**Alcohol Consumption and Depression in Older Adults- Analysis of the HRS Study**

**Introduction**

Alcohol consumption and depression are closely linked, with effects varying based on drinking patterns. Individuals with alcohol use disorders (AUD) are at higher risk for depression, and co-occurring AUD and depression often worsen outcomes for both conditions.1 Heavy drinking increases depression risk due to alcohol's impact on mood, while moderate consumption may offer protective benefits, suggesting a possible "J-shaped" relationship.2,3 However, the protective effects of moderate drinking, especially in older adults, remain poorly understood, and the risk of escalating to problematic drinking persists,2 as does the role of gender in modifying this relationship. Biological and social differences between men and women may influence how alcohol impacts mental health, potentially altering the association between alcohol use and depression.1,4 This study uses data from the Health and Retirement Study (HRS) to examine how varying levels of alcohol consumption are associated with depression in older adults, with a focus on sex differences as an effect modifier.

**Methods**

Study Population and Data Source

I examined the association between alcohol consumption and depression among older adults using cross-sectional data from Wave 10 (2010) of the Health and Retirement Study (HRS), a nationally representative cohort of U.S. adults aged 50 and older. From the original HRS cohort established in 1992 with 42,405 participants, Wave 10 included 19,887 participants. After excluding those with missing depression data (n=19), I had 19,868 participants. I further excluded participants with missing data on key covariates: age (n=0), gender (n=5), smoking history (n=12), and physical activity (n=12), resulting in my final analytical sample of 19,839 participants (Figure 1). I chose complete case analysis for both exposure and outcome variables to maintain the strongest internal validity, as missing data in these key variables could introduce substantial bias into my primary association of interest.

Variables and Operational Definitions

I assessed the exposure variable, alcohol consumption, based on self-reported drinking frequency in the past year. I converted weekly drinking frequency to estimated monthly drinking days (frequency × 4). Based on both current drinking patterns and lifetime alcohol use history, I initially categorized participants into five groups. However, after examining the distribution of drinking patterns (Figure 2) and to increase statistical power, I merged the Low-Frequency (1-5 days/month) and Moderate-Frequency (6-15 days/month) categories into a single Low/Moderate Drinker category (1-15 days/month). My final four categories were: Never Drinker (those who reported never consuming alcohol), Former Drinker (those who reported ever drinking but either reported 0 current drinking days or had missing current drinking data), Low/Moderate Drinker (current drinkers consuming alcohol 1-15 days per month), and High-Frequency Drinker (>15 days/month).

I measured depression, the primary outcome, as a binary variable, defined by participants reporting depressive symptoms lasting at least two consecutive weeks in the past year. This measure aligns with established clinical criteria for depressive episodes. Covariates included age (continuous), gender (male/female), smoking history (ever smoked: yes/no), and physical activity (any activity: yes/no). I coded physical activity as positive if participants reported any vigorous activity, moderate activity, or mild activity, and negative if they reported no activity across all three levels. I selected these covariates based on their potential confounding roles in the relationship between alcohol use and depression, as illustrated in my directed acyclic graph (Figure 3).

Data Cleaning and Statistical Analysis

For my data cleaning process, I converted character variables (mod\_activity and alc\_days) to numeric, with "NA" or blank values treated as missing. I employed complete case analysis for my primary exposure and outcome variables to ensure the strongest possible internal validity. For alcohol consumption, I classified participants with missing current drinking data but who reported previous drinking as Former Drinkers, as this represented a meaningful category rather than missing data. For other covariates, I also used complete case analysis, removing participants with missing values (n=29 total across all covariates) rather than imputation, as the proportion of missing data was small (<0.2% for any variable) and unlikely to substantially impact my results (Table 1).

I summarized baseline characteristics of participants across alcohol consumption categories using frequencies for categorical variables and means with standard deviations for continuous variables. I performed Chi-square tests to assess differences in categorical variables across alcohol consumption groups. I employed multiple logistic regression models using Fisher's scoring as the optimization technique. First, I examined the unadjusted association between alcohol use and depression, with Never Drinkers as the reference category. I then fitted a fully adjusted model incorporating all covariates to estimate adjusted odds ratios (AORs) and 95% confidence intervals.

To evaluate potential effect modification by gender, I conducted separate analyses for males (n=8,220) and females (n=11,619) while maintaining the same covariate structure (Table 3). I assessed model fit using Akaike Information Criterion (AIC), Schwarz Criterion (SC), and -2 Log Likelihood statistics. I used the Wald test to evaluate the statistical significance of individual predictors. I conducted all statistical analyses using SAS version 9.4 (SAS Institute Inc., Cary, NC), with statistical significance set at p<0.05.

**Results**

In my final analytical sample of 19,839 participants (Figure 1), I found that 42.44% were Never Drinkers, 19.29% were Former Drinkers, 26.31% were Low/Moderate Drinkers, and 11.95% were High-Frequency Drinkers (Table 1). The mean age was highest among Never Drinkers (67.80 ± 11.21 years) compared to Former Drinkers (63.70 ± 10.87 years), High-Frequency Drinkers (65.98 ± 10.90 years), and Low/Moderate Drinkers (62.32 ± 10.26 years). My sample comprised more females (58.57%) than males (41.43%). I observed significant gender distribution variations across drinking categories (p<0.001), with females more likely to be Never Drinkers (48.41% of females vs 34.01% of males) and males more likely to be Low/Moderate Drinkers (30.76% of males vs 23.17% of females). Physical activity and smoking history also showed significant variations across drinking categories (p<0.001 for both), as detailed in Table 1.

In my fully adjusted model (Table 2), I found that both Former Drinkers (AOR=0.841, 95% CI: 0.756-0.935) and Low/Moderate Drinkers (AOR=0.683, 95% CI: 0.616-0.757) had lower odds of depression compared to Never Drinkers. My gender-stratified analyses (Table 3) revealed similar patterns but with varying magnitudes. Among males, Former Drinkers (AOR=0.794, 95% CI: 0.653-0.965) and Low/Moderate Drinkers (AOR=0.595, 95% CI: 0.497-0.713) showed lower odds of depression compared to Never Drinkers. Similarly, among females, both Former Drinkers (AOR=0.861, 95% CI: 0.759-0.978) and Low/Moderate Drinkers (AOR=0.734, 95% CI: 0.647-0.832) had reduced odds of depression compared to Never Drinkers. I identified other significant predictors of depression including age (AOR=0.966, 95% CI: 0.963-0.970), male gender (AOR=0.566, 95% CI: 0.519-0.617), physical activity (AOR=0.581, 95% CI: 0.535-0.631), and smoking history (AOR=1.513, 95% CI: 1.393-1.644).

**Discussion**

Our study highlights the complex relationship between alcohol consumption and depression in older adults. We found that Low/Moderate Drinkers had lower odds of depression compared to Never Drinkers, aligning with findings from Gemes et al., who reported that long-term abstainers from alcohol were at a higher risk for depressive episodes than those who drank moderately.5 This suggests that moderate drinking might offer some protective benefits against depression in older adults. Both men and women who were Low/Moderate Drinkers showed lower odds of depression, though the effect was stronger in men. This gender difference is consistent with Choi and DiNitto's research, which found significant links between drinking patterns and depressive symptoms in older men but not as strongly in women.4 Keyes et al. also noted that moderate alcohol consumption is often linked to better mental health outcomes, emphasizing that not all drinking patterns have the same effects.6 A key strength of our study lies in its use of the nationally representative HRS dataset, which enhances generalizability and provides detailed measures of both alcohol consumption and depression. My analytical approach, using complete case analysis for primary variables and careful categorization of alcohol consumption based on observed patterns (Figure 2), helped maintain internal validity while maximizing statistical power.

However, several limitations warrant consideration when interpreting these findings. The cross-sectional design limits my ability to establish causality, and self-reported alcohol use may introduce recall bias. While my complete case analysis approach maintains internal validity, it could introduce selection bias, though this impact is likely minimal given the small proportion (<0.2%) of missing data for key covariates. Despite adjusting for key confounders, residual confounding from unmeasured variables such as genetic predispositions and socioeconomic factors may influence my results. Future longitudinal studies are needed to better understand these temporal relationships and explore the gender-specific dynamics observed in my findings, particularly the stronger protective effect of moderate alcohol consumption against depression in men compared to women, as shown in my stratified analyses (Table 3). Such research could inform more effective, tailored mental health interventions for older populations.

**Tables and Figures**

Table 1- Baseline Characteristics of the Study Sample by Alcohol Consumption Category (HRS Wave 10, 2010)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Characteristics | Never Drinker (N=8423) | Former Drinker  (N=3829) | Low/Moderate Drinker (n=5222) | High Frequency Drinker (n=2371) |
| Sample Size (n, %) | 8,423 (42.44%) | 3,829 (19.29%) | 5,222 (26.31%) | 2,371 (11.95%) |
| Age (mean, SD) | 67.80 (11.22) | 63.70 (10.87) | 62.32 (10.26) | 65.98 (10.90) |
| **Gender (n, %)** |  |  |  |  |
| Male | 2,796 (34.01%) | 1,522 (18.51%) | 2,529 (30.76%) | 1,375 (16.72%) |
| Female | 5,627 (48.41%) | 2,307 (19.85%) | 2,693 (23.17%) | 996 (8.57%) |
| **Smoking Status (n, %)** |  |  |  |  |
| Yes | 4,072 (48.34%) | 2,226 (58.14%) | 3,220 (61.66%) | 1,647 (69.46%) |
| No | 4,351 (51.66%) | 1,603 (41.86%) | 2,002 (38.34%) | 724 (30.54%) |
| **Physical Activity (n, %)** |  |  |  |  |
| Yes | 5,144 (61.07%) | 2,688 (70.20%) | 4,026 (77.10%) | 1,868 (78.79%) |
| No | 3,279 (38.93%) | 1,141 (29.80%) | 1,196 (22.90%) | 503 (21.21%) |

1. Alcohol Categories: "Never Drinker" = never drank; "Former Drinker" = drank previously but not in the last 3 months; "Low-Frequency" = 1–5 days/month; "Moderate-Frequency" = 6–15 days/month; "High-Frequency" = >15 days/month.
2. Data Source: Health and Retirement Study, Wave 10 (2010).
3. Percentages: Percentages are calculated within each alcohol category.
4. Abbreviations: N, sample size; %, proportion in percentages; SD, standard

deviation.

Table 2: Adjusted Odds Ratios for Depression by Alcohol Consumption Category (HRS Wave 10, 2010)

|  |  |
| --- | --- |
| Alcohol Consumption Categories | Adjusted\* Odds Ratio (95% CI) |
| Never Drinker | 1.0 |
| Former Drinker | 0.841 (0.756-0.935) |
| Low/Moderate Drinker | 0.683 (0.616-0.757) |
| High Frequency Drinker | 0.719 (0.625-0.828) |

\* Adjusted for age, gender, smoking status, and physical activity.

1. Reference Group: "Never Drinker" is the reference category for all odds ratios.

2. Adjusted OR: Adjusted for age, gender, smoking history, and physical activity to control for confounding.

3. Data Source: Health and Retirement Study, Wave 10 (2010)

4. Abbreviations- 95% CI, 95% Confidence Intervals

Table 3: Stratified Analysis of Adjusted Odds Ratios for Depression by Alcohol Consumption Category and Gender (HRS Wave 10, 2010)

|  |  |  |
| --- | --- | --- |
| Alcohol Consumption Categories | Male (95% CI)  (N=8222) | Female (95% CI)  (N=11623) |
| Never Drinker | 1.0 | 1.0 |
| Former Drinker | 0.794 (0.653-0.965) | 0.861 (0.759-0.978) |
| Low/ Moderate Drinker | 0.595 (0.497-0.713) | 0.734 (0.647-0.832) |
| High Frequency Drinker | 0.702 (0.567-0.869) | 0.714 (0.591-0.863) |

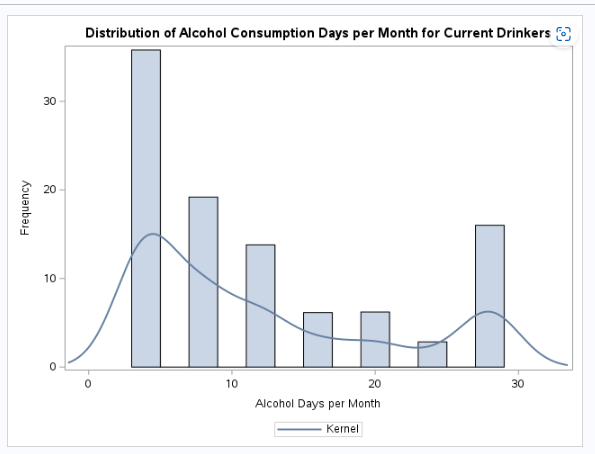
1. Reference Group: "Never Drinker" is the reference category for all odds ratios.
2. Adjusted OR: Adjusted for age, gender, smoking history, and physical activity to control for confounding.
3. Data Source: Health and Retirement Study, Wave 10 (2010)
4. Abbreviations- 95% CI, 95% Confidence Intervals

Figure 1: Flowchart of Sample Selection and Alcohol Consumption Categories (HRS Wave 10, 2010)



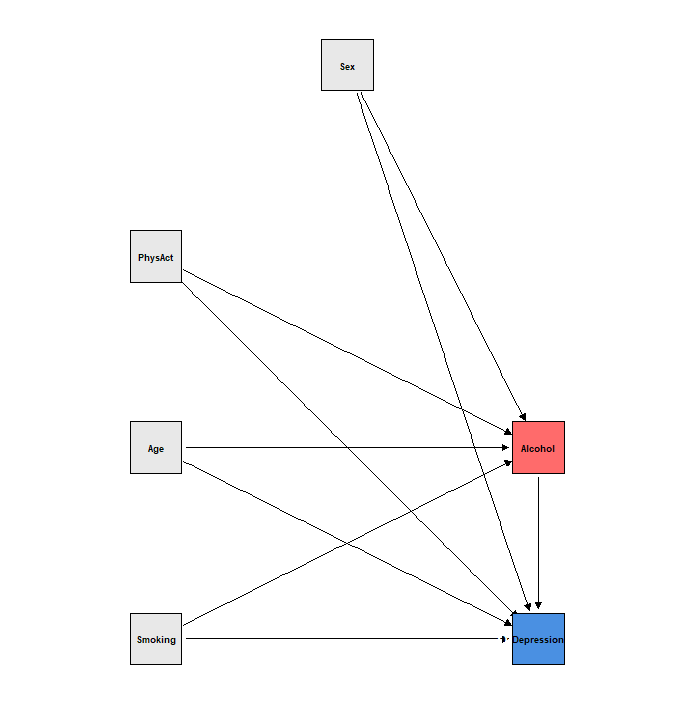
Description: This flowchart illustrates the steps taken to arrive at the final analytical sample. Participants were categorized into five alcohol consumption groups: Never Drinker, Former Drinker, Low/Moderate Frequency Drinker and High-Frequency Drinker.

Figure 2: Distribution of Alcohol Consumption Days per Month for Current Drinkers



Description: This histogram shows the distribution of estimated drinking days per month for participants who reported current alcohol use. A kernel density curve is overlaid to illustrate the distribution pattern, highlighting distinct peaks that informed the categorization of alcohol consumption frequency.

Figure 3: Directed Acyclic Graph (DAG) of the Relationship Between Alcohol Consumption and Depression



Description: This DAG illustrates the hypothesized relationships between alcohol consumption (exposure) and depression (outcome), including covariates and potential confounders. Variables include age, smoking status (Smoking), physical activity (PhysAct), and gender (Sex). Arrows indicate assumed causal pathways, with gender (Sex) considered both a confounder and an effect modifier.

*Abbreviations*: DAG = Directed Acyclic Graph, PhysAct = Physical Activity.

SAS/R Code-

1. **R Code for DAG**

library(ggplot2)

library(dagitty)

# Create the DAG with specified layout

dag <- dagitty("dag {

Alcohol [pos=\"2,0\"]

Depress [pos=\"2,-1\"]

Age [pos=\"0,0\"]

Sex [pos=\"1,2\"]

PhysAct [pos=\"0,1\"]

Smoking [pos=\"0,-1\"]

Alcohol -> Depress

Smoking -> Depress

Smoking -> Alcohol

PhysAct -> Depress

PhysAct -> Alcohol

Age -> Depress

Age -> Alcohol

Sex -> Depress

Sex -> Alcohol

}")

# Convert the DAG into a tidy format

tidy\_dag <- tidy\_dagitty(dag)

# Create a customized plot with larger square nodes

ggdag(tidy\_dag) +

theme\_dag() +

geom\_dag\_point(aes(fill = name),

size = 22.5, # Increased size from 12 to 20

shape = 22, # Square shape

stroke = 0.5) +

geom\_dag\_text(size = 3.5, # Slightly reduced text size for better fit

color = "black") +

scale\_fill\_manual(values = c(

"Depress" = "#4A90E2", # Blue for outcome

"Alcohol" = "#FF6B6B", # Coral for exposure

"Sex" = "#E8E8E8", # Light grey for covariates

"Age" = "#E8E8E8",

"PhysAct" = "#E8E8E8",

"Smoking" = "#E8E8E8"

)) +

theme(legend.position = "none") +

coord\_fixed(ratio = 1)

1. **R Code for Flowchart**

# Install and load required packages if not already installed

# install.packages("DiagrammeR")

library(DiagrammeR)

flowchart <- grViz("

digraph flowchart {

# Node attributes

node [shape = rectangle, fontname = Helvetica, fontsize = 10,

style = filled, fillcolor = lightblue, width = 4]

# Define nodes

orig [label = 'Original HRS Cohort (1992): 42,405']

wave [label = 'Wave 10 (2010) Cross-Sectional Sample: 19,887']

miss [label = 'Removed Missing Depression Values: 19,868']

cat [label = 'Categorized Alcohol Consumption: 19,845']

# Define category nodes

never [label = 'Never Drinker:\n8,423 (42.44%)']

former [label = 'Former Drinker:\n3,829 (19.29%)']

lowmod [label = 'Low/Moderate Drinker:\n5,222 (26.31%)']

high [label = 'High-Frequency Drinker:\n2,371 (11.95%)']

# Define edges

orig -> wave

wave -> miss

miss -> cat

cat -> never

cat -> former

cat -> lowmod

cat -> high

# Graph attributes

graph [rankdir = TB]

# Rank specifications

{rank = same; never former lowmod high}

}

")

flowchart

1. **SAS Code**

/\* ----------------------------------------------------------------- \*/

/\* Library Definition and Data Loading Section \*/

/\* ----------------------------------------------------------------- \*/

/\* Defined the library reference for accessing the data files \*/

libname lib '/home/u63883323/sasuser.v94';

/\* Loaded the Health and Retirement Study (HRS) dataset into a working dataset \*/

data project;

set lib.hrs\_data;

run;

/\* ----------------------------------------------------------------- \*/

/\* Initial Data Quality Assessment and Exploration \*/

/\* ----------------------------------------------------------------- \*/

/\* Examined the dataset structure including variables, types, and formats \*/

proc contents data=project;

title "Structure of Original Dataset";

run;

/\* Analyzed the initial distributions of key variables and identified missing patterns

for depression scores and alcohol consumption metrics \*/

proc freq data=project;

tables depress ever\_alc alc\_days / missing;

title "Initial Distributions and Missing Patterns";

run;

/\* Performed quality check by cross-tabulating alcohol consumption variables

to verify data consistency and identify potential discrepancies \*/

proc freq data=project;

tables alc\_days \* ever\_alc / list missing;

title "Cross-tabulation of Alcohol Use Variables";

run;

/\* ----------------------------------------------------------------- \*/

/\* Data Cleaning and Variable Transformation \*/

/\* ----------------------------------------------------------------- \*/

/\* Step 1: Removed records with missing depression scores to ensure

complete outcome data for primary analysis \*/

data project\_1;

set project;

if depress ^= .;

run;

/\* Step 2: Performed alcohol consumption variable transformations:

- Converted character alcohol days to numeric format

- Calculated monthly drinking frequency

- Handled NA values and missing data appropriately \*/

data project\_2;

set project\_1;

/\* Converted character alcohol days to numeric, properly handling NA values \*/

if alc\_days not in ('NA', '') then alc\_days\_num = input(alc\_days, 8.);

else alc\_days\_num = .;

/\* Calculated standardized monthly drinking frequency for analysis \*/

if alc\_days\_num ne . then alc\_days\_per\_month = alc\_days\_num \* 4;

run;

/\* Verified the accuracy of numeric conversions and monthly calculations \*/

proc freq data=project\_2;

tables alc\_days \* alc\_days\_num

alc\_days \* alc\_days\_per\_month / list missing;

title "Verification of Alcohol Variable Conversions";

run;

/\* ----------------------------------------------------------------- \*/

/\* Figure 2: Alcohol Consumption Distribution Analysis \*/

/\* ----------------------------------------------------------------- \*/

/\* Created a subset of current drinkers for distribution analysis \*/

data current\_drinkers;

set project\_2;

if alc\_days\_per\_month > 0;

run;

/\* Generated visualization of alcohol consumption patterns:

- Created histogram with 2-day bin widths

- Overlaid kernel density estimate

- Set appropriate axis limits and labels \*/

ods graphics on;

proc sgplot data=current\_drinkers;

histogram alc\_days\_per\_month / binwidth=2;

density alc\_days\_per\_month / type=kernel;

xaxis label="Alcohol Days per Month" min=0 max=30;

yaxis label="Frequency";

title "Distribution of Alcohol Consumption Days per Month for Current Drinkers";

run;

/\* Calculated detailed distribution statistics for alcohol consumption \*/

proc univariate data=current\_drinkers;

var alc\_days\_per\_month;

histogram alc\_days\_per\_month / endpoints=0 to 30 by 2;

title "Summary Statistics for Alcohol Consumption Days";

run;

ods graphics off;

/\* Created comprehensive variable categorizations and transformations:

- Developed alcohol status categories

- Processed physical activity data

- Standardized gender coding \*/

data project\_3;

set project\_2;

/\* Created alcohol consumption categories based on frequency and history \*/

length alcohol\_status $25;

if ever\_alc = 0 then alcohol\_status = 'Never Drinker';

else if ever\_alc = 1 then do;

if alc\_days\_num = 0 or alc\_days = 'NA' then alcohol\_status = 'Former Drinker';

else if 1 <= alc\_days\_per\_month <= 15 then alcohol\_status = 'Low/Moderate Drinker';

else if alc\_days\_per\_month > 15 then alcohol\_status = 'High-Frequency Drinker';

else alcohol\_status = 'Former Drinker';

end;

/\* Derived physical activity status from vigorous and moderate activity data \*/

if vig\_activity in (1,2,7) then phys\_activity = 1;

else if mod\_activity not in ('NA', '') then do;

mod\_activity\_num = input(mod\_activity, 8.);

if mod\_activity\_num in (1,2,7) then phys\_activity = 1;

else if mod\_activity\_num in (3,4) and vig\_activity in (3,4) then phys\_activity = 0;

else phys\_activity = .;

end;

else if vig\_activity in (3,4) then phys\_activity = 0;

else phys\_activity = .;

/\* Standardized gender coding for analysis \*/

if sex = 1 then gender = 'Female';

else if sex = 2 then gender = 'Male';

else gender = '';

format alcohol\_status $25. gender $10.;

run;

/\* Validated physical activity categorization accuracy \*/

proc freq data=project\_3;

tables mod\_activity\*phys\_activity / missing;

title "Check Physical Activity Categorization";

run;

/\* ----------------------------------------------------------------- \*/

/\* Final Analysis Dataset Preparation \*/

/\* ----------------------------------------------------------------- \*/

/\* Created complete case analysis dataset by removing missing values

on all key variables required for the analysis \*/

data project\_analysis;

set project\_3;

if age = . then delete;

if gender = '' then delete;

if ever\_smoke = . then delete;

if phys\_activity = . then delete;

if alcohol\_status = '' then delete;

run;

/\* ----------------------------------------------------------------- \*/

/\* Statistical Analysis and Results Generation \*/

/\* ----------------------------------------------------------------- \*/

/\* Generated Table 1: Analyzed baseline characteristics across alcohol consumption groups \*/

proc freq data=project\_analysis;

tables alcohol\_status\*(gender ever\_smoke phys\_activity) / chisq;

title "Table 1: Baseline Characteristics by Alcohol Consumption";

run;

/\* Calculated age distribution statistics by alcohol consumption category \*/

proc means data=project\_analysis n mean std;

class alcohol\_status;

var age;

title "Age Distribution by Alcohol Consumption";

run;

/\* Performed adjusted logistic regression analysis for depression risk \*/

proc logistic data=project\_analysis;

class alcohol\_status(ref='Never Drinker') gender(ref='Female') / param=ref;

model depress(event='1') = alcohol\_status age gender phys\_activity ever\_smoke;

title "Table 2: Adjusted Analysis of Depression Risk";

run;

/\* Conducted gender-stratified analyses to examine effect modification \*/

proc logistic data=project\_analysis;

where gender = 'Male';

class alcohol\_status(ref='Never Drinker') / param=ref;

model depress(event='1') = alcohol\_status age phys\_activity ever\_smoke;

title "Table 3: Males - Stratified Analysis";

run;

proc logistic data=project\_analysis;

where gender = 'Female';

class alcohol\_status(ref='Never Drinker') / param=ref;

model depress(event='1') = alcohol\_status age phys\_activity ever\_smoke;

title "Table 3: Females - Stratified Analysis";

run;

/\* Verified final sample size and distribution \*/

proc freq data=project\_analysis;

tables gender\*alcohol\_status / nocol nopercent;

title "Final Analysis Sample Distribution";

run;

/\* Cleared all titles \*/

title;

**References**

1. McHugh RK, Weiss RD. Alcohol Use Disorder and Depressive Disorders. *Alcohol Res Curr Rev*. 2019;40(1):arcr.v40.1.01. doi:10.35946/arcr.v40.1.01

2. Nunes EV. Alcohol and the Etiology of Depression. *Am J Psychiatry*. 2023;180(3):179-181. doi:10.1176/appi.ajp.20230004

3. Li J, Wang H, Li M, et al. Effect of alcohol use disorders and alcohol intake on the risk of subsequent depressive symptoms: a systematic review and meta-analysis of cohort studies. *Addiction*. 2020;115(7):1224-1243. doi:10.1111/add.14935

4. Choi NG, Dinitto DM. Heavy/binge drinking and depressive symptoms in older adults: gender differences. *Int J Geriatr Psychiatry*. 2011;26(8):860-868. doi:10.1002/gps.2616

5. Keyes KM, Allel K, Staudinger UM, Ornstein KA, Calvo E. Alcohol consumption predicts incidence of depressive episodes across 10 years among older adults in 19 countries. *Int Rev Neurobiol*. 2019;148:1-38. doi:10.1016/bs.irn.2019.09.001

6. Gémes K, Forsell Y, Janszky I, et al. Moderate alcohol consumption and depression – a longitudinal population-based study in Sweden. *Acta Psychiatr Scand*. 2019;139(6):526-535. doi:10.1111/acps.13034