Independent Project Final

Bathsheba Aklilu

2025-04-06

Introduction

Breast cancer, like many other health issues, is characterized by inequitable survival rates, particularly among marginalized groups (Giaquinto et al., 2022). Disparities in health outcomes, such as age at diagnosis, quality of life, and overall survival, have been shown to vary significantly by race and ethnicity. These disparities may be exacerbated by the stressors and social determinants of health (SDOH) that individuals from marginalized communities face, in addition to their breast cancer diagnosis. Social determinants of health, such as income and access to healthcare, can contribute to increased inflammation and stress, which are often reflected in biomarkers like C-reactive protein (CRP) and cortisol levels (Antoni et al., 2020). Black and Latina/Hispanic women are disproportionately affected by these health disparities in breast cancer (Yedjou et al., 2019), yet they remain underrepresented in much of oncological research (Duma et al., 2018). C-reactive protein (CRP) is an important biomarker for inflammation, frequently utilized in clinical and epidemiological research to assess risk factors for cardiovascular disease, chronic inflammation, and overall health status. CRP levels can be influenced by various factors, including social determinants of health (SDOH) such as socioeconomic status, education, and access to healthcare. In particular, low socioeconomic status (SES) and limited access to healthcare are often associated with chronic stress, which can contribute to elevated CRP levels. While much of the research has focused on the role of SDOH in these disparities, few studies have explored how stress factors are related to these outcomes, particularly in a diverse cohort such as the NHANES dataset.

This analysis aims to investigate the relationship between CRP levels and social determinants of health among participants with history of breast cancer using the **NHANES 2021-2023 dataset**. We will focus on CRP as the dependent variable, exploring how social determinants of health such as income, education, and marital status impact inflammation. The National Health and Nutrition Examination Survey (NHANES) 2021-2023, includes detailed information on various health indicators, including CRP levels, income, and cancer history. We will examine CRP levels in relation to social determinants of health, including family income and poverty level, as well as demographic factors such as age and race/ethnicity. The primary statistical approaches utilized in this analysis include descriptive statistics, ANOVA, GLM, and regression modeling.

Research Question and Hypothesis

- **Research Question**: How do social determinants of health (SDOH) such as income, education, and marital status influence CRP levels among women with hisptory of breast cancer in the NHANES 2021-2023 cohort?
- **Null Hypothesis**: There is no association between C-reactive protein levels and social determinants of health across breast cancer suriviors by racial and ethnic groups.
- **Alternative Hypothesis**: There is a significant association between C-reactive protein levels and social determinants of health across breast cancer suriviors by racial and ethnic groups.
- **Hypothesis (alternative)**: We hypothesize that individuals with lower income, lower education levels, and divorced marital status will have higher CRP levels, reflecting higher inflammation due to chronic stress and poor health.

Aim 1: Determine the association between social determinants of health (income, education, and marital status) and C-reactive protein by race/ethnicity among women with history of breast cancer.

Data Installation

```
# Install and load necessary packages
#install.packages("tidyverse")
library(tidyverse)
library(haven)

# loading data
crp_data <- read_xpt("data/HSCRP_L.xpt")
demo_data <- read_xpt("data/DEMO_L.xpt")
income_data <- read_xpt("data/INQ_L.xpt")
medical_data <- read_xpt("data/MCQ_L.xpt")
```

Data Cleaning and Preparation for Breast Cancer Survivors

```
# Clean and prepare CRP data (select necessary columns only)
crp_data_clean <- crp_data %>%
  select(SEQN, LBXHSCRP)

# Clean and prepare demographics data (select necessary columns only)
demo_data_clean <- demo_data %>%
  select(SEQN, RIAGENDR, RIDAGEYR, RIDRETH3, RIDRETH1, INDFMPIR, DMDEDUC2, DMDMARTZ)

# Clean and prepare income data (select necessary columns only)
income_data_clean <- income_data %>%
  select(SEQN, INDFMMPI)

# Clean and prepare medical conditions data (select necessary columns only, specifically, participants
breast_cancer_data <- medical_data %>%
  filter(MCQ230A == 14) %>%
  select(SEQN, MCQ230A)

# Merge the datasets by SEQN (participant id)
merged_data <- demo_data_clean %>%
  left_join(crp_data_clean, by = "SEQN") %>%
  left_join(income_data_clean, by = "SEQN") %>%
  inner_join(breast_cancer_data, by = "SEQN")

# Change coding of variables to characters
cleaned_data_labeled <- merged_data %>%
  mutate(
    RIDRETH3 = recode(RIDRETH3,
      `1` = "Mexican American",
      `2` = "Other Hispanic",
      `3` = "Non-Hispanic White",
      `4` = "Non-Hispanic Black",
      `6` = "Non-Hispanic Asian",
```

```r
      `7` = "Other Race - Including Multi-Racial",
      .default = NA_character_
    ),
    RIDRETH1 = recode(RIDRETH1,
      `1` = "Mexican American",
      `2` = "Other Hispanic",
      `3` = "Non-Hispanic White",
      `4` = "Non-Hispanic Black",
      `5` = "Other Race - Including Multi-Racial",
      .default = NA_character_
    ),
    DMDMARTZ = recode(DMDMARTZ,
      `1` = "Married/Living with partner",
      `2` = "Widowed/Divorced/Separated",
      `3` = "Never married",
      `77` = "Refused",
      `99` = "Don't know",
      .default = NA_character_
    ),
    DMDEDUC2 = recode(DMDEDUC2,
      `1` = "Less than 9th grade",
      `2` = "9-11th grade (Includes 12th grade with no diploma)",
      `3` = "High school graduate/GED or equivalent",
      `4` = "Some college or AA degree",
      `5` = "College graduate or above",
      `7` = "Refused",
      `9` = "Don't know",
      .default = NA_character_
    )) %>%
  filter(
    DMDMARTZ != "Refused",
    DMDMARTZ != "Don't know",
    !is.na(DMDMARTZ),
    !is.na(RIDRETH1),
    !is.na(RIDRETH3),
    DMDEDUC2 != "Refused",
    DMDEDUC2 != "Don't know",
    !is.na(DMDEDUC2))


# Drop last nas if any

cleaned_compiled_allbc <- cleaned_data_labeled %>%
  drop_na()

# Create SES category from INDFMPIR
cleaned_compiled_allbc$SES <- cut(
  cleaned_compiled_allbc$INDFMPIR,
  breaks = c(-Inf, 1.30, 3.50, Inf),
  labels = c("Low", "Moderate", "High")
)
```

```
cleaned_compiled_allbc
```

```
## # A tibble: 108 x 12
##      SEQN RIAGENDR RIDAGEYR RIDRETH3         RIDRETH1 INDFMPIR DMDEDUC2 DMDMARTZ
##     <dbl>    <dbl>    <dbl> <chr>            <chr>       <dbl> <chr>    <chr>
##  1 130392        2       74 Non-Hispanic Wh~ Non-His~     3.04 College~ Married~
##  2 130407        2       73 Non-Hispanic Wh~ Non-His~     4.37 College~ Widowed~
##  3 130523        2       61 Non-Hispanic Wh~ Non-His~     5    College~ Married~
##  4 130826        2       79 Non-Hispanic Wh~ Non-His~     3.3  Some co~ Married~
##  5 131137        2       67 Other Hispanic   Other H~     4.05 High sc~ Widowed~
##  6 131169        2       78 Non-Hispanic Wh~ Non-His~     5    Some co~ Widowed~
##  7 131342        2       61 Non-Hispanic Wh~ Non-His~     4.67 College~ Married~
##  8 131450        2       80 Non-Hispanic Wh~ Non-His~     3.68 College~ Widowed~
##  9 131509        2       45 Non-Hispanic Wh~ Non-His~     5    College~ Widowed~
## 10 131554        2       69 Non-Hispanic Wh~ Non-His~     5    College~ Widowed~
## # i 98 more rows
## # i 4 more variables: LBXHSCRP <dbl>, INDFMMPI <dbl>, MCQ230A <dbl>, SES <fct>
```

##Data Visualization (CRP and Race/Ethnicity)

```
# Loading Okabe-Ito palette (colorblind-friendly color palette)
okabe_ito_colors <- palette.colors(palette = "Okabe-Ito")

# Participants with breast cancer
ggplot(cleaned_compiled_allbc, aes(x = RIDRETH3, y = LBXHSCRP, fill = RIDRETH3)) +
  geom_boxplot() +
  geom_hline(yintercept = 3.0, linetype = "dashed", color = "red") +
  labs(
    x = "Race/Ethnicity",
    y = "High-Sensitivity CRP (mg/L)",
    title = "C-Reactive Protein Levels of Breast Cancer Survivors by Race/Ethnicity"
  ) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),

  ) +
  scale_fill_manual(values = okabe_ito_colors)
```

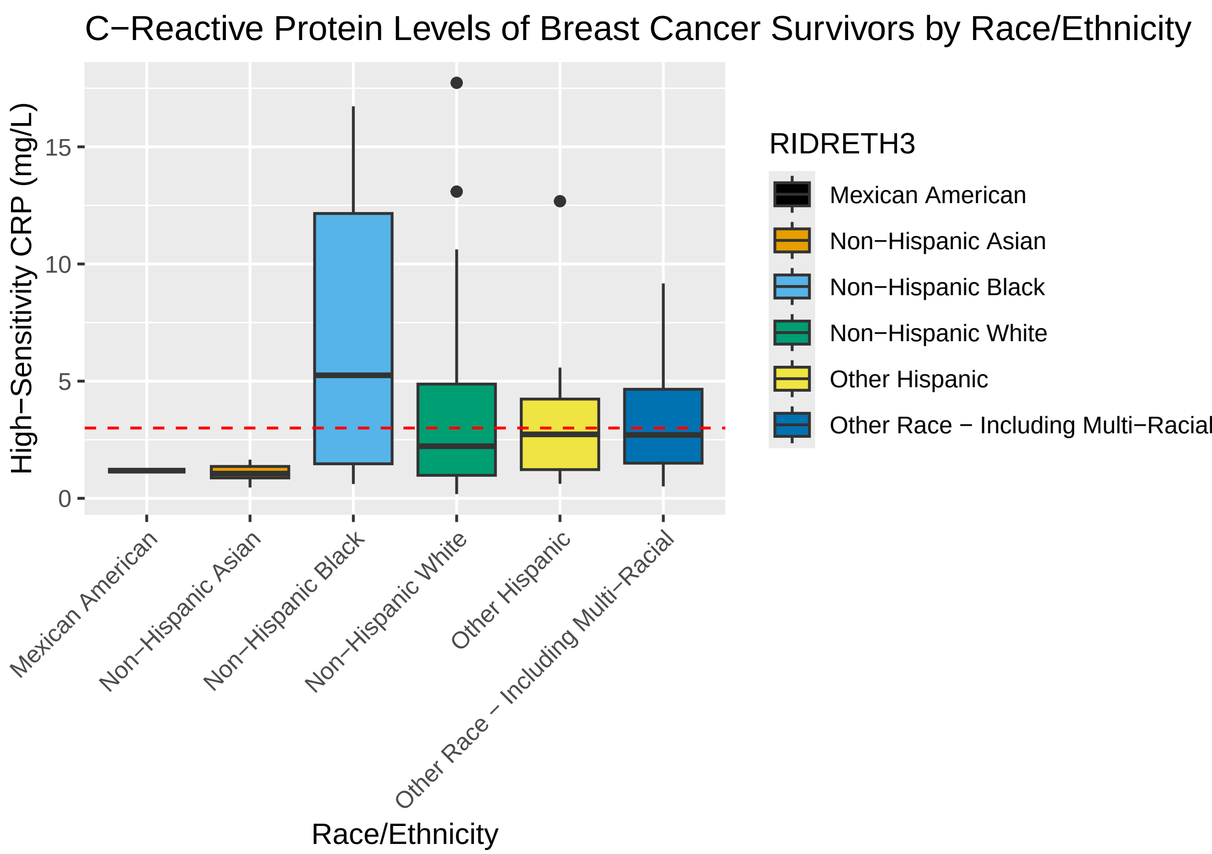

Figure 1: Boxplot depicting high-sensitivity c-reactive protein (CRP) levels (mg/L) in serum or plasma among participants with history of breast cancer stratified by race/ethnicity. C-reactive protein was collected by serum/plasma, not time-specific. Data collected through the CDC NHANES Research Program. Values above 3.0 mg/L suggest elevated cardiovascular or chronic inflammation risk and a dashed red line on the y-axis.

Interpretation: While black, white, hispanic and multiracial groups have ranges above the 0.3 threshold for high inflammation, black women are the only group to have a median above the threshold. This graph and disparity between black women and white women only confirms what many other studies have often concluded in the past (Nazmi & Victora, 2007). Overall, women of color, and especially black women typically exhibit higher inflammation and cortisol levels. With data on stage at diagnosis, type of breast cancer, etc. we would be able to investigate the relationship of race to when participants were diagnosed and have a better

##Data Visualization (CRP and Education Level)

```
##I want the groups on the x axis to go from least to highest education.
cleaned_compiled_allbc$education_ranked <- factor(cleaned_compiled_allbc$DMDEDUC2, ordered = TRUE,
  levels = c("Less than 9th grade",
             "9-11th grade (Includes 12th grade with no diploma)",
             "High school graduate/GED or equivalent",
             "Some college or AA degree",
             "College graduate or above"))

# Participants with breast cancer
ggplot(cleaned_compiled_allbc, aes(x = education_ranked, y = LBXHSCRP, fill = education_ranked)) +
  geom_boxplot() +
  geom_hline(yintercept = 3.0, linetype = "dashed", color = "red") +
```

```
  labs(
    x = "Education Level",
    y = "High-Sensitivity CRP (mg/L)",
    title = "C-Reactive Protein Levels of Breast Cancer Survivors by Education Level"
  ) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    legend.position = "none"
  )
```

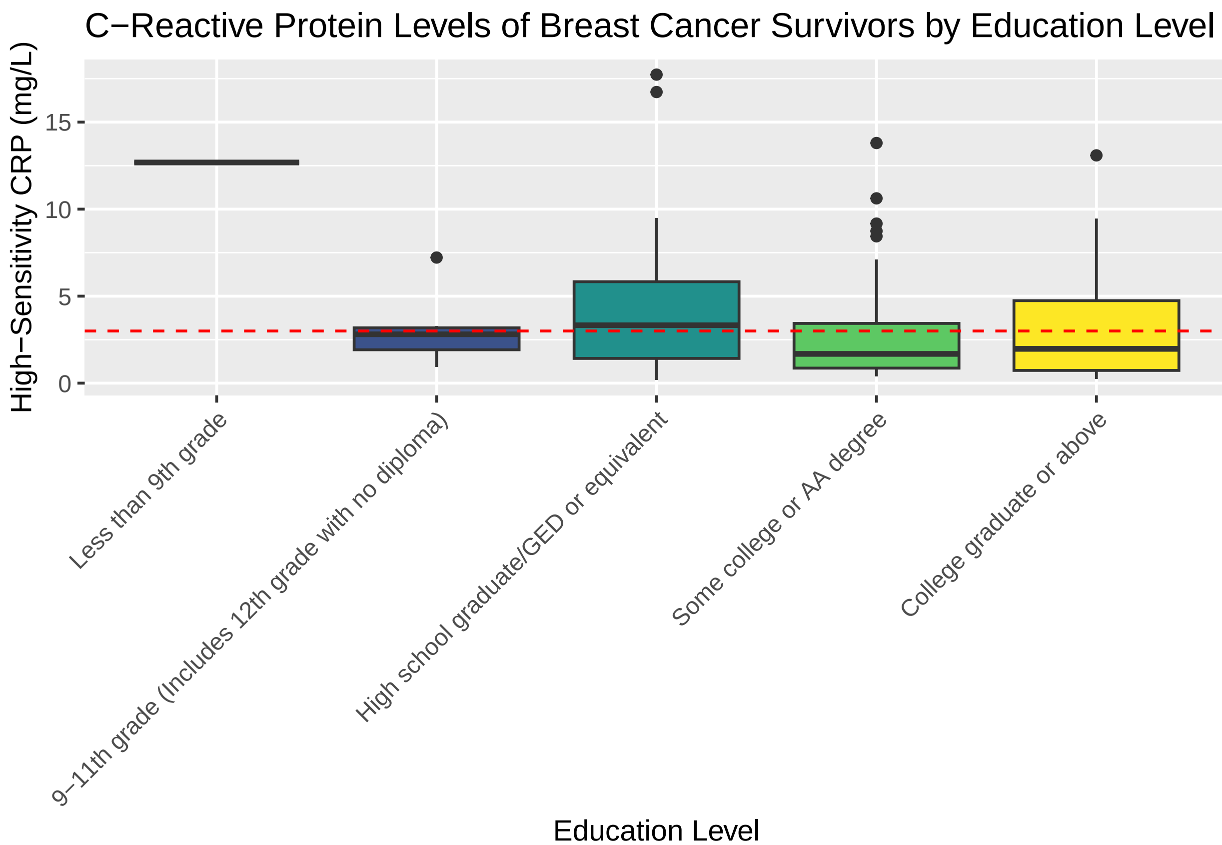

Figure 2: Boxplot depicting high-sensitivity c-reactive protein (CRP) levels (mg/L) in serum or plasma among participants with no history of breast cancer stratified by education level. Education level data was collected through survey. C-reactive protein was collected by serum/plasma, not time-specific. Data collected through the CDC NHANES Research Program. Values above 3.0 mg/L suggest elevated cardiovascular or chronic inflammation risk and a dashed red line on the y-axis.

Interpretation: Participants with high school or less than 9th grade education are the only groups to have a median above the threshold. This graph and disparity could be influenced by the jobs that participants are able to obtain and their exposure to inflammation-causing materials. Additonally this could be influenced by access to healthcare/quality of healthcare these jobs afford.

##Data Visualization (CRP and Marital Status)

```
# Participants with breast cancer
ggplot(cleaned_compiled_allbc, aes(x = DMDMARTZ, y = LBXHSCRP, fill = DMDMARTZ)) +
  geom_boxplot() +
  geom_hline(yintercept = 3.0, linetype = "dashed", color = "red") +
```

```
  labs(
    x = "Marital Status",
    y = "High-Sensitivity CRP (mg/L)",
    title = "C-Reactive Protein Levels of Breast Cancer Survivors by Marital Status"
  ) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    legend.position = "none"
  )
```

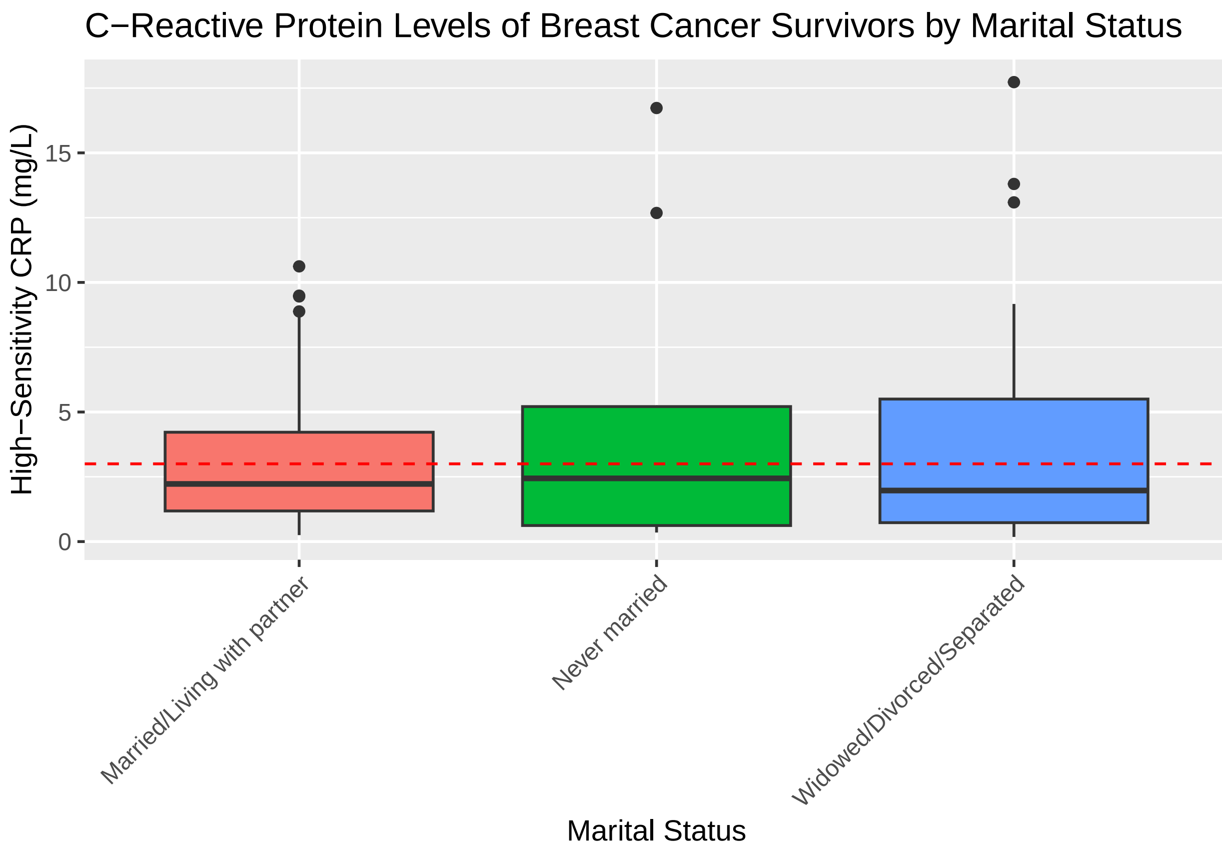

Figure 3: Boxplot depicting high-sensitivity c-reactive protein (CRP) levels (mg/L) in serum or plasma among participants with no history of breast cancer stratified by marital status. Marital status data was collected through survey. C-reactive protein was collected by serum/plasma, not time-specific. Data collected through the CDC NHANES Research Program. Values above 3.0 mg/L suggest elevated cardiovascular or chronic inflammation risk and a dashed red line on the y-axis.

Interpretation: Interestingly, while all groups range over the threshold, no groups had a median above the threshold. However, the widowed/divorced/separated group has the highest range. The lack of disparity could be due to the small sample size, or simply no association between marital status and inflammation. I would like to test social support in the future as that may be the primary factor involved here.

##CRP and Income-to-Poverty Ratio (recoded into SES level based on HHS standards)

```
# Participants with breast cancer
ggplot(cleaned_compiled_allbc, aes(x = SES, y = LBXHSCRP, fill = SES)) +
  geom_boxplot() +
  geom_hline(yintercept = 3.0, linetype = "dashed", color = "red") +
```

```r
labs(
    x = "Socioeconomic Status (recoded from Income-to-Poverty Ratio)",
    y = "High-Sensitivity CRP (mg/L)",
    title = "C-Reactive Protein Levels of Breast Cancer Survivors by Socioeconomic Status"
  ) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    legend.position = "none"
  )
```

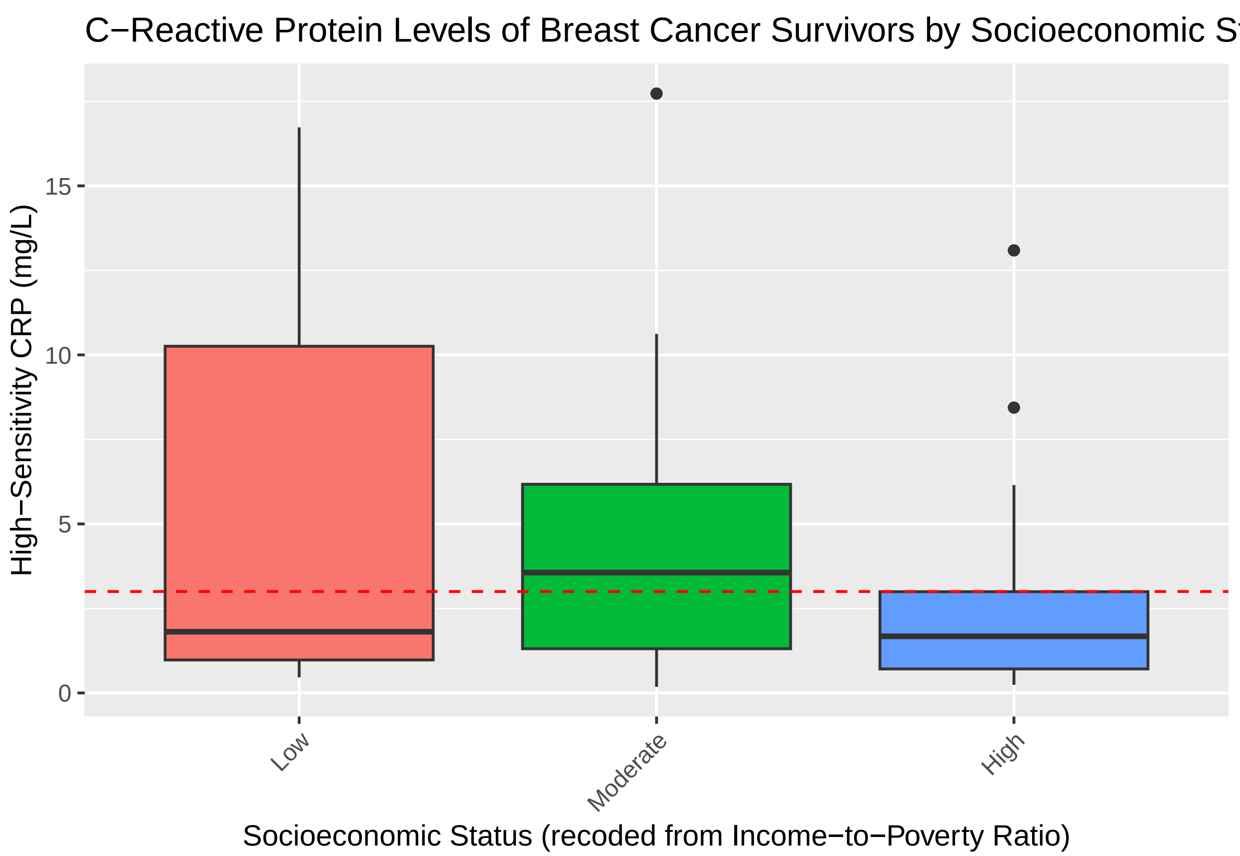

Figure 4: Boxplot depicting high-sensitivity c-reactive protein (CRP) levels (mg/L) in serum or plasma among participants with no history of breast cancer stratified by Socioeconomic Status. Socioeconomic status data was dervied from Income-to-Poverty Ratio based of Health and Human Services standards. C-reactive protein was collected by serum/plasma, not time-specific. Data collected through the CDC NHANES Research Program. Values above 3.0 mg/L suggest elevated cardiovascular or chronic inflammation risk and a dashed red line on the y-axis.

Interpretation: Interestingly, while all groups range over the threshold, moderate SES was the only group above the threshold. This could be due to the way the groups were categorized. Participants in the low socioeconomic status group qualified for government assistance, and health insurance, while the moderate and high groups did not. The government assistance with Health insurance could be providing participants with low SES access to healthcare that the moderate group may not have, or have insufficiently.

Data transformation and statistical analysis: Shapiro-Wilks, Bartlett, ANOVA(Tukeys)

```
#Because the data is naturally not normal, and CRP is normally right skewed, I transformed CRP data by

cleaned_compiled_allbc$log_CRP <- log(cleaned_compiled_allbc$LBXHSCRP + 0.01)
crp_log_aov <- aov(log_CRP ~ DMDEDUC2 + RIDRETH3, data = cleaned_compiled_allbc)

# CRP and Race/Ethnicity
# Shapiro test on residuals, ensuring over 2 participants per group
shapiro.test(residuals(crp_log_aov))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(crp_log_aov)
## W = 0.98312, p-value = 0.1885
```

```
df_filtered <- cleaned_compiled_allbc %>%
  group_by(RIDRETH3) %>%
  filter(n() >= 2) %>%
  ungroup()

# Bartlett's test for equal variances
bartlett.test(log_CRP ~ RIDRETH3, data = df_filtered)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  log_CRP by RIDRETH3
## Bartlett's K-squared = 5.1029, df = 4, p-value = 0.2769
```

```
TukeyHSD(crp_log_aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = log_CRP ~ DMDEDUC2 + RIDRETH3, data = cleaned_compiled_allbc)
##
## $DMDEDUC2
##                                                                                         dif
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)           -0.3662809
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma)  0.1207521
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                  1.5930247
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)           -0.3223604
## High school graduate/GED or equivalent-College graduate or above                        0.4870331
## Less than 9th grade-College graduate or above                                           1.9593056
## Some college or AA degree-College graduate or above                                     0.0439205
## Less than 9th grade-High school graduate/GED or equivalent                              1.4722725
## Some college or AA degree-High school graduate/GED or equivalent                       -0.4431125
## Some college or AA degree-Less than 9th grade                                          -1.9153851
##                                                                                         lwr
```

```
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)           -1.6416011
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma) -1.2376008
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                   -1.5764654
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)             -1.6189374
## High school graduate/GED or equivalent-College graduate or above                         -0.2884495
## Less than 9th grade-College graduate or above                                            -1.0074971
## Some college or AA degree-College graduate or above                                      -0.6174133
## Less than 9th grade-High school graduate/GED or equivalent                               -1.5311588
## Some college or AA degree-High school graduate/GED or equivalent                         -1.2530778
## Some college or AA degree-Less than 9th grade                                            -4.8913873
##                                                                                                 upr
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)           0.9090392
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma) 1.4791051
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                   4.7625149
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)             0.9742165
## High school graduate/GED or equivalent-College graduate or above                         1.2625158
## Less than 9th grade-College graduate or above                                            4.9261084
## Some college or AA degree-College graduate or above                                      0.7052544
## Less than 9th grade-High school graduate/GED or equivalent                               4.4757039
## Some college or AA degree-High school graduate/GED or equivalent                         0.3668526
## Some college or AA degree-Less than 9th grade                                            1.0606170
##                                                                                               p adj
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)           0.9305408
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma) 0.9991637
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                   0.6311298
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)             0.9580056
## High school graduate/GED or equivalent-College graduate or above                         0.4113666
## Less than 9th grade-College graduate or above                                            0.3594151
## Some college or AA degree-College graduate or above                                      0.9997361
## Less than 9th grade-High school graduate/GED or equivalent                               0.6529965
## Some college or AA degree-High school graduate/GED or equivalent                         0.5518256
## Some college or AA degree-Less than 9th grade                                            0.3859777
##
## $RIDRETH3
##                                                          diff        lwr
## Non-Hispanic Asian-Mexican American                 -0.3361514 -3.6512325
## Non-Hispanic Black-Mexican American                  0.9864698 -2.3286113
## Non-Hispanic White-Mexican American                  0.4754712 -2.6123553
## Other Hispanic-Mexican American                      0.2119221 -3.0691572
## Other Race - Including Multi-Racial-Mexican American   0.7308986 -2.5841825
## Non-Hispanic Black-Non-Hispanic Asian                1.3226212 -0.4493642
## Non-Hispanic White-Non-Hispanic Asian                0.8116225 -0.4863920
## Other Hispanic-Non-Hispanic Asian                    0.5480735 -1.1594544
## Other Race - Including Multi-Racial-Non-Hispanic Asian  1.0670500 -0.7049354
## Non-Hispanic White-Non-Hispanic Black               -0.5109986 -1.8090131
## Other Hispanic-Non-Hispanic Black                   -0.7745477 -2.4820755
## Other Race - Including Multi-Racial-Non-Hispanic Black -0.2555712 -2.0275565
## Other Hispanic-Non-Hispanic White                   -0.2635491 -1.4720857
## Other Race - Including Multi-Racial-Non-Hispanic White  0.2554275 -1.0425870
## Other Race - Including Multi-Racial-Other Hispanic     0.5189765 -1.1885513
##                                                         upr     p adj
## Non-Hispanic Asian-Mexican American                 2.9789297 0.9996924
## Non-Hispanic Black-Mexican American                 4.3015509 0.9538753
## Non-Hispanic White-Mexican American                 3.5632976 0.9976734
```

```
## Other Hispanic-Mexican American                        3.4930015 0.9999668
## Other Race - Including Multi-Racial-Mexican American   4.0459797 0.9875775
## Non-Hispanic Black-Non-Hispanic Asian                  3.0946065 0.2613991
## Non-Hispanic White-Non-Hispanic Asian                  2.1096370 0.4593911
## Other Hispanic-Non-Hispanic Asian                      2.2556013 0.9369889
## Other Race - Including Multi-Racial-Non-Hispanic Asian 2.8390354 0.5023338
## Non-Hispanic White-Non-Hispanic Black                  0.7870159 0.8614925
## Other Hispanic-Non-Hispanic Black                      0.9329802 0.7741069
## Other Race - Including Multi-Racial-Non-Hispanic Black 1.5164142 0.9982988
## Other Hispanic-Non-Hispanic White                      0.9449876 0.9881810
## Other Race - Including Multi-Racial-Non-Hispanic White 1.5534420 0.9926234
## Other Race - Including Multi-Racial-Other Hispanic     2.2265044 0.9496260
```

```
crp_log_aov <- aov(log_CRP ~ RIDRETH3, data = cleaned_compiled_allbc)

##CRP and education
# Shapiro test on residuals, ensuring over 2 participants per group
shapiro.test(residuals(crp_log_aov))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(crp_log_aov)
## W = 0.97835, p-value = 0.07532
```

```
df_filtered <- cleaned_compiled_allbc %>%
  group_by(DMDEDUC2) %>%
  filter(n() >= 2) %>%
  ungroup()

# Bartlett's test for equal variances
bartlett.test(log_CRP ~ DMDEDUC2, data = df_filtered)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  log_CRP by DMDEDUC2
## Bartlett's K-squared = 2.1389, df = 3, p-value = 0.5441
```

```
TukeyHSD(crp_log_aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = log_CRP ~ RIDRETH3, data = cleaned_compiled_allbc)
##
## $RIDRETH3
##                                                               diff        lwr
## Non-Hispanic Asian-Mexican American                    -0.16216589 -3.5312498
## Non-Hispanic Black-Mexican American                     1.16045527 -2.2086286
## Non-Hispanic White-Mexican American                     0.54537569 -2.5927515
## Other Hispanic-Mexican American                         0.69798076 -2.6365475
```

```
## Other Race - Including Multi-Racial-Mexican American    0.72357855 -2.6455053
## Non-Hispanic Black-Non-Hispanic Asian                    1.32262116 -0.4782299
## Non-Hispanic White-Non-Hispanic Asian                    0.70754159 -0.6116176
## Other Hispanic-Non-Hispanic Asian                        0.86014666 -0.8751969
## Other Race - Including Multi-Racial-Non-Hispanic Asian  0.88574444 -0.9151066
## Non-Hispanic White-Non-Hispanic Black                   -0.61507958 -1.9342388
## Other Hispanic-Non-Hispanic Black                       -0.46247451 -2.1978180
## Other Race - Including Multi-Racial-Non-Hispanic Black -0.43687672 -2.2377278
## Other Hispanic-Non-Hispanic White                        0.15260507 -1.0756187
## Other Race - Including Multi-Racial-Non-Hispanic White  0.17820285 -1.1409564
## Other Race - Including Multi-Racial-Other Hispanic      0.02559778 -1.7097457
##                                                              upr     p adj
## Non-Hispanic Asian-Mexican American                     3.2069180 0.9999923
## Non-Hispanic Black-Mexican American                     4.5295391 0.9167822
## Non-Hispanic White-Mexican American                     3.6835029 0.9958910
## Other Hispanic-Mexican American                         4.0325090 0.9902422
## Other Race - Including Multi-Racial-Mexican American    4.0926624 0.9890252
## Non-Hispanic Black-Non-Hispanic Asian                   3.1234722 0.2787769
## Non-Hispanic White-Non-Hispanic Asian                   2.0267008 0.6278561
## Other Hispanic-Non-Hispanic Asian                       2.5954902 0.7028010
## Other Race - Including Multi-Racial-Non-Hispanic Asian 2.6865955 0.7096038
## Non-Hispanic White-Non-Hispanic Black                   0.7040796 0.7538378
## Other Hispanic-Non-Hispanic Black                       1.2728690 0.9712873
## Other Race - Including Multi-Racial-Non-Hispanic Black 1.3639744 0.9809994
## Other Hispanic-Non-Hispanic White                       1.3808288 0.9991757
## Other Race - Including Multi-Racial-Non-Hispanic White 1.4973621 0.9987647
## Other Race - Including Multi-Racial-Other Hispanic     1.7609413 1.0000000
```

```r
crp_log_aov <- aov(log_CRP ~ DMDEDUC2, data = cleaned_compiled_allbc)

#CRP and marital status
# Shapiro test on residuals
shapiro.test(residuals(crp_log_aov))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(crp_log_aov)
## W = 0.9861, p-value = 0.3263
```

```r
df_filtered <- cleaned_compiled_allbc %>%
  group_by(DMDMARTZ) %>%
  filter(n() >= 2) %>%
  ungroup()

# Bartlett's test for equal variances
bartlett.test(log_CRP ~ DMDMARTZ, data = df_filtered)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  log_CRP by DMDMARTZ
## Bartlett's K-squared = 3.2644, df = 2, p-value = 0.1955
```

```r
TukeyHSD(crp_log_aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = log_CRP ~ DMDEDUC2, data = cleaned_compiled_allbc)
##
## $DMDEDUC2
##                                                                                          dif
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)            -0.3662809
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma)  0.1207521
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                   1.5930247
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)            -0.3223604
## High school graduate/GED or equivalent-College graduate or above                         0.4870331
## Less than 9th grade-College graduate or above                                            1.9593056
## Some college or AA degree-College graduate or above                                      0.0439205
## Less than 9th grade-High school graduate/GED or equivalent                               1.4722725
## Some college or AA degree-High school graduate/GED or equivalent                        -0.4431125
## Some college or AA degree-Less than 9th grade                                           -1.9153851
##                                                                                          lwr
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)            -1.6480359
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma) -1.2444545
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                  -1.5924575
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)            -1.6254794
## High school graduate/GED or equivalent-College graduate or above                        -0.2923623
## Less than 9th grade-College graduate or above                                           -1.0224665
## Some college or AA degree-College graduate or above                                     -0.6207501
## Less than 9th grade-High school graduate/GED or equivalent                              -1.5463130
## Some college or AA degree-High school graduate/GED or equivalent                        -1.2571646
## Some college or AA degree-Less than 9th grade                                           -4.9064031
##                                                                                         upr
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)            0.9154740
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma) 1.4859588
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                  4.7785069
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)            0.9807586
## High school graduate/GED or equivalent-College graduate or above                        1.2664286
## Less than 9th grade-College graduate or above                                           4.9410778
## Some college or AA degree-College graduate or above                                     0.7085912
## Less than 9th grade-High school graduate/GED or equivalent                              4.4908581
## Some college or AA degree-High school graduate/GED or equivalent                        0.3709394
## Some college or AA degree-Less than 9th grade                                           1.0756328
##                                                                                       p adj
## College graduate or above-9-11th grade (Includes 12th grade with no diploma)            0.9319691
## High school graduate/GED or equivalent-9-11th grade (Includes 12th grade with no diploma) 0.9991839
## Less than 9th grade-9-11th grade (Includes 12th grade with no diploma)                  0.6363278
## Some college or AA degree-9-11th grade (Includes 12th grade with no diploma)            0.9589079
## High school graduate/GED or equivalent-College graduate or above                        0.4172830
## Less than 9th grade-College graduate or above                                           0.3652219
## Some college or AA degree-College graduate or above                                     0.9997425
## Less than 9th grade-High school graduate/GED or equivalent                              0.6580273
## Some college or AA degree-High school graduate/GED or equivalent                        0.5574894
## Some college or AA degree-Less than 9th grade                                           0.3918553
```

```r
crp_log_aov <- aov(log_CRP ~ SES, data = cleaned_compiled_allbc)

#CRP and SES
# Shapiro test on residuals
shapiro.test(residuals(crp_log_aov))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(crp_log_aov)
## W = 0.98284, p-value = 0.1788
```

```r
df_filtered <- cleaned_compiled_allbc %>%
  group_by(SES) %>%
  filter(n() >= 2) %>%
  ungroup()

# Bartlett's test for equal variances
bartlett.test(log_CRP ~ SES, data = df_filtered)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  log_CRP by SES
## Bartlett's K-squared = 1.6704, df = 2, p-value = 0.4338
```

```r
TukeyHSD(crp_log_aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = log_CRP ~ SES, data = cleaned_compiled_allbc)
##
## $SES
##                     diff        lwr        upr     p adj
## Moderate-Low   0.1652746 -0.6945044  1.0250536 0.8913926
## High-Low      -0.5157135 -1.3606214  0.3291945 0.3187334
## High-Moderate -0.6809880 -1.1794203 -0.1825558 0.0044149
```

```r
crp_log_aov <- aov(log_CRP ~ RIDRETH3, data = cleaned_compiled_allbc)
```

Shapiro test an batrlett both indicate not-normal distribution.

Statistical Analysis Continued

- **Kruskal-Wallis Test**: Traditionally, we cannot run ANOVA because many groups violate normality (which is why I transformed the data).

```r
# Kruskal-Wallis for Education
kruskal.test(LBXHSCRP ~ DMDEDUC2, data = cleaned_compiled_allbc)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  LBXHSCRP by DMDEDUC2
## Kruskal-Wallis chi-squared = 6.2027, df = 4, p-value = 0.1845
```

```r
# Kruskal-Wallis for Marital status
kruskal.test(LBXHSCRP ~ DMDMARTZ, data = cleaned_compiled_allbc)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  LBXHSCRP by DMDMARTZ
## Kruskal-Wallis chi-squared = 0.062187, df = 2, p-value = 0.9694
```

```r
# Kruskal-Wallis for SES
kruskal.test(LBXHSCRP ~ SES, data = cleaned_compiled_allbc)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  LBXHSCRP by SES
## Kruskal-Wallis chi-squared = 10.366, df = 2, p-value = 0.00561
```

```r
# Kruskal-Wallis for Race/Ethnicity
kruskal.test(LBXHSCRP ~ RIDRETH3, data = cleaned_compiled_allbc)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  LBXHSCRP by RIDRETH3
## Kruskal-Wallis chi-squared = 5.6454, df = 5, p-value = 0.3423
```

- **Spearman's**: Due to lack of normal distribution, we are no longer using pearsons, and instead using spearmans.

```r
# Correlation between CRP and income-to-poverty ratio
cor.test(cleaned_compiled_allbc$LBXHSCRP, cleaned_compiled_allbc$INDFMPIR, method = "spearman")
```

```
## Warning in cor.test.default(cleaned_compiled_allbc$LBXHSCRP,
## cleaned_compiled_allbc$INDFMPIR, : Cannot compute exact p-value with ties
```

```
##
##  Spearman's rank correlation rho
##
## data:  cleaned_compiled_allbc$LBXHSCRP and cleaned_compiled_allbc$INDFMPIR
## S = 265448, p-value = 0.005683
```

```
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.2644349
```

```r
# Correlation between CRP and Education (by ranking education first)
cleaned_compiled_allbc$education_ranked <- factor(cleaned_compiled_allbc$DMDEDUC2, ordered = TRUE,
  levels = c("Less than 9th grade",
             "9-11th grade (includes 12th grade with no diploma)",
             "High school graduate/GED or equivalent",
             "Some college or AA degree",
             "College graduate or above"))

cor.test(as.numeric(cleaned_compiled_allbc$education_ranked), cleaned_compiled_allbc$LBXHSCRP, method =
```

```
## Warning in
## cor.test.default(as.numeric(cleaned_compiled_allbc$education_ranked), : Cannot
## compute exact p-value with ties
```

```
##
##  Spearman's rank correlation rho
##
## data:  as.numeric(cleaned_compiled_allbc$education_ranked) and cleaned_compiled_allbc$LBXHSCRP
## S = 205415, p-value = 0.1048
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.1615156
```

Cross Referencing with GLM

```r
glm_crp <- glm(log_CRP ~ DMDEDUC2 + DMDMARTZ + SES + RIDRETH3,
               data = cleaned_compiled_allbc, family = gaussian())
glm_gamma_crp1 <- glm(LBXHSCRP ~ DMDEDUC2,
                      data = cleaned_compiled_allbc,
                      family = Gamma(link = "log"))
glm_gamma_crp2 <- glm(LBXHSCRP ~ DMDMARTZ,
                      data = cleaned_compiled_allbc,
                      family = Gamma(link = "log"))
glm_gamma_crp3 <- glm(LBXHSCRP ~ SES,
                      data = cleaned_compiled_allbc,
                      family = Gamma(link = "log"))
glm_gamma_crp4 <- glm(LBXHSCRP ~ RIDRETH3,
                      data = cleaned_compiled_allbc,
                      family = Gamma(link = "log"))

summary(glm_crp)
```

```
##
## Call:
## glm(formula = log_CRP ~ DMDEDUC2 + DMDMARTZ + SES + RIDRETH3,
```

```
##     family = gaussian(), data = cleaned_compiled_allbc)
##
## Coefficients:
##                                              Estimate Std. Error t value
## (Intercept)                                   0.23861    1.23192   0.194
## DMDEDUC2College graduate or above             0.08362    0.59071   0.142
## DMDEDUC2High school graduate/GED or equivalent  0.25590    0.56128   0.456
## DMDEDUC2Less than 9th grade                   2.14084    1.27821   1.675
## DMDEDUC2Some college or AA degree            -0.24157    0.54523  -0.443
## DMDMARTZNever married                        -0.55391    0.42651  -1.299
## DMDMARTZWidowed/Divorced/Separated           -0.21258    0.22045  -0.964
## SESModerate                                   0.17692    0.41594   0.425
## SESHigh                                      -0.54698    0.43527  -1.257
## RIDRETH3Non-Hispanic Asian                   -0.01461    1.15665  -0.013
## RIDRETH3Non-Hispanic Black                    1.39191    1.16230   1.198
## RIDRETH3Non-Hispanic White                    0.82450    1.05619   0.781
## RIDRETH3Other Hispanic                        0.71528    1.16187   0.616
## RIDRETH3Other Race - Including Multi-Racial   1.06792    1.11883   0.954
##                                              Pr(>|t|)
## (Intercept)                                    0.8468
## DMDEDUC2College graduate or above              0.8877
## DMDEDUC2High school graduate/GED or equivalent   0.6495
## DMDEDUC2Less than 9th grade                    0.0973 .
## DMDEDUC2Some college or AA degree              0.6587
## DMDMARTZNever married                          0.1972
## DMDMARTZWidowed/Divorced/Separated             0.3374
## SESModerate                                    0.6716
## SESHigh                                        0.2120
## RIDRETH3Non-Hispanic Asian                     0.9899
## RIDRETH3Non-Hispanic Black                     0.2341
## RIDRETH3Non-Hispanic White                     0.4370
## RIDRETH3Other Hispanic                         0.5396
## RIDRETH3Other Race - Including Multi-Racial    0.3423
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.05489)
##
##     Null deviance: 123.54  on 107  degrees of freedom
## Residual deviance:  99.16  on  94  degrees of freedom
## AIC: 327.27
##
## Number of Fisher Scoring iterations: 2
```

```
summary(glm_gamma_crp1)
```

```
##
## Call:
## glm(formula = LBXHSCRP ~ DMDEDUC2, family = Gamma(link = "log"),
##     data = cleaned_compiled_allbc)
##
## Coefficients:
##                                              Estimate Std. Error t value
## (Intercept)                                   1.13676    0.41190   2.760
```

```
## DMDEDUC2College graduate or above             -0.05270    0.43850  -0.120
## DMDEDUC2High school graduate/GED or equivalent  0.44984    0.46705   0.963
## DMDEDUC2Less than 9th grade                     1.40326    1.08978   1.288
## DMDEDUC2Some college or AA degree              -0.01938    0.44581  -0.043
##                                                Pr(>|t|)
## (Intercept)                                     0.00685 **
## DMDEDUC2College graduate or above               0.90457
## DMDEDUC2High school graduate/GED or equivalent  0.33772
## DMDEDUC2Less than 9th grade                     0.20075
## DMDEDUC2Some college or AA degree               0.96542
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 1.017965)
##
##     Null deviance: 112.71  on 107  degrees of freedom
## Residual deviance: 105.52  on 103  degrees of freedom
## AIC: 488.73
##
## Number of Fisher Scoring iterations: 5
```

```r
summary(glm_gamma_crp2)
```

```
##
## Call:
## glm(formula = LBXHSCRP ~ DMDMARTZ, family = Gamma(link = "log"),
##     data = cleaned_compiled_allbc)
##
## Coefficients:
##                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           1.1187     0.1415   7.906 2.81e-12 ***
## DMDMARTZNever married                 0.4858     0.3623   1.341    0.183
## DMDMARTZWidowed/Divorced/Separated   0.1612     0.2011   0.801    0.425
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 1.001193)
##
##     Null deviance: 112.71  on 107  degrees of freedom
## Residual deviance: 110.57  on 105  degrees of freedom
## AIC: 490.55
##
## Number of Fisher Scoring iterations: 6
```

```r
summary(glm_gamma_crp3)
```

```
##
## Call:
## glm(formula = LBXHSCRP ~ SES, family = Gamma(link = "log"), data = cleaned_compiled_allbc)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.6675     0.3037   5.490 2.81e-07 ***
```

```
## SESModerate  -0.1671     0.3365  -0.497   0.6204
## SESHigh      -0.8379     0.3307  -2.534   0.0128 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.9225807)
##
##     Null deviance: 112.708  on 107  degrees of freedom
## Residual deviance:  99.362  on 105  degrees of freedom
## AIC: 477.28
##
## Number of Fisher Scoring iterations: 6
```

```r
summary(glm_gamma_crp4)
```

```
##
## Call:
## glm(formula = LBXHSCRP ~ RIDRETH3, family = Gamma(link = "log"),
##     data = cleaned_compiled_allbc)
##
## Coefficients:
##                                              Estimate Std. Error t value
## (Intercept)                                   0.16551    0.96180   0.172
## RIDRETH3Non-Hispanic Asian                   -0.08701    1.03886  -0.084
## RIDRETH3Non-Hispanic Black                    1.79247    1.03886   1.725
## RIDRETH3Non-Hispanic White                    1.04610    0.96765   1.081
## RIDRETH3Other Hispanic                        1.18256    1.02821   1.150
## RIDRETH3Other Race - Including Multi-Racial  1.11031    1.03886   1.069
##                                              Pr(>|t|)
## (Intercept)                                    0.8637
## RIDRETH3Non-Hispanic Asian                     0.9334
## RIDRETH3Non-Hispanic Black                     0.0875 .
## RIDRETH3Non-Hispanic White                     0.2822
## RIDRETH3Other Hispanic                         0.2528
## RIDRETH3Other Race - Including Multi-Racial    0.2877
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.9250596)
##
##     Null deviance: 112.71  on 107  degrees of freedom
## Residual deviance: 102.04  on 102  degrees of freedom
## AIC: 486.56
##
## Number of Fisher Scoring iterations: 6
```

Conclusion:

While there is variation in C-reactive protein levels of breast cancer survivors in terms of whether or not CRP was above normal levels, no comparison had a statistically significant correlation except for CRP and SES between the moderate and high socioeconomic status. This could be due to the fact that one of the factors used to categorize SES is eligibility for government assisted programs like EBT and healthcare. With participants that were moderate, they did not have access to healthcare through the government. This could be a driving factor as access to healthcare has been shown to have significant associations with poor cancer outcomes, which could be contributing to inflammation in the body. Overall, we hope this research can be used to inform public policymakers on the associations of SDOH and inflammation.

References:

Anandapadmanathan, & Kresten. (2019, July 24). How to add dashed horizontal line with label in ggplot [Online forum post]. Stack Overflow. https://stackoverflow.com/questions/57177608/how-to-add-dashed-horizontal-line-with-label-in-ggplot

Antoni, M. H., Lechner, S. C., Kilbourn, K. M., & Phillips, K. A. (2020). Behavioral, physical, and psychological predictors of cortisol and C-reactive protein in breast cancer survivors: A longitudinal study. Psycho-Oncology, 29(8), 1237-1245. https://doi.org/10.1002/pon.5397

Bewick, V., Cheek, L., & Ball, J. (2003). Statistics review 7: Correlation and regression. Critical care (London,England), 7(6), 451–459. https://doi.org/10.1186/cc2401

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. Journal of Economic Surveys, 22(1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Cohen, S., Doyle, W. J., & Baum, A. (2006). Socioeconomic status is associated with stress hormones. Psychosomatic medicine, 68(3), 414–420. Doi: 10.1097/01.psy.0000221236.37158.b9. PMID: 16738073.

Coughlin S. S. (2019). Social determinants of breast cancer risk, stage, and survival. Breast cancer research and treatment, 177(3), 537–548. https://doi.org/10.1007/s10549-019-05340-7

DeSantis, C. E., Ma, J., Sauer, A. G., Newman, L. A., & Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. CA: A Cancer Journal for Clinicians, 67(6), 439-448. https://doi.org/10.3322/caac.21412

Ding, Z., Mangino, M., Aviv, A., Spector, T., Durbin, R., & UK10K Consortium (2014). Estimating telomere length from whole genome sequence data. Nucleic acids research, 42(9), e75. https://doi.org/10.1093/nar/gku181

Duma, N., Vera Aguilera, J., Paludo, J., Haddox, C. L., Gonzalez Velez, M., Wang, Y., Leventakos, K., Hubbard, J. M., Mansfield, A. S., Go, R. S., & Adjei, A. A. (2018). Representation of Minorities and Women in Oncology Clinical Trials: Review of the Past 14 Years. Journal of oncology practice, 14(1), e1–e10. https://doi.org/10.1200/JOP.2017.025288

Current hematologic malignancy reports, 18(6), 284–291. https://doi.org/10.1007/s11899-023-00717-4

Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, et al.. Breast cancer statistics, 2022. CA: Cancer J Clin (2022) 0:1–18. doi: 10.3322/caac.21754

Guo, L., Liu, S., Zhang, S., Chen, Q., Zhang, M., Quan, P., Lu, J., & Sun, X. (2015). C-reactive protein and risk of breast cancer: A systematic review and meta-analysis. Scientific reports, 5, 10508. https://doi.org/10.1038/srep10508

He, X.-Y., Gao, Y., Ng, D., Michalopoulou, E., George, S., Adrover, J. M., . . . & Egeblad, M. (2023). Chronic stress increases metastasis via neutrophil-mediated changes to the microenvironment. Nature, 616 (7956), 563–572. https://doi.org/10.1038/s41586-023-06020-3

Hopper, J.L., Dite, G.S., MacInnis, R.J. et al. Age-specific breast cancer risk by body mass index and familial risk: prospective family study cohort (ProF-SC). Breast Cancer Res 20, 132 (2018). https://doi.org/10.1186/s13058-018-1056-1

Islami, F., Ward, E. M., Sung, H., Cronin, K. A., Tangka, F. K. L., Sherman, R. L., Zhao, J., Anderson, R. N., Henley, S. J., Yabroff, K. R., Jemal, A., & Benard, V. B. (2021). Annual Report to the Nation on the Status of Cancer, Part 1: National Cancer Statistics. JNCI: Journal of the National Cancer Institute, 113(12), 1648–1669. https://doi.org/10.1093/jnci/djab131

PMC7048405. Mikkelsen, M. K., Lindblom, N. A. F., Dyhl-Polk, A., Juhl, C. B., Johansen, J. S., & Nielsen, D. (2022). Systematic review and meta-analysis of C-reactive protein as a biomarker in breast cancer. Critical 3 Reviews in Clinical Laboratory Sciences, 59(7), 480–500. https://doi.org/10.1080/10408363.2022.2050886

Nazmi, A., & Victora, C. G. (2007). Socioeconomic and racial/ethnic differentials of C-reactive protein levels: A systematic review of population-based studies - BMC Public Health. BioMed Central. https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-7-212

Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. Journal of health and social behavior, 51 Suppl, S28–S40. doi: 10.1177/0022146510383498.PMID: 20943581.

Solorio, S., Murillo-Ortíz, B., Hernández-González, M., Guillén-Contreras, J., Arenas-Aranda, D., Solorzano-Zepeda, F. J., Ruiz-Avila, R., Mora-Villalpando, C., de la Roca-Chiapas, J. M., & MalacaraHernández, J. M. (2011). Association between telomere length and C-reactive protein and the development of coronary collateral circulation in patients with coronary artery disease. Angiology, 62(6), 467–472. doi: 10.1177/0003319710398007. PMID: 21441231.

Wang, F., Giskeødegård, G. F., Skarra, S., Engstrøm, M. J., Hagen, L., Geisler, J., Mikkola, T. S., Tikkanen, M. J., Debik, J., Reidunsdatter, R. J., & Bathen, T. F. (2023). Association of serum cortisol and cortisone levels and risk of recurrence after endocrine treatment in breast cancer. Clinical and experimental medicine, 23(7), 3883–3893. https://doi.org/10.1007/s10238-023-01109-x

Williams, D. R., Priest, N., & Anderson, N. B. (2016). Understanding associations among race, socioeconomic status, and health: Patterns and prospects. Health psychology : official journal of the Division of Health Psychology, American Psychological Association, 35(4), 407–411. doi: 10.1037/hea0000242. PMID: 27018733; PMCID: PMC4817358.

Wong, J. Y., De Vivo, I., Lin, X., Fang, S. C., & Christiani, D. C. (2014). The relationship between inflammatory biomarkers and telomere length in an occupational prospective cohort study. PloS one, 9(1), e87348. doi: 10.1371/journal.pone.0087348. PMID: 24475279; PMCID: PMC3903646.

Yedjou, C. G., Sims, J. N., Miele, L., Noubissi, F., Lowe, L., Fonseca, D. D., Alo, R. A., Payton, M., & Tchounwou, P. B. (2019). Health and Racial Disparity in Breast Cancer. Advances in Experimental Medicine and Biology, 1152, 31. https://doi.org/10.1007/978-3-030-20301-6_3