoke/rsquare;

run;

**data** epigroup6;

set epigroup6;

if BMXBMI < **18.5** then BMI = **0**;

else if **18.5** <= BMXBMI <= **24.9** then BMI = **1**;

else if **25** <= BMXBMI <= **29.9** then BMI = **2**;

else if BMXBMI >= **30** then BMI = **3**;

run;

**proc** **freq** data = epigroup6;

table BMI;

run;

proc logistic;

model breastcancer(event='1')= alcohol BMI/rsquare;

run;

**data** epigroup6;

set epigroup6;

if RIDAGEYR < **45** then menopause = **0**;

else menopause = **1**;

run;

**proc** **freq** data = epigroup6;

table menopause;

run;

proc logistic;

model breastcancer(event='1')= alcohol menopause/rsquare;

run;

**data** epigroup6;

set epigroup6;

if RHQ131 = **1** then pregnancy = **0**;

else if RHQ131 = **7** or **9** and **.** then pregnancy = **.**;

else pregnancy = **1**;

run;

**proc** **freq** data = epigroup6;

table pregnancy;

run;

proc logistic;

model breastcancer(event='1')= alcohol pregnancy/rsquare;

run;

**data** epigroup6;

set epigroup6;

if RHQ420 = **2** then bc = **0**;

else if RHQ420 = **7** and **9** and **.** then bc = **.**;

else bc = **1**;

run;

**proc** **freq** data = epigroup6;

table bc;

run;

proc logistic;

model breastcancer(event='1')= alcohol bc/rsquare;

run;

**data** epigroup6;

set epigroup6;

if RHQ540 = **2** then fhormone = **0**;

else if RHQ540 = **7** and **9** and **.** then fhormone = **.**;

else fhormone = **1**;

run;

**proc** **freq** data = epigroup6;

table fhormone;

run;

proc logistic;

model breastcancer(event='1')= alcohol fhormone/rsquare;

run;

**data** epigroup6;

set epigroup6;

if RHQ570 = **1** then ep\_pills = **1**;

else if RHQ570 = **2** then ep\_pills = **0**;

if RHQ570 = **.** then ep\_pills = **.**;

run;

**proc** **freq** data = epigroup6;

table ep\_pills;

run;

\*/////////FINAL MODEL////////;

**proc** **logistic** data=epigroup6;

class alcohol (ref='0');

model breastcancer(event='1')= alcohol BMI smoke pregnancy bc fhormone menopause/rsquare;

run;

**data** epigroup6;

set epigroup6;

inter1 = alcohol\*BMI;

run;

proc logistic;

model breastcancer(event='1')= alcohol BMI inter1;

run;

**data** epigroup6;

set epigroup6;

inter2 = alcohol\*smoke;

run;

proc logistic;

model breastcancer(event='1')= alcohol smoke inter2;

run;

**data** epigroup6;

set epigroup6;

inter3 = alcohol\*pregnancy;

run;

proc logistic;

model breastcancer(event='1')= alcohol pregnancy inter3;

run;

**data** epigroup6;

set epigroup6;

inter4 = alcohol\*bc;

run;

proc logistic;

model breastcancer(event='1')= alcohol bc inter4;

run;

**data** epigroup6;

set epigroup6;

inter5 = alcohol\*fhormone;

run;

proc logistic;

model breastcancer(event='1')= alcohol fhormone inter5;

run;

**data** epigroup6;

set epigroup6;

inter7 = alcohol\*menopause;

run;

proc logistic;

model breastcancer(event='1')= alcohol menopause inter7;

run;

**proc** **freq** data = epigroup6;

table breastcancer;

run;

**proc** **freq** data = epigroup6;

table alcohol\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table race\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table smoke\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table BMI\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table menopause\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table pregnancy\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table bc\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table fhormone\*breastcancer;

run;

**proc** **freq** data = epigroup6;

table RIDAGEYR\*breastcancer;

run;