Exploring the Associations between HDL Cholesterol Levels, Family History of Heart Attack Modifiers, and Incident of Coronary Heart Disease [CHD] in Older Adults ≥ 65 years old

Sophie Whikehart [smwhikeh@uw.edu], Tessa Fujisaki [tfuj@uw.edu], Thania Solar [tsolar@uw.edu], and Sian Siam [ssiam@uw.edu]

1 BACKGROUND

Cardiovascular diseases, including coronary heart disease (CHD) and stroke, account for 17.9 million deaths per year, making them the leading cause of death worldwide, and the leading cause of morbidity and mortality in the U.S.[1,2] Dyslipidemia is a well-known risk factor for CHD, but investigations differ on whether the importance of cholesterol levels in predicting CHD diminishes with age and their impact on the older population remains unclear. This study aims to determine which variables are associated with CHD, if HDL cholesterol levels are associated with CHD, and if family history of heart attack (FHHA) modifies this association. By addressing these factors, we hope to contribute to the advancement of preventive measures and therapeutic interventions, reducing the global burden of CHD.[3]

2 METHODS

This study utilized data from the Cardiovascular Health Study, a longitudinal investigation in a cohort of 3200 adults aged ≥ 65 over 6 years. R Studio (Version 2023.06.2+561) was used to analyze the relationships between CHD and variables such as FHHA, smoking status (SS), and HDL levels. For descriptive analysis, we created a summary table for the quantitative continuous variables that produced the mean, median, range and skewness. We used the results of this table to report either the mean or IQR when compared to incidence of CHD based on the skewness. For the categorical data, we created a summary table with frequency distribution and % of variables for FHHA, CHD, and SS. We then conducted a univariate analysis to estimate the distribution of each variable separately for frequencies and distribution through bar charts of CHD, FHHA and SS, and histograms for HDL, insulin and age. For bivariate analysis we created a bar chart for FHHA, HDL, and CHD, and box plots to compare CHD by HDL and age by CHD. Finally, two summary tables showed variables of age, alcohol consumption, HDL, serum insulin, SS, and FHHA when compared to CHD. We divided individuals based on HDL levels of < or ≥ 40 mg/dL.[4] We stratified the relationship between HDL and CHD by FHHA (yes or no) to compare the proportions to determine if FHHA was an effect modifier.

3 RESULTS

The results of the descriptive analysis for quantitative variables [Fig. 1] showed us that age and HDL were both slightly positively skewed and confirmed with histogram [Fig. 2&3] so mean should be reported. Alcohol consumption was highly skewed with lots of zero values, so we divided it into categories to view frequency distribution. Serum insulin levels were also heavily positively skewed with a lot of small values and a few larger values so we reported the median (IQR) [Fig. 4]. The categorical variable summary table [Fig. 5]

found that overall for CHD, 86.5% reported no, 13.5% reported yes. For FHHA, 69.5% reported no and 30.8% reported yes. For SS, 48.7% never smoked, 39.6% are former smokers, and 11.8% are current smokers. For alcohol consumption, 48% consumed 0 alcoholic drinks/week 38.8% consumed >0 to 7, and 13.2% consumed 7. Bar charts were created to view frequency for FFHA [Fig. 6], SS [Fig. 7] and alcohol consumption [Fig. 8]. Boxplot visualization of HDL by CHD [Fig. 9] shows there are outliers outside of the median and IQR. Histogram visualization on HDL confirms the distribution [Fig. 3]. The two sample t-test for HDL and CHD [Fig. 10] showed the t-value: 6.1893, degrees of freedom (df): 600.74, p-value: 1.119e-09 with a 95% CI for the difference in means (3.190391, 6.156090). The results of the bivariate analysis [Fig. 11] show that those with CHD have a slightly higher median age [72] compared to those without. Higher percentage of individuals with CHD consume > 7 drinks/week than those without CHD. Mean HDL levels are lower [51 mg/dL] in individuals with CHD than those without. Medium serum insulin level is higher in individuals with CHD[13 IU/mL] than those without[12 IU/mL], a higher percentage of individuals with CHD are current smokers [12.8%], and a higher percentage of individuals with CHD have FHHA [16.5%].

For the inferential analysis, a table for CHD and HDL alone [Fig. 12] was created to visualize the percent frequency of those with an HDL level of either < or ≥ 40 mg/dL by if they had CHD or not and a bar chart was made [Fig. 13]. For those who had an HDL of <40 mg/dL, 80% had CHD, and for those with an HDL of ≥ 40 mg/dL, 88% had CHD. We then analyzed if FHHA modifies the relationship between HDL and CHD by creating two tables that stratified the HDL and CHD relationship by if the individual had a FHHA or not (Fig. 14 FFHA = Yes and Fig. 15 FFHA = No). When comparing those who did and didn't have a FHHA and CHD with an HDL of <40 mg/dL, their frequencies were 25% and 17%, respectively. The 8% difference between the two groups demonstrates a large enough difference for FHHA to be an effect modifier. Between those that did and didn't have a FHHA and CHD with an HDL of ≥ 40 mg/dL, the values were 15% and 11%, with a 4% difference.

4 CONCLUSION

The study findings clarify that HDL levels are significantly associated with CHD, with lower HDL levels being associated with an increased risk of CHD. Additionally, FHHA was found to modify the association between HDL and CHD, suggesting that individuals with a FHHA may be more susceptible to the adverse effects of low HDL levels. Those with CHD tend to be older, have FHHA, lower HDL cholesterol, and higher serum insulin levels, alcohol consumption and prevalence of smoking. These findings emphasize the significance of public health interventions to promote healthy lifestyles, consider FHHA in assessing CHD risk, and tailor preventive and treatment strategies accordingly. Further research is required to understand underlying mechanisms and improve global CHD prevention measures.

5 APPENDIX 1

Figure 1 Conduct descriptive analysis for the quantitative variables by creating a summary table for the quantitative variables [age, alcoh, hdl, insulin]

Variable	N = 3,200
Age (y)	
Mean (SD)	71.6 (4.4)
Median (IQR)	71.0 (68.0, 75.0)
Min to Max	65.0, 82.0
Skewness	0.5
Alcohol consumption (drinks/wk)	
Mean (SD)	2.6 (6.5)
Median (IQR)	0.0 (0.0, 1.3)
Min to Max	0.0, 77.0
Skewness	4.4
HDL cholesterol (mg/dL)	
Mean (SD)	54.9 (15.6)
Median (IQR)	53.0 (44.0, 64.0)
Min to Max	15.0, 149.0
Skewness	0.9
Serum insulin (IU/ml)	
Mean (SD)	16.1 (20.0)
Median (IQR)	13.0 (10.0, 17.0)
Min to Max	3.0, 400.0
Skewness	12.8

Figure 2 Histogram to show skewness distribution of Alcohol Consumption (drinks/wk)

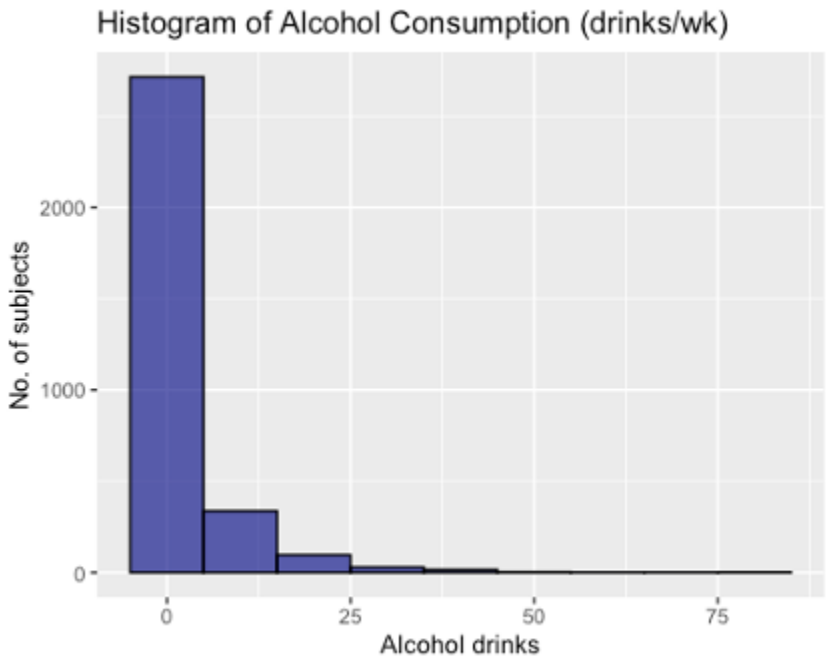

Figure 3 Histogram to show skewness distribution of HDL

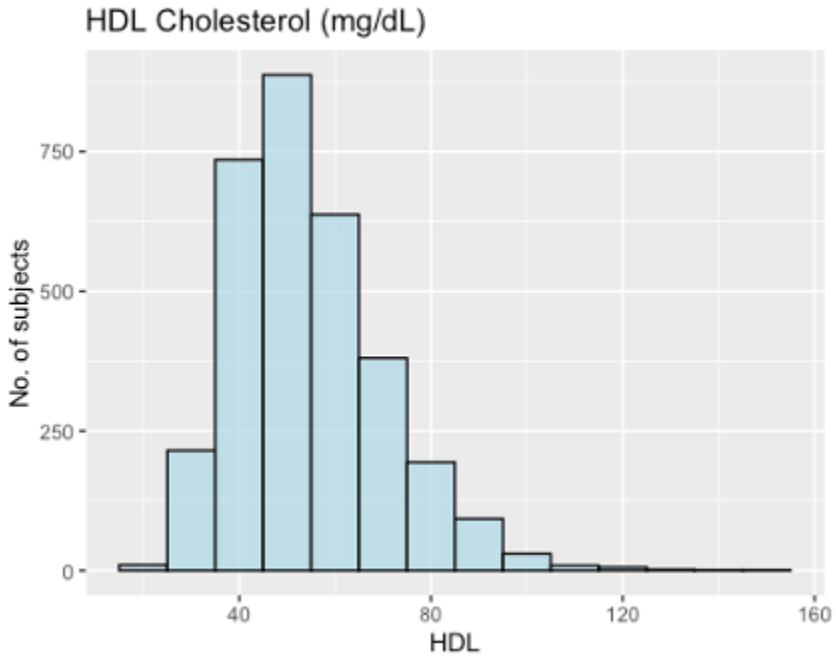

Figure 4 Histogram to show skewness distribution of serum insulin levels

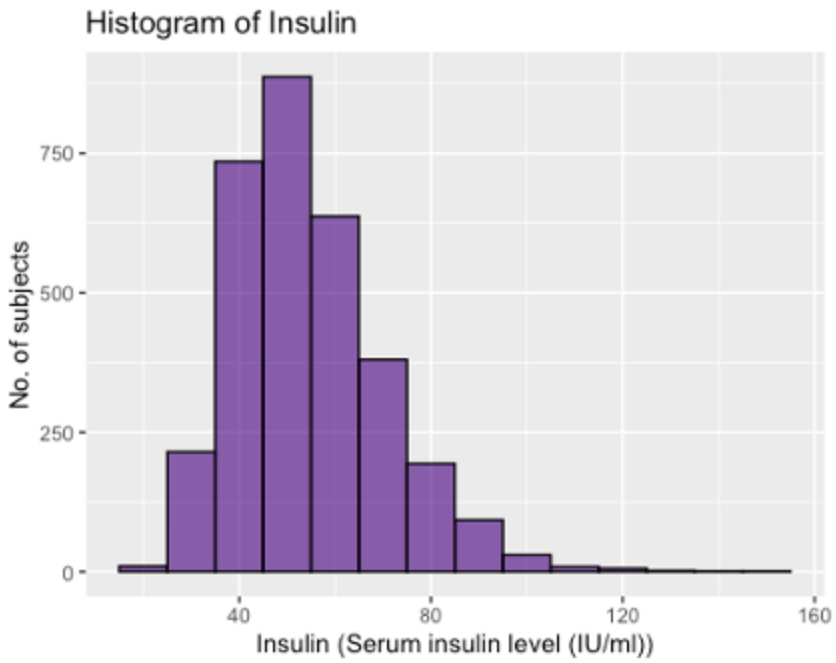

Figure 5 Frequency distribution of categorical variables

Variable	N = 3,200
Family history MI, n (%)	
No	2,216 (69.3)
Yes	984 (30.8)
Coronary heart disease, n (%)	
No	2,769 (86.5)
Yes	431 (13.5)
Smoking status, n (%)	
Never	1,558 (48.7)
Former	1,266 (39.6)
Current	376 (11.8)
Average consumption of alcohol per week, n (%)	
0	1,537 (48.0)
>0 to 7	1,240 (38.8)

>7	423 (13.2)

Figure 6 Bar chart of Family History of Heart Attack

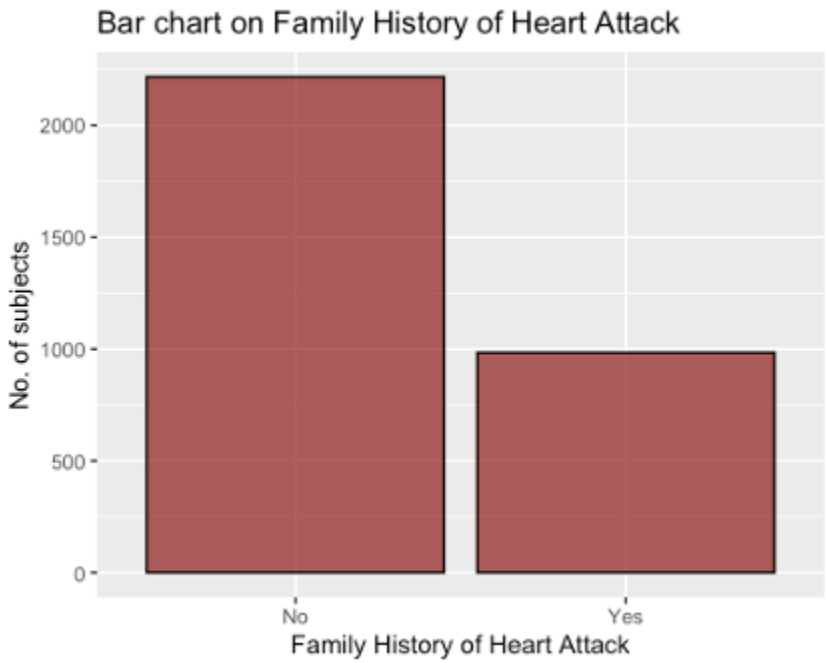

Figure 7 Bar chart on smoking status

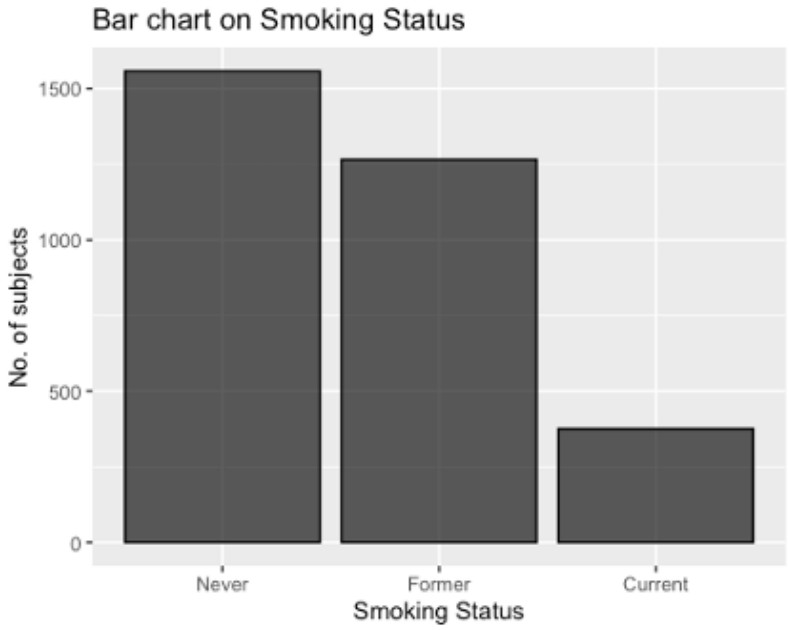

Figure 8 Bar chart to show alcohol consumption

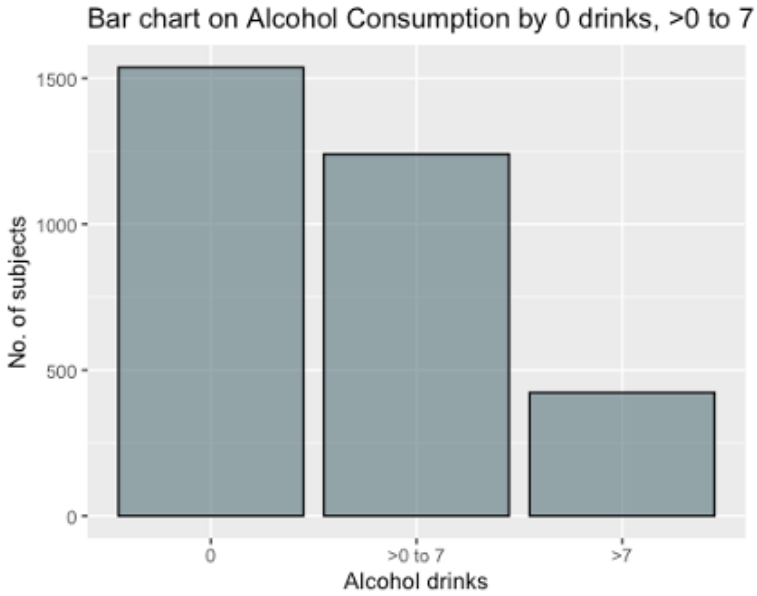

Figure 9 Box plot of CHD and HDL Cholestrol

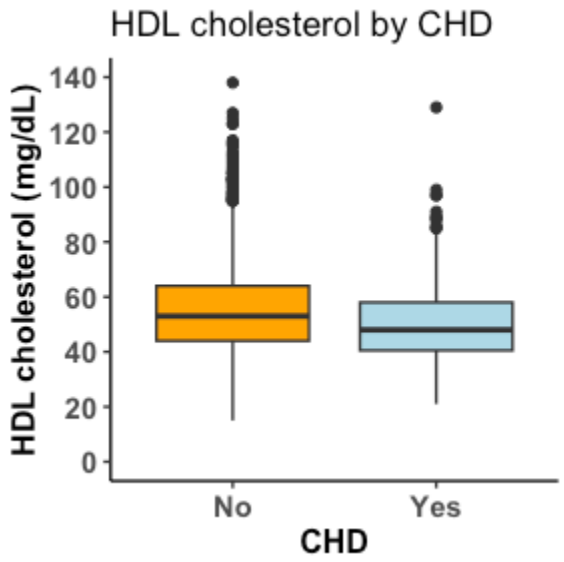

Figure 10 Two sample t-test to determine if HDL cholesterol levels are associated with CHD

```
Welch Two Sample t-test

data:  hdl by incchd.f
t = 6.1893, df = 600.74, p-value = 1.119e-09
alternative hypothesis: true difference in means between group No and group Yes is not equal to 0
95 percent confidence interval:
 3.190391 6.156090
sample estimates:
 mean in group No mean in group Yes
         55.54099          50.86775
```

Figure 11 Final summary table reporting HDl cholesterol, alcohol, serum insulin, smoking status and family history of heart attack by CHD

		Coronary Heart Disease		
Variable	**Overall**, N = 3,200	**No**, N = 2,769	**Yes**, N = 431	**p-value**[1]
Age (years), Median (IQR)	71 (68, 75)	71 (68, 75)	72 (69, 77)	<0.001
Average number of alcoholic drinks per week, n (%)				0.046
0	1,537 (100.0)	1,311 (85.3)	226 (14.7)	
>0 to 7	1,240 (100.0)	1,078 (86.9)	162 (13.1)	
>7	423 (100.0)	380 (89.8)	43 (10.2)	
HDL cholesterol (mg/dL), Mean (SD)	55 (16)	56 (16)	51 (14)	<0.001
Serum insulin (IU/ml), Median (IQR)	13 (10, 17)	12 (9, 17)	13 (10, 19)	0.002
Smoking status, n (%)				0.005
Never	1,558 (100.0)	1,376 (88.3)	182 (11.7)	
Former	1,266 (100.0)	1,065 (84.1)	201 (15.9)	
Current	376 (100.0)	328 (87.2)	48 (12.8)	

Family history of heart attack, n (%)				<0.001
No	2,216 (100.0)	1,947 (87.9)	269 (12.1)	
Yes	984 (100.0)	822 (83.5)	162 (16.5)	

[1] Wilcoxon rank sum test; Pearson's Chi-squared test

Figure 12 Overall HDL cholesterol by CHD

```
(-Inf,40] (40, Inf]
        550      2650


<40 >=40
550 2650
```

Overall HDL and CHD		
Characteristic	**No**, N = 2,769	**Yes**, N = 431
HDL cholesterol, n (%)		
<40	442 (80%)	108 (20%)
>=40	2,327 (88%)	323 (12%)

Figure 13 Bar chart on categorized HDL

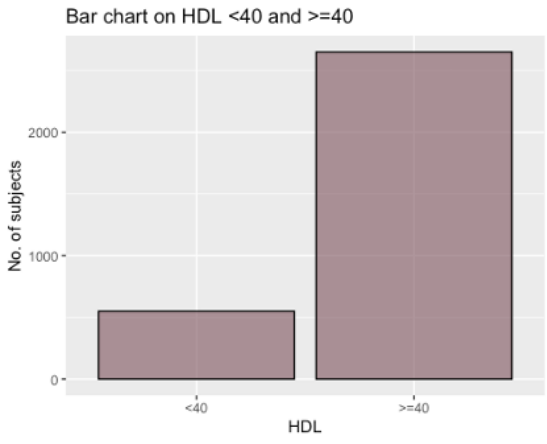

Figure 14 HDL cholesterol by CHD - Family History of Heart Attack = Yes

```
(-Inf,40] (40, Inf]
         550      2650


<40 >=40
550 2650
```

	Family History of Heart Attack = Yes to CHD	
Variable	**No**, N = 822	**Yes**, N = 162
HDL cholesterol (mg/dL), n (%)		
<40	135 (75%)	44 (25%)
>=40	687 (85%)	118 (15%)

Figure 15 HDL cholesterol by CHD - Family History of Heart Attack = No

```
(-Inf,40] (40, Inf]
         550      2650


<40 >=40
550 2650
```

	Family History of Heart Attack = No to CHD	
Variable	**No**, N = 1,947	**Yes**, N = 269
HDL cholesterol (mg/dL), n (%)		
<40	307 (83%)	64 (17%)
>=40	1,640 (89%)	205 (11%)

APPENDIX 2

```
knitr::opts_chunk$set(echo = TRUE,
                      warning = FALSE,
                      message = FALSE,
                      comment = "")
# load libraries
library(e1071) # Needed for skewness()
library(tidyverse) # Need to summarize data and for %>% function

library(ggplot2) # Needed to create graphs using the ggplot2 package
library(flextable) # Used to create a nice summary table in Word

library(gtsummary) # Used to create a nice summary table in Word

#import data
chdclass <- read.csv(file = "chd3200 (1).csv")
#create factor variables
# MI family history
chdclass$fhha.f <- factor(chdclass$fhha, levels = c(0,1),
                      labels = c("No", "Yes"))
# Coronary heart disease (CHD)
chdclass$incchd.f <- factor(chdclass$incchd, levels = c(0,1),
                      labels = c("No", "Yes"))
# Smoking status
chdclass$smoke.f <- factor(chdclass$smoke, levels = c(1,2,3),
                           labels = c("Never", "Former", "Current"))
```

Figure 1 Conduct descriptive analysis for the quantitative variables by creating a summary table for the quantitative variables [age, alcoh, hdl, insulin]

```
chdclass %>% select(age, alcoh, hdl, insulin) %>%
  tbl_summary(missing = "ifany",
              type = all_continuous() ~ "continuous2",
              statistic = all_continuous() ~ c("{mean} ({sd})",
                                               "{median} ({p25}, {p75})",
                                               "{min}, {max}",
                                               "{skewness}"),
              digits = list(all_continuous() ~ 1),
              label = list(age ~ "Age (y)",
                           alcoh  ~ "Alcohol consumption (drinks/wk)",
                           hdl ~ "HDL cholesterol (mg/dL)",
                           insulin ~ "Serum insulin (IU/ml)")
              ) %>%
  modify_header(label = "**Variable**") %>%
  add_stat_label(label = all_continuous() ~ c("Mean (SD)",
                                              "Median (IQR)",
                                              "Min to Max", "Skewness"))
```

Figure 2 Histogram to show distribution of alcohol

```
ggplot(data=chdclass, aes(x=alcoh)) +
geom_histogram(color="black", fill="blue4", binwidth=10, position="identity",
alpha=0.7) +
labs(x="Alcohol drinks",
y="No. of subjects",
title="Histogram of Alcohol Consumption (drinks/wk)")
```

Figure 3 Histogram to show distribution of hdl

```
ggplot(data=chdclass, aes(x=hdl)) +
geom_histogram(color="black", fill="lightblue", binwidth=10,
position="identity", alpha=0.7) +
labs(x="HDL",
y="No. of subjects",
title="HDL Cholesterol (mg/dL)")
```

```
#create categories
chdclass$alcoh <- cut(chdclass$alcoh, breaks = c(-Inf,0,7.000,Inf))

#check categories
table(chdclass$alcoh, exclude=FALSE)
```

```
(-Inf,0]    (0,7] (7, Inf]
      1537  1240        423
```

```
#create a factor variable
chdclass$alcoh.f <- factor(chdclass$alcoh,
                        levels = c("(-Inf,0]", "(0,7]",
                        "(7, Inf]"),
                        labels = c("0", ">0 to 7", ">7"))
#check factor variable
table(chdclass$alcoh.f)
```

```
     0 >0 to 7   >7
  1537     1240  423
```

Figure 4 Histogram to show distribution of serum insulin levels

```
ggplot(data=chdclass, aes(x=hdl)) +
geom_histogram(color="black", fill="purple4", binwidth=10, position="identity",
alpha=0.7) +
labs(x="Insulin (Serum insulin level (IU/ml))",
y="No. of subjects",
title="Histogram of Insulin")
```

Figure 5 Frequency distribution of categorical variables

```
theme_gtsummary_journal(journal = "jama") # Use the JAMA theme

chdclass %>% select(fhha.f, incchd.f, smoke.f, alcoh.f) %>%
  tbl_summary(missing = "ifany",
              type = list(fhha.f ~ "categorical",
                          incchd.f ~ "categorical"),
              digits = list(all_categorical() ~ c(0,1)),
              label = list(fhha.f ~ "Family history MI",
                          incchd.f ~ "Coronary heart disease",
                          smoke.f ~ "Smoking status",
                          alcoh.f ~ "Average consumption of alcohol per week")
              ) %>%
  modify_header(label = "**Variable**")
```

Figure 6 Bar chart of Family History of Heart Attack

```
ggplot(data=chdclass, aes(x=fhha.f)) +
geom_bar(color="black", fill="red4", binwidth=10, position="identity",
alpha=0.7) +
labs(x="Family History of Heart Attack",
y="No. of subjects",
title="Bar chart on Family History of Heart Attack")
```

Figure 7 Bar chart on smoking status

```
ggplot(data=chdclass, aes(x=smoke.f)) +
geom_bar(color="black", fill="black", binwidth=10, position="identity",
alpha=0.7) +
labs(x="Smoking Status",
y="No. of subjects",
title="Bar chart on Smoking Status")
```

Figure 8 Bar chart to show alcohol consumption

```
ggplot(data=chdclass, aes(x=alcoh.f)) +
geom_bar(color="black", fill="lightblue4", binwidth=10, position="identity",
alpha=0.7) +
labs(x="Alcohol drinks",
y="No. of subjects",
title="Bar chart on Alcohol Consumption by 0 drinks, >0 to 7 and >7 drinks per
week")
```

Figure 9 Box plot of CHD and HDL Cholestrol

```
ggplot(data=chdclass, aes(x=incchd.f, y=hdl, fill=incchd.f)) +
  geom_boxplot() +
  labs(x="CHD",
      y="HDL cholesterol (mg/dL)",
      title="HDL cholesterol by CHD") +
  scale_y_continuous(limits=c(0, 140), breaks=seq(0,150,20)) +
  theme_classic() +
  theme(legend.position="none") +
   theme(
      legend.position = "none",
      legend.title = element_text(colour="black", size=10,
                                  face="bold"),
      legend.text = element_text(colour="black", size=10,
                                  face="bold"),
      axis.title.x = element_text(face="bold", size=12),
      axis.text.x = element_text(face="bold", size=11),
      axis.title.y = element_text(face="bold", size=12),
      axis.text.y = element_text(face="bold", size=11)) +
  scale_fill_manual(values=c("orange","lightblue"))
```

Figure 10 Two sample t-test to determine if HDL cholesterol levels are associated with CHD

```
# t.test
t.test(hdl ~ incchd.f, conf.level=0.95, var.equal = FALSE, data=chdclass,
alternative= "two.sided")
```

```
      Welch Two Sample t-test

data:  hdl by incchd.f
t = 6.1893, df = 600.74, p-value = 1.119e-09
alternative hypothesis: true difference in means between group No and group Yes
is not equal to 0
95 percent confidence interval:
 3.190391 6.156090
sample estimates:
 mean in group No mean in group Yes
      55.54099          50.86775
```

Figure 11 Final summary table reporting HDl cholesterol, alcohol, serum insulin, smoking status and family history of heart attack by CHD

```
theme_gtsummary_journal(journal = "jama") # Use the JAMA theme

chdclass %>% select(incchd.f, age, alcoh.f, hdl, insulin, smoke.f, fhha.f) %>%
  tbl_summary(by = incchd.f,
              missing = "ifany",
              statistic = list(age ~ c("{median} ({p25}, {p75})"),
                               hdl ~ c("{mean} ({sd})"),
                               insulin ~ c("{median} ({p25}, {p75})")
                               ),
              percent = "row",
              type = list(fhha.f ~ "categorical",
                          alcoh.f ~ "categorical"),
              digits = list(all_categorical() ~ c(0,1)),
              label = list(hdl ~ "HDL cholesterol (mg/dL)",
                           insulin ~ "Serum insulin (IU/ml)",
                           smoke.f ~ "Smoking status",
                           age ~ "Age (years)",
                           alcoh.f ~ "Average number of alcoholic drinks per week",
                           fhha.f ~ "Family history of heart attack")
              ) %>%
  add_overall() %>%
  modify_header(label = "**Variable**") %>%
  modify_spanning_header(stat_1 ~ "Coronary Heart Disease",
                         stat_2 ~ "Coronary Heart Disease") %>%
  add_p()
```

Figure 12 Overall HDL cholesterol by CHD

```
#Categorizing HDL cholesterol
chdclass$hdlcat <- cut(chdclass$hdl, breaks = c(-Inf,40,Inf))

table(chdclass$hdlcat, exclude = FALSE)


(-Inf,40] (40, Inf]
      550    2650

# Create a factor variable
chdclass$hdlcat.f <- factor(chdclass$hdlcat,
                            levels = c("(-Inf,40]", "(40, Inf]"),
                            labels = c("<40", ">=40"))
```

```
# Check factor variable
table(chdclass$hdlcat.f)


 <40 >=40
 550 2650

#Create cross-tab for overall HDL cholesterol and CHD incidence
chdclass |>
  select(incchd.f, hdlcat.f) |>
  tbl_summary(by = incchd.f,
              missing = "no",
              percent = "row",
              statistic = all_categorical() ~ "{n} ({p}%)",
              label = list(incchd.f ~ "Incidence of CHD",
                           hdlcat.f ~ "HDL cholesterol")) %>%
  modify_spanning_header(stat_1 ~ "Overall HDL and CHD")
```

Figure 13 Bar chart on categorized HDL

```
ggplot(data=chdclass, aes(x=hdlcat.f)) +
geom_bar(color="black", fill="pink4", binwidth=10, position="identity",
alpha=0.7) +
labs(x="HDL",
y="No. of subjects",
title="Bar chart on HDL <40 and >=40")
```

Figure 14 HDL cholesterol by CHD - Family History of Heart Attack = Yes

```
#Categorizing HDL cholesterol
chdclass$hdlcat <- cut(chdclass$hdl, breaks = c(-Inf,40,Inf))

table(chdclass$hdlcat, exclude = FALSE)


(-Inf,40] (40, Inf]
      550   2650

# Create a factor variable
chdclass$hdlcat.f <- factor(chdclass$hdlcat,
                            levels = c("(-Inf,40]", "(40, Inf]"),
                            labels = c("<40", ">=40"))

# Check factor variable
table(chdclass$hdlcat.f)


 <40 >=40
 550 2650

#Create cross-tab for FFHA = YES by HDL cholesterol and CHD incidence
chdclass %>%
  filter(fhha.f == "Yes") %>%
  select(incchd.f, hdlcat.f) %>%
  tbl_summary(by = incchd.f,
              missing = "no",
```

```
            percent = "row",
            statistic = all_categorical() ~ "{n} ({p}%)",
            label = list(incchd.f ~ "Incidence of CHD",
                         hdlcat.f ~ "HDL cholesterol (mg/dL)")) %>%
modify_header(label = "**Variable**") %>%
  modify_spanning_header(stat_1 ~ "Family History of Heart Attack = Yes to CHD")
```

Figure 15 HDL cholesterol by CHD - Family History of Heart Attack = No

```
#Categorizing HDL cholesterol
chdclass$hdlcat <- cut(chdclass$hdl, breaks = c(-Inf,40,Inf))

table(chdclass$hdlcat, exclude = FALSE)


(-Inf,40] (40, Inf]
      550   2650

# Create a factor variable
chdclass$hdlcat.f <- factor(chdclass$hdlcat,
                            levels = c("(-Inf,40]", "(40, Inf]"),
                            labels = c("<40", ">=40"))

# Check factor variable
table(chdclass$hdlcat.f)


 <40 >=40
 550 2650

#Create cross-tab for FFHA = NO by HDL cholesterol and CHD incidence
chdclass %>%
  filter(fhha.f == "No") %>%
  select(incchd.f, hdlcat.f) %>%
  tbl_summary(by = incchd.f,
            missing = "no",
            percent = "row",
            statistic = all_categorical() ~ "{n} ({p}%)",
            label = list(incchd.f ~ "Incidence of CHD",
                         hdlcat.f ~ "HDL cholesterol (mg/dL)")) %>%
modify_header(label = "**Variable**") %>%
  modify_spanning_header(stat_1 ~ "Family History of Heart Attack = No to CHD")
```

7 REFERENCES

1. World Health Organization. (2023, November 1). Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-%28cvds%29

2. Centers for Disease Control and Prevention. (2023, November). Heart disease facts. https://www.cdc.gov/heartdisease/index.htm

3. Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Beaton, A. Z., Bittencourt, M. S., Boehme, A. K., ... & American Heart Association. (2023). Heart disease and stroke statistics—2023 update: A report from the American Heart Association. Circulation, 147(e93–e621). https://doi.org/10.1161/CIR.0000000000001087

4. Patient education: High cholesterol and lipids (Beyond the Basics)—UpToDate. (n.d.). Retrieved December 6, 2023, from https://www.uptodate.com/contents/high-cholesterol-and-lipids-beyond-the-basics/print