

Alcohol-Related Disease Impact (ARDI) Custom Data User Manual and Sample Code



March 2022

**Alcohol Program
Division of Population Health
National Center for Chronic Disease Prevention and Health Promotion
Centers for Disease Control and Prevention**

Contents

Introduction.....	3
Prevalence of Alcohol Consumption	4
Relative Risks and Direct Alcohol-Attributable Fractions	7
Total Deaths	9
Viewing Results in the Custom Data Portal	13
Generating Reports	13
Saving Reports and Changing Report View	13
Appendix A. Supporting Documents	15
1. ARDI ICD Codes and Alcohol-Attributable Fraction Sources.....	15
2. Cause of Death Abbreviations and Age Groups	15
3. Locations for Custom Data File	17
Appendix B. Sample SAS Codes.....	18
1. Age Group Codes.....	18
2. Age Group Conversion	19
3. Total Deaths by State	20
4. Prevalence of Average Daily Alcohol Consumption Calculation and Sample SAS code.	35

Introduction

The Alcohol-Related Disease Impact ([ARDI](#)) application is an online tool developed by the Centers for Disease Control and Prevention's (CDC) Alcohol Program to estimate alcohol-related harms including alcohol-attributable deaths (AAD), years of potential life lost (YPLL), and alcohol-attributable fractions (AAF). Estimates of AAD and YPLL are available for the United States overall and individually for all 50 states and the District of Columbia. Currently, there are 58 causes of death included in the ARDI estimates, and results are available by sex and age groups.

In addition to the public use site of ARDI, ARDI allows users to create customized estimates of AAD and YPLL with the appropriate data sources. Custom data reports could include estimates of alcohol-related harms for different geographic locations, such as populous counties or metropolitan areas, or different years than those provided in the public use site.

The Q&A that follows will help you understand the purpose of this manual and see how it will help you create custom data sets in the ARDI application.

Q&A

What is the purpose of this user manual?

You'll use this manual to create custom data sets for the ARDI application and to know how to use the Custom Data Portal in the ARDI application.

What are two examples of custom data analyses I can perform by following the steps in this manual?

1. Average annual number of alcohol-attributable deaths and years of potential life lost for geographic locations other than the state-level data provided in the public use site, such as analyses for populous counties or metropolitan areas.
2. Years different than those provided in the public use site.

Will I have access to sample SAS codes to use in creating my custom data analyses?

Yes. You'll find sample SAS codes in [Appendix B](#) of this manual.

What customized data sets can I learn about in this manual?

- Prevalence of alcohol consumption (page 4).
- Relative risks and direct alcohol-attributable fractions (page 7).
- Total deaths (page 9).

When I'm doing a custom data analysis, can I use some of the default data sets available in ARDI?

Yes. For example, you might use the default relative risks and direct alcohol-attributable fractions data set with your customized prevalence of alcohol consumption and total deaths data sets.

How can I get technical assistance with my custom data analyses?

Click on “[Contact Us About ARDI](#)” on any of the ARDI web pages. Please be as specific as possible with your question with sufficient detail so that someone can assist you.

Prevalence of Alcohol Consumption

In this section you’ll find instructions on preparing a prevalence of alcohol consumption data set and the sample code you need.

The nine steps show how to prepare your custom data set on the prevalence of alcohol consumption.

1. Familiarize yourself with the methodology for calculating the prevalence of average daily alcohol consumption in the ARDI application.

ARDI uses two core BRFSS alcohol questions in the calculation of average daily alcohol consumption in ARDI: 1) the frequency of alcohol use in the past 30 days and 2) the average quantity of alcohol consumed on days a person reports drinking.

The ARDI application adjusts the prevalence of low, medium, and high average daily alcohol consumption to account for the underreporting of self-reported alcohol consumption in [BRFSS](#) using per capita alcohol sales data. Individual-level average daily alcohol consumption estimates from the BRFSS are adjusted using a correction factor so that overall average daily consumption reaches 73% of national per capita sales. This method is explained by [Esser et al. \(2022\)](#).

After applying the per capita alcohol sales adjustment for males and females, average daily alcohol consumption is compared to cut points (for low, medium, high) specified by the meta-analyses used to obtain risk estimates for a given cause of death. Prevalence estimates are calculated for the low, medium, high consumption levels. The cut points for defining these levels are described in the ARDI [Methods](#).

ARDI generally uses the prevalence of US adults aged 20 years and older (stratified by sex) who reported average daily alcohol consumption at various levels, but there are exceptions.

If you want to create your own prevalence data set, your first step is to review the sample SAS code provided (see [Appendix B](#)).

2. Acquire Behavioral Risk Factor Surveillance System (BRFSS) data files.

- a. Go to the [BRFSS](#) web page. The BRFSS is a state-based system that collects information on a variety of health risk behaviors, including alcohol use. Additional information on the BRFSS is available [online](#).
- b. Understand that some of the BRFSS variable names may differ by year, depending on the years of data used.

- c. Determine the jurisdiction and years for which you would like to calculate the prevalence of alcohol consumption.
- Jurisdiction/Location: If you are estimating alcohol-attributable deaths below the state level, use BRFSS data for the corresponding location, if possible.

If this is not possible, use state-level estimates to estimate local alcohol consumption. Note that local prevalence estimates may be different than those reported for the state.

- Years: Use multiple years of data to calculate the average daily alcohol consumption. Doing so will help to account for any year-to-year fluctuations in the data that may not be related to actual changes in alcohol consumption.

If state-level data are not available in the BRFSS for a certain year (e.g., New Jersey in 2019), sample SAS code is provided to address this issue (see [Appendix B](#)).

3. Use the sample SAS code for calculating the prevalence of average daily alcohol consumption using the methodology in ARDI (see [Appendix B](#)).

Two tips for understanding your prevalence data set using the sample SAS code:

- When preparing your data set, even if there are zero records for people aged 85 and older, keep the age range as 80–85 rather than changing it to 80–84.
- There are steps to calculate the prevalence of average daily alcohol consumption among women aged 18–44, which is used in the calculations of alcohol-attributable infant deaths. After calculating the prevalence estimates among women aged 18–44, the SAS program is designed to apply those same estimates in the output for men, which is necessary to calculate deaths among infants who are boys.

4. Put your data files in CSV format. Each row must contain the following columns of data (see Table 1):

- a. Column A: Location (full location name).
- b. Column B: Year (four digits).
- c. Column C: Study (e.g., English).
- d. Column D: Male low prevalence (percentage expressed as a fraction: 0.50 for 50%).
- e. Column E: Male medium prevalence.
- f. Column F: Male high prevalence.
- g. Column G: Female low prevalence.
- h. Column H: Female medium prevalence.
- i. Column I: Female high prevalence.

Table 1. Example of format for prevalence data set

	A	B	C	D	E	F	G	H	I
2	UNITED STATES	2019	English	0.37	0.05	0.1	0.24	0.07	0.05
3	UNITED STATES	2019	Standardiz	0.4	0.1	0.1	0.32	0.08	0.08
4	UNITED STATES	2019	1844	0.35	0.09	0.08	0.35	0.09	0.08
5	UNITED STATES	2019	2064	0.42	0.1	0.1	0.34	0.09	0.08
6	UNITED STATES	2019	Zhao	0.32	0.1	0.13	0.31	0.06	0.04

5. **Remove all column headers from the data set.** You'll receive an error message if you haven't removed them.
6. **Remove all blank rows at the bottom of the data set in CSV file.** A row without data will interfere with the upload. You'll receive an error message if you haven't removed them.
7. **You have two options to input your data set** (see Figure 1).
 - a. Click on "Upload Your Own Prevalence Data Set" and browse to upload the correct file.
 - b. Click on "Create Your Own Prevalence Data Set"
 - Carefully follow the instructions on the screen to enter prevalence data into the ARDI application.
 - Name your data set. It combines the state and the year of your data set (for example, "Georgia 2022").
 - Click save.

Figure 1. Two options to input your prevalence data set

Select Data Sets

Note: It is suggested to select data sets with consistent locations, such as from the same state or the state where the large city is located if you have a sub-state data set. The title of the custom reports will be generated based on the location of the total deaths data set.

Prevalence of Alcohol Consumption

View Data Set

[Create Your Own Prevalence Data Set](#)
[Upload Your Own Prevalence Data Set](#)

8. **Select your customized data set for your analysis.** It will appear at the bottom of the prevalence data sets drop-down menu.
9. **Delete files you no longer need.** If you upload multiple versions of the same data set, it is challenging to differentiate files that use the same state name; no date/time stamp appears on the uploaded file. If you find that you have multiple data sets with the same name, check carefully that you are analyzing the desired one before proceeding.

Having trouble? Click on "[Contact Us About ARDI](#)" on any of the ARDI web pages. Please be as specific as possible with your question. For example, instead of saying, "I have a question about how to calculate the prevalence of average daily alcohol consumption," state your question (for example, "I don't understand why there are different sets of prevalence estimates for different causes of death, could you explain this to me?").

Relative Risks and Direct Alcohol-Attributable Fractions

In this section you'll find instructions on preparing a relative risks and direct alcohol-attributable fractions (AAF) data set and the sample code you need.

The six steps for acquiring and entering direct AAFs into the ARDI application are:

- 1. Familiarize yourself with the methodology for using AAF in ARDI calculations.**
ARDI uses AAF to calculate the proportion of deaths and years of potential life lost that are due to alcohol use. ARDI uses either direct estimates of AAF or calculates AAF using pooled relative risk estimates and prevalence of alcohol consumption data. Read about the different types of AAF in the ARDI [Methods](#).
- 2. Click on “Create Your Own Acute Cause AAF Data Set”** beneath the Relative Risks and Direct AAF data set to see the causes of death that have customizable AAF.
- 3. Acquire location specific AAF data for acute causes of death, if data are available.** In general, you can enter your own directly measured AAF for acute causes of death. There are only four chronic causes with customizable AAF (see Figure 2). You cannot enter your own relative risks or AAF for deaths due to most of the deaths from chronic causes; instead, the default relative risks are used for these. If you are looking at a particular state or county, you can use AAF collected from death data in that state or county. If you do not have access to location specific AAF, you can use the default AAF that are pre-populated on the custom portal.

Figure 2. Example of some causes of death with customizable alcohol-attributable fractions

Chronic Causes				
		Age Group	Males	Females
Esophageal varices	Enter age specific	All	0.680	0.680
Gastroesophageal hemorrhage	Enter age specific	All	0.470	0.470
Liver cirrhosis, unspecified	Enter age specific	All	0.400	0.400
Portal hypertension	Enter age specific	All	0.680	0.680
Acute Causes				
		Age Group	Males	Females
Air-space transport	Enter age specific	All	0.180	0.180
Aspiration	Enter age specific	All	0.180	0.180
Child maltreatment	Enter age specific	All	0.160	0.160
Drowning	Enter age specific	All	0.340	0.340
Fall injuries	Enter age specific	All	0.320	0.320

- 4. When you are ready to create your customized data set, click on “Create Your Own Acute Cause AAF Data Set”** beneath the Relative Risks and Direct AAF data set.
- 5. Enter the name of the data set, location, and year to represent the AAF estimates.**
 - Type the AAF estimates for both males and females for all ages for each cause of death that you have data to customize.

- b. For motor vehicle traffic crashes, enter age specific AAF for males and females (Figure 3).

Figure 3. Example of customizable alcohol-attributable fractions for motor vehicle traffic crash deaths

Motor-vehicle traffic crashes	Enter all ages	0-14	0.190	0.210
		15-19	0.260	0.240
		20-24	0.430	0.380
		25-34	0.460	0.400
		35-44	0.440	0.360
		45-54	0.410	0.330
		55-64	0.340	0.250
		65+	0.170	0.130

6. Click “Save” at the bottom of the page.

Having trouble? Click on “[Contact Us About ARDI](#)” on any of the ARDI web pages. Please be as specific as possible with your question. For example, instead of saying, “I have a question about entering my AAF,” state your question (for example, “Do I need to enter age specific AAF for conditions other than motor vehicle traffic crashes?”).

Total Deaths

In this section you'll find steps for generating a total deaths data set in the ARDI application.

The 11 steps for generating a total deaths data set are:

1. **Familiarize yourself with the requirements for preparing a total deaths data set for ARDI.** ARDI is primarily used to calculate the average number of AAD and YPLL over more than one year. Calculating annual average estimates of AAD and YPLL helps to smooth out year-to-year fluctuations in these estimates due to extraneous factors that are unrelated to alcohol, which helps to improve the interpretation of ARDI estimates. Do not use ARDI to calculate AAD and YPLL over less than 12 months.
2. **Acquire state or local death data** by contacting the government agencies that manage vital statistics in your state, or perhaps by arranging a data use agreement with National Center for Health Statistics (NCHS).
 - About death data: The state data from death certificates are shared with NCHS. With your data, remove any deaths where the location of the death does not correspond to the location you want to assess in ARDI, or where the year of death does not correspond to the period you want to assess.
 - About death data from locations outside the United States: ARDI was primarily designed to assess alcohol-attributable deaths and YPLL in the United States. You can use the ARDI application to estimate AAD and YPLL for locations outside of the United States, but you'll need to obtain the data required (such as alcohol consumption prevalence estimates or death data) from other sources. When you name your data file, adhere to the formats and naming conventions within the ARDI application which are based on locations within the United States.
3. **Use the sample SAS code provided for calculating the total deaths from alcohol-related causes** using the methodology in ARDI (see [Appendix B](#)). In this SAS program, you will:
 - **Create a population level death data set for the location specified** (in other words, you won't upload individual death data).
 - **Use International Classification of Disease Tenth Edition ([ICD-10](#)) codes to define the 58 causes of death** included in ARDI. Include only those deaths with an underlying cause that corresponds to one of the 58 conditions specified in ARDI application. You'll receive errors if you enter data on causes of death other than the 58 conditions that are currently included in the application.
 - **Base the data on the *underlying* cause of death only.** On some death certificates, an underlying (or primary) cause of death is listed in addition to contributing causes of death. The AAF in ARDI are based on population estimates of the contribution of alcohol to the underlying cause of death. Therefore, ARDI only includes those deaths where the underlying cause of

death corresponds to one of the 58 alcohol-related conditions included in the application.

- **Identify the alcohol-related causes of death using a six-letter code** (see [Appendix A](#)).
 - **Stratify the death data by sex and 5-year age groups.** Generally chronic causes of death are collected for people aged 20 years and older and acute causes of death for people aged 15 years and older. However, there are several exceptions to this rule, such as infant deaths and motor vehicle traffic crash deaths (see [Appendix A](#)).
4. **Check your death data set** to ensure that the location where the death took place and the year of death correspond to the location and period you want to assess using ARDI. For example, if you want to estimate the AAD for Atlanta, Georgia, for 2015–2019, the death data should be for Atlanta, Georgia, for 2015–2019.
 5. **Name your CSV file using three elements.** The name should have three elements and use an underscore (“_”) as the separator between parts (for example, 2015-2019_Georgia_2022.csv):
 - a. Range of years the data set is covering (2015–2019)
 - b. State for which the data set covers (Georgia) (see [Appendix A](#))
 - c. Current year the data set is uploaded (2022)
 6. **Assign your total deaths data set file a name and a location.**
 - The data set name can be anything you specify.
 - The location name must be either the United States, one of the 50 states, or the District of Columbia (see [Appendix A](#)). If you have more than one file with the same year and location, the file names will be sequentially numbered when they appear in your drop-down menu. The file names will be appended with “_Custom_#”.
 7. **Put your data files in CSV format** (see Table 2); each row must contain the following columns of data:
 - a. Column A: Location ([full location name](#)).
 - b. Column B: Year (4 digits).
 - c. Column C: [Six-letter code](#) for the cause of death abbreviations.
 - d. Column D: [Age group code](#).
 - e. Column E: Number of male deaths.
 - f. Column F: Number female deaths.

Table 2. Examples of total death data in CSV format

Example 1. Drowning

MISSISSIPPI,2019,DROWN,15-19,231,123

MISSISSIPPI,2019,DROWN,20-24,314,101

Where:

- MISSISSIPPI is the location.
- 2019 is the final year of data included.
- DROWN is the cause of death abbreviation for drowning.
- 15-19 and 20-24 are the age groups.
- 231 and 314 are the number of male deaths.
- 123 and 101 are the number of female deaths.

Example 2. Prostate cancer

UTAH,2021,CANPRO,70-75,89,0

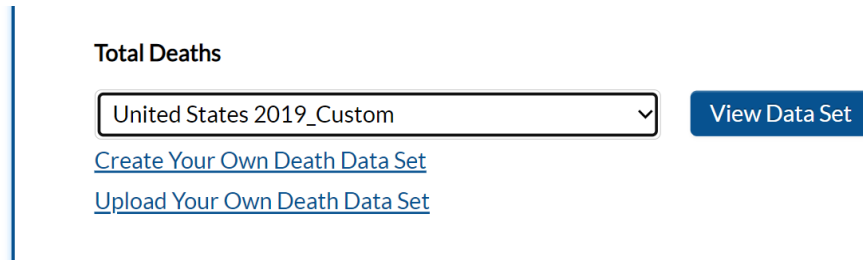
Where:

- Utah is the location.
- 2021 is the final year of data included.
- CANPRO is the cause of death abbreviation for prostate cancer.
- 70-75 is the age group.
- 89 is the number of male deaths.
- 0 is the number of female deaths.

- 8. Do not leave any columns blank.** If a cause of death is sex-specific (e.g., breast cancer, prostate cancer), put zeros in the column that pertains to the sex that is not affected by the cause of death. Male breast cancer is not included due to limited available science at the time of ARDI development.
- 9. Remove all column headers from the data set.** You'll receive an error message if you haven't removed them.
- 10. Check that the age groups are formatted correctly.** If the file is opened in Excel, you might need to change the format of the age groups if date formatting occurs automatically. For example, the age group of "1-4" might become Jan-4. To correct the formatting of the age groups:
 - Save the file in the CSV format.
 - Then, open the file in a text-based program such as Notepad.
 - Scan for the above conditions and manually change the dates to the age ranges.
 - Save it as a CSV file. After completing this, **do not** reopen the file in Excel; if you do, the age ranges might default back to date formatting.
- 11. You have two options to input your data set** (see Figure 4).
 - Click on "Upload Your Own Death Data Set" and browse to upload the correct file.
 - This is the preferred method to avoid data entry errors.

- b. Click on “Create Your Own Death Data Set”
- Enter the following information in the “Data Set Properties” box:
 - Name of the data set file, which can contain multiple years (such as, Name = My Data 2015–2019). You can customize the name of the data set in this box only.
 - Choose the appropriate location.
 - A single year to represent the data (such as the final year of data included).

Figure 4. Two options to input your total deaths data set



Total Deaths

United States 2019_Custom ▼

[Create Your Own Death Data Set](#)

[Upload Your Own Death Data Set](#)

View Data Set

Having trouble? Click on “[Contact Us About ARDI](#)” on any of the ARDI web pages. Please be as specific as possible with your question. For example, instead of saying, “I have a question about how to create a total deaths data set,” state your question (for example, “Can I include contributing causes of death in my total deaths data set?”).

Viewing Results in the Custom Data Portal

Generating Reports

In this section you'll find steps for using custom data sets to generate reports on alcohol-attributable deaths or years of potential life lost.

The six steps show you how to use custom data sets to these generate reports.

1. **Go to the Custom Data Portal homepage.**
2. **See your data set(s)** in the drop-down menus, with their file names appended with “_Custom”.
3. **Select the data sets on the Prevalence of Alcohol Consumption data set, AAF, and Total Deaths for consistent locations.**
4. **Select to view either the AAD or YPLL report** (see Figure 5). The YPLL report uses data from the default life expectancy data programmed in ARDI, as explained in the ARDI [Methods](#). The life expectancy data set cannot be modified in the custom analyses.

Figure 5. Types of reports to generate using your custom data



5. **Check that you are viewing results for the desired data sets** by scrolling to the bottom of the report and seeing the “Data Sets Used to Generate this Report.”
6. **Review your data.** As you review, there might be suppressed estimates. Data are suppressed in cells of state reports with an estimate of fewer than 10 deaths or in which presenting data would provide information to derive the estimate for another cell that has fewer than 10 deaths. Category subtotals might still be available even if estimates are suppressed for specific causes of death.

Saving Reports and Changing Report View

In this section you'll find instructions on viewing, downloading, saving, and printing reports and on changing the report view.

The three steps show you how to view, download, save, and print reports and to change the report view.

1. **Go to the Alcohol-Attributable Deaths report or the Years of Potential Life Lost report.**
2. **Click to download, save, or print the results.**

3. As desired, **change the report view** to show alcohol-attributable deaths for only those under age 21. You may also change the report view to show alcohol-attributable deaths for any alcohol use (i.e., all levels of alcohol consumption). Finally, you can also change the report view to show alcohol-attributable deaths by age groups. (Note: Due to rounding, the death report by age group may differ from the total deaths report.)

Appendix A. Supporting Documents

1. [ARDI ICD Codes and Alcohol-Attributable Fraction Sources](#)

2. Cause of Death Abbreviations and Age Groups

Use the following abbreviations for cause of death and appropriate age groupings for setting up custom data.

Cause of Death	Cause of Death Abbreviation	Age Range (using 5-year age groupings)
Chronic causes		
100% alcohol-attributable		
Alcoholic psychosis	ALCPSY	20–85+
Alcohol abuse	ALCABU	20–85+
Alcohol dependence syndrome	ALCDSY	20–85+
Alcohol polyneuropathy	ALCPOL	20–85+
Degeneration of nervous system due to alcohol	ALCNER	20–85+
Alcoholic myopathy	ALCMYO	20–85+
Alcohol cardiomyopathy	ALCCAR	20–85+
Alcoholic gastritis	ALCGAS	20–85+
Alcoholic liver disease	ALCLID	20–85+
Fetal alcohol syndrome	ALCFAS	<1, 1–4, 5–9, 10–14, 15–19, 20–85+
Fetus and newborn affected by maternal use of alcohol	ALCMAU	<1, 1–4
Alcohol-induced acute pancreatitis	ALCACP	20–85+
Alcohol-induced chronic pancreatitis	ALCCRP	20–85+
Direct alcohol-attributable fraction estimates		
Esophageal varices	ESOVAR	20–85+
Gastroesophageal hemorrhage	GASHEM	20–85+
Liver cirrhosis, unspecified	LIVCIR	20–85+
Portal hypertension	PORHYP	20–85+
Indirect alcohol-attributable fraction estimates		
Atrial fibrillation	ATRFIB	20–85+
Cancer, breast (female only)	CANBRE	20–85+
Cancer, colorectal	CANCOL	20–85+
Cancer, esophageal	CANESO	20–85+
Cancer, laryngeal	CANLAR	20–85+

Cause of Death	Cause of Death Abbreviation	Age Range (using 5-year age groupings)
Cancer, liver	CANLIV	20–85+
Cancer, oral cavity and pharyngeal	CANOCP	20–85+
Cancer, pancreatic	CANPAN	20–85+
Cancer, prostate (male only)	CANPRO	20–85+
Cancer, stomach	CANSTO	20–85+
Chronic hepatitis	CHRHEP	20–85+
Gallbladder	GALLBL	20–85+
Hypertension	HYPERT	20–85+
Infant death, low birth weight	INFLBW	<1, 1–4
Infant death, preterm birth	INFPRE	<1, 1–4
Infant death, small for gestational age	INFSGA	<1, 1–4
Coronary heart disease	ISCHTD	20–85+
Acute pancreatitis	ACUPAN	20–85+
Chronic pancreatitis	CHRPAN	20–85+
Pneumonia	PNEUMO	20–64
Stroke, ischemic	STRISC	20–85+
Stroke, hemorrhagic	STRHEM	20–85+
Unprovoked seizures, epilepsy or seizure disorder	SEIZUR	20–85+
Acute causes		
100% attributable-attributable		
Alcohol poisoning	ALCPOI	15–85+
Suicide by and exposure to alcohol	ALCSUC	15–85+
Direct alcohol-attributable fraction estimates		
Air-space transport	AIRTRA	15–85+
Aspiration	ASPIRA	15–85+
Child maltreatment	CHLMAL	<1, 1–4, 5–9, 10–14
Homicide	HOMICI	15–85+
Drowning	DROWNI	15–85+
Fall injuries	FALLIN	15–69
Fire injuries	FIREIN	15–85+
Firearm injuries	FIREAR	15–85+
Hypothermia	HYPOTH	15–85+
Motor vehicle non-traffic crashes	MTVNTR	15–85+

Cause of Death	Cause of Death Abbreviation	Age Range (using 5-year age groupings)
Motor vehicle traffic crashes	MTVTRA	<1, 1–4, 5–9, 10–14, 15–85+
Occupation and machine injuries	MACHOB	15–85+
Other road vehicle crashes	ORVACC	15–85+
Poisoning (not alcohol)	OTHPOI	15–85+
Suicide	SUICID	15–85+
Water transport	WATTRA	15–85+

3. Locations for Custom Data File

Use one of the following location names for the state field in custom data sets:

- United States
- Alabama
- Alaska
- Arizona
- Arkansas
- California
- Colorado
- Connecticut
- Delaware
- District of Columbia
- Florida
- Georgia
- Hawaii
- Idaho
- Illinois
- Indiana
- Iowa
- Kansas
- Kentucky
- Louisiana
- Maine
- Maryland
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Missouri
- Montana
- Nebraska
- Nevada
- New Hampshire
- New Jersey
- New Mexico
- New York
- North Carolina
- North Dakota
- Ohio
- Oklahoma
- Oregon
- Pennsylvania
- Rhode Island
- South Carolina
- South Dakota
- Tennessee
- Texas
- Utah
- Vermont
- Virginia
- Washington
- West Virginia
- Wisconsin
- Wyoming

Appendix B. Sample SAS Codes

1. Age Group Codes

Use the following age group codes for the mortality data set.

Age Group	Age Group Code
<1	AgeA
1–4	AgeB
5–9	AgeC
10–14	AgeD
15–19	AgeE
20–24	AgeF
25–29	AgeG
30–34	AgeH
35–39	AgeI
40–44	AgeJ
45–49	AgeK
50–54	AgeL
55–59	AgeM
60–64	AgeN
65–69	AgeO
70–74	AgeP
75–79	AgeQ
80–84	AgeR
85+	AgeS
Unknown*	AgeU*

*Delete age groups that are labeled with “AgeU” before uploading into ARDI.

2. Age Group Conversion

```
/******  
/* SAS Version 9.4  
/* ARDI  
/* Program: Recode Age Groupings for export to SQL  
/* Data sets used: Mortality  
/******  
/* The following code is written for the underlying cause of death file from  
the National Center for Health Statistics. */  
/* All conditions use ICD-10 coding.  
  
/* DEFINE YOUR OWN LIBNAME */  
/* Example: libname ardi ' Q:\ARDI\MORT\DATA' */  
  
libname name 'specify location of data';  
  
/* SPECIFY NAME OF DATA SET "mortwrk" IS THE EXAMPLE DATA SET NAME USED HERE.  
*/  
  
data mortwrk;  
set "data";  
  
length agegrp $5;  
  
/* CODES FOR DEFINING THE AGE GROUPS USED IN ARDI*/  
  
if AGE in ( 01 02) then agegrp = "AgeA"; /* <1 yrs */  
if AGE in ( 03 04 05 06) then agegrp = "AgeB"; /* 1-4 yrs*/  
if AGE = 07 then agegrp = "AgeC"; /* 5-9 yrs */  
if AGE = 08 then agegrp = "AgeD"; /* 10-14 yrs */  
if AGE = 09 then agegrp = "AgeE"; /* 15-19 yrs */  
if AGE = 10 then agegrp = "AgeF"; /* 20-24 yrs */  
if AGE = 11 then agegrp = "AgeG"; /* 25-29 yrs */  
if AGE = 12 then agegrp = "AgeH"; /* 30-34 yrs */  
if AGE = 13 then agegrp = "AgeI"; /* 35-39 yrs */  
if AGE = 14 then agegrp = "AgeJ"; /* 40-44 yrs */  
if AGE = 15 then agegrp = "AgeK"; /* 45-49 yrs */  
if AGE = 16 then agegrp = "AgeL"; /* 50-54 yrs */  
if AGE = 17 then agegrp = "AgeM"; /* 55-59 yrs */  
if AGE = 18 then agegrp = "AgeN"; /* 60-64 yrs */  
if AGE = 19 then agegrp = "AgeO"; /* 65-69 yrs */  
if AGE = 20 then agegrp = "AgeP"; /* 70-74 yrs */  
if AGE = 21 then agegrp = "AgeQ"; /* 75-79 yrs */  
if AGE = 22 then agegrp = "AgeR"; /* 80-84 yrs */  
if AGE in (23 24 25 26) then agegrp = "AgeS"; /* 85 + yrs */  
if AGE = 27 then agegrp = "AgeU"; /* unknown */  
run;  
  
/* CHECK YOUR DATA TO MAKE SURE ALL AGES ARE LABELED WITH A CODE*/  
proc print data= mortwrk (obs =100);  
run;  
  
proc freq data = mortwrk;  
tables agegrp / missing;  
run;
```

3. Total Deaths by State

When using the sample SAS program from this manual for single state analyses, you'll adapt the program. Single state analyses do not require the data from other states, but this code includes the United States, as well as all 50 states and the District of Columbia. It is assumed that you are using national formats for the codes. However, some states may use a different coding scheme for their subcodes (e.g., three digits). The sample code in this manual does not take this into consideration, so it may not be applied accurately for data sets of other formats.

```
/******  
/* SAS Version 9.4  
*/  
/* ARDI CHRONIC and ACUTE CONDITIONS BY DATASETNAME */  
/* Program: To Generate Frequencies */  
/* Data sets used: Mortality: "NAME OF DATA SET" */  
/******  
  
/******  
  
/* CREATE MACROS FOR AGE GROUPS DEFINED IN THE DOCUMENT: "Age Group  
Codes for SAS XML Import"*/  
  
/* AGE 0 TO 4 YEARS: USED FOR INFANT DEATH FROM MATERNAL USE OF  
ALCOHOL, LOW BIRTH WEIGHT, PRETERM BIRTH AND SMALL FOR GESTATIONAL AGE  
*/  
%macro infantage;  
  (agegrp in ("AgeA", "AgeB"))  
%mend;  
  
/* AGE 0 TO 14 YEARS: USED FOR CHILD MALTREATMENT DEATHS*/  
%macro juvenileage;  
  (agegrp in ("AgeA", "AgeB", "AgeC", "AgeD"))  
%mend;  
  
/* AGE 15 YEARS AND OLDER: USED FOR MOST ACUTE CAUSES OF DEATH*/  
%macro acutecond;  
  (agegrp in ("AgeE", "AgeF", "AgeG", "AgeH", "AgeI", "AgeJ", "AgeK",  
"AgeL", "AgeM", "AgeN", "AgeO", "AgeP", "AgeQ", "AgeR", "AgeS"))  
%mend;  
  
/* AGE 20 YEARS AND OLDER: USED FOR MOST CHRONIC CAUSES OF DEATH*/  
%macro chroniccond;  
  (agegrp in ("AgeF", "AgeG", "AgeH", "AgeI", "AgeJ", "AgeK", "AgeL",  
"AgeM", "AgeN", "AgeO", "AgeP", "AgeQ", "AgeR", "AgeS"))  
%mend;  
  
/* AGE 20 TO 64 YEARS: USED FOR PNEUMONIA ONLY */  
%macro pneumoniaage;
```

```

(agegrp in ("AgeF", "AgeG", "AgeH", "AgeI", "AgeJ", "AgeK", "AgeL",
"AgeM", "AgeN"))
%mend;

/* AGE 15 TO 69 YEARS: USED FOR FALLS ONLY */
%macro fallsage;
(agegrp in ("AgeE", "AgeF", "AgeG", "AgeH", "AgeI", "AgeJ", "AgeK",
"AgeL", "AgeM", "AgeN", "AgeO"))
%mend;

/* DEFINE YOUR OWN LIBNAME */
/* EXAMPLE: libname ardi ' Q:\ARDI\MORT\DATA' */

libname in '\\FILELOCATION\';

/*CREATE DATA FILE THAT HAS THE FOLLOWING COMPONENTS: 1) DATASETNAME 2)
STATE 3) AGE GROUP CODE 4) CAUSE 5) DEATHS FOR MALES 6) DEATHS FOR
FEMALES */

data mortc;
set in.mort1519;
length cause $6;
/* CREATE YOUR DATASETNAME FIELD*/
/* EXAMPLE datasetname = "United States"; */
/* DO NOT USE A YEAR OR OTHER NUMBER IN THE DATASETNAME*/
datasetname = 'United States';

cause = "UNASIGN";
/*
create permanent dataset
proc sort data = ardi.mort1519;
  by icd10dx;
run;*/

/* SELECT ARDI CAUSE OF DEATH "cause" FIELD */
select;

/***** CHRONIC CAUSES OF DEATH *****/
/***** 100% ALCOHOL-ATTRIBUTABLE *****/

/* Alcoholic psychosis F10.3 - F10.9 */
when ((icd10dx in ('F103', 'F104', 'F105', 'F106', 'F107', 'F108',
'F109')) and %chroniccond) cause="ALCPSY";

/* Alcohol abuse F10.0,10.1 */
when ((icd10dx in ('F100', 'F101')) and %chroniccond) cause = "ALCABU";

/* Alcohol dependence syndrome F10.2 */
when ((icd10dx = 'F102') and %chroniccond) cause = "ALCDSY";

/* Alcohol polyneuropathy G62.1*/

```

```

when ((icd10dx = 'G621') and %chroniccond) cause = "ALCPOL";

/* Degeneration of nervous system due to Alcohol G31.2*/
when ((icd10dx = 'G312') and %chroniccond) cause = "ALCNER";

/* Alcohol myopathy G72.1*/
when ((icd10dx = 'G721') and %chroniccond) cause = "ALCMYO";

/* Alcohol cardiomyopathy I42.6 */
when ((icd10dx = 'I426') and %chroniccond) cause= "ALCCAR";

/* Alcoholic gastritis K29.2 */
when ((icd10dx = 'K292') and %chroniccond) cause= "ALCGAS";

/* Alcoholic liver disease K70-K70.4, K70.9 */
when ((icd10dx in ('K700', 'K701', 'K702', 'K703', 'K704') OR (icd10dx
= 'K709'))
and %chroniccond) cause= "ALCLID";

/* Alcohol induced chronic pancreatitis K86.0 */
when ((icd10dx = 'K860') and %chroniccond) cause= "ALCCRP";

/* Alcohol-induced acute pancreatitis K85.2 */
when ((icd10dx = 'K852')and %chroniccond) cause= "ALCACP";

/* Fetal alcohol syndrome Q86.0, age macros not used with this
condition */
when (icd10dx = 'Q860') cause= "ALCFAS";

/* Fetus and newborn affected by maternal use of alcohol P04.3 */
when ((icd10dx = 'P043') and %infantage) cause= "ALCMAU";

/*****CHRONIC Conditions - DIRECT AAF ESTIMATES *****/

/* Esophageal varices I85 */
when ((icd10dx =: 'I85') and %chroniccond) cause= "ESOVAR";

/* Gastroesophageal hemorrhage K22.6 */
when ((icd10dx = 'K226') and %chroniccond) cause= "GASHEM";

/* Liver cirrhosis, unspecified K74.0-K74.2, K74.6, K76.0, K76.7, K
76.9*/
when ((icd10dx in ('K740', 'K741', 'K742', 'K746', 'K760', 'K767',
'K769')) and %chroniccond) cause= "LIVCIR";

/* Portal Hypertension (K76.6) */
when ((icd10dx = 'K766') and %chroniccond) cause="PORHYP";

/***** CHRONIC Conditions - INDIRECT AAF ESTIMATES *****/

```

```

/* Atrial fibrillation I48 */
when ((icd10dx =: 'I48') and %chroniccond) cause="ATRFIB";

/* Cancer, female breast (C50) */
when ((icd10dx =: 'C50') and %chroniccond) cause="CANBRE";

/* Cancer, colorectal (C18,C20) */
when (((icd10dx =: 'C18') OR (icd10dx =: 'C20')) and %chroniccond)
cause= "CANCOL";

/* Cancer, esophageal (C15) */
when ((icd10dx =: 'C15') and %chroniccond) cause= "CANESO";

/* Cancer, laryngeal (C32) */
when ((icd10dx =: 'C32') and %chroniccond) cause= "CANLAR";

/* Cancer, liver (C22) */
when ((icd10dx =: 'C22') and %chroniccond) cause= "CANLIV";

/* Cancer, oral cavity and pharyngeal (C01-C06,C09-C10,C12-C14) */
when (((icd10dx=: 'C01') or (icd10dx=: 'C02') or (icd10dx=: 'C03')
      or (icd10dx=: 'C04') or (icd10dx=: 'C05') or (icd10dx=: 'C06')
      or (icd10dx=: 'C09') or (icd10dx=: 'C10') or (icd10dx=: 'C12')
      or (icd10dx=: 'C13') or (icd10dx=: 'C14')) and %chroniccond) cause
= "CANOCP";

/* Cancer, pancreatic (C25) */
when ((icd10dx =: 'C25') and %chroniccond) cause= "CANPAN";

/* Cancer, male prostate (C61) */
when ((icd10dx =: 'C61') and %chroniccond) cause= "CANPRO";

/* Cancer, stomach (C16) */
when ((icd10dx =: 'C16') and %chroniccond) cause= "CANSTO";

/* Chronic hepatitis (K73) */
when ((icd10dx =: 'K73') and %chroniccond) cause= "CHRHEP";

/* Gallbladder disease (K80,K81,K83) */
when (((icd10dx =: 'K80') OR (icd10dx =: 'K81') OR (icd10dx =: 'K83'))
and %chroniccond) cause= "GALLBL";

/* Hypertension (I10-I13,I15) */
when (((icd10dx =: 'I10') or (icd10dx =: 'I11') or (icd10dx =: 'I12')
or (icd10dx =: 'I13') or (icd10dx =: 'I15'))
and %chroniccond) cause= "HYPERT";

/* Infant death, low birth weight (P05.0, P07.0, P07.1) */
when ((icd10dx in ('P050','P070','P071')) and %infantage) cause=
"INFLBW";

```

```

/* Infant death, pre-term birth (P07.2, P07.3) */
when ((icd10dx in ('P072','P073')) and %infantage) cause= "INFPRE";

/* Infant death, small for gestational age (P05.1) */
when ((icd10dx in ('P051')) and %infantage) cause= "INFSGA";

/* Coronary heart disease (I20-I25) */
when (((icd10dx =: 'I20') or (icd10dx =: 'I21') or (icd10dx =: 'I22')
or (icd10dx =: 'I23') or (icd10dx =: 'I24') or (icd10dx =: 'I25'))
and %chroniccond) cause= "ISCHTD";

/* Pancreatitis, acute (K85.0, K85.1, K85.3, K85.8, K85.9) */
when ((icd10dx in ('K850','K851','K853','K858','K859')) and
%chroniccond) cause="ACUPAN";

/* Pancreatitis, chronic (K86.1) */
when ((icd10dx = 'K861') and %chroniccond) cause= "CHRPAN";

/* Pneumonia (J12-J16, J18) */
when (((icd10dx =: 'J12') or (icd10dx =: 'J13') or (icd10dx =: 'J14')
or (icd10dx =: 'J15') or (icd10dx =: 'J16') or (icd10dx =: 'J18'))
and %pneumoniaage) cause= "PNEUMO";
/* FOOTNOTE: Deaths among those aged 20-64 years */

/* Stroke, ischemic (G45, I63, I65-I67, I69.3) */
when (((icd10dx =: 'G45') OR (icd10dx =: 'I63') OR (icd10dx =: 'I65')
OR (icd10dx =: 'I66') or (icd10dx =: 'I67') or (icd10dx =: 'I693'))
and %chroniccond) cause= "STRISC";

/* Stroke, hemorrhagic (I60-I62, I69.0-I69.2) */
when (((icd10dx =: 'I60') OR (icd10dx =: 'I61') or (icd10dx =: 'I62')
or (icd10dx in ('I690', 'I691', 'I692'))))
and %chroniccond) cause= "STRHEM";

/* Unprovoked seizures, epilepsy or seizure disorder (G40, G41, R56.8)
*/
when (((icd10dx =: 'G40') or (icd10dx =: 'G41') or (icd10dx =: 'R568'))
and %chroniccond) cause= "SEIZUR";

/***** ACUTE CAUSES OF DEATH *****/
/***** 100% ALCOHOL-ATTRIBUTABLE *****/

/* Alcohol poisoning (X45, Y15) */
when (((icd10dx =: 'X45') or (icd10dx =: 'Y15'))
and %acutecond) cause= "ALCPOI";

/* Suicide by and exposure to alcohol (X65) */
when ((icd10dx =: 'X65') and %acutecond) cause= "ALCSUC";

/***** DIRECT AAF ESTIMATE *****/

```



```

/* Air space transport (V95-V97) */
when (((icd10dx =: 'V95') or (icd10dx =:'V96') or (icd10dx =:'V97'))
      and %acutecond) cause= "AIRTRA";

/* Aspiration (W78-W79) */
when (((icd10dx =: 'W78') or (icd10dx =:'W79'))
      and %acutecond) cause= "ASPIRA";

/* Child Maltreatment / Homicide (X85 -Y09, Y87.1) */
/* SAME CODES BUT DIFFERENT AGE GROUPS */
when (((icd10dx =: 'X85') or (icd10dx =: 'X86') or (icd10dx =: 'X87')
      or (icd10dx =: 'X88') or (icd10dx =: 'X89') or (icd10dx =: 'X90')
      or (icd10dx =: 'X91') or (icd10dx =: 'X92') or (icd10dx =: 'X93')
      or (icd10dx =: 'X94') or (icd10dx =: 'X95') or (icd10dx =: 'X96')
      or (icd10dx =: 'X97') or (icd10dx =: 'X98') or (icd10dx =: 'X99')
      or (icd10dx =: 'Y00') or (icd10dx =: 'Y01') or (icd10dx =: 'Y02')
      or (icd10dx =: 'Y03') or (icd10dx =: 'Y04') or (icd10dx =: 'Y05')
      or (icd10dx =: 'Y06') or (icd10dx =: 'Y07') or (icd10dx =: 'Y08')
      or (icd10dx =: 'Y09') or (icd10dx =: 'Y871')) do;
  if %juvenileage then cause="CHLMAL";
  else cause = "HOMICI";
end;
/* FOOTNOTE: Deaths among those aged 0-14 years only for CHLMAL
otherwise HOMICI */

/* Drowning (W65-W70, W73, W74,Y21 ) */
when (((icd10dx =: 'W65') or (icd10dx =: 'W66') or (icd10dx =: 'W67')
      or (icd10dx =: 'W68') or (icd10dx =: 'W69') or (icd10dx =: 'W70')
      or (icd10dx =: 'W73') or (icd10dx =: 'W74') or (icd10dx =: 'Y21'))
      and %acutecond) cause= "DROWNI";

/* Falls (W00-W19, Y30) */
when (((icd10dx =: 'W00') or (icd10dx =: 'W01') or (icd10dx =: 'W02')
      or (icd10dx =: 'W03') or (icd10dx =: 'W04') or (icd10dx =: 'W05')
      or (icd10dx =: 'W06') or (icd10dx =: 'W07') or (icd10dx =: 'W08')
      or (icd10dx =: 'W09') or (icd10dx =: 'W10') or (icd10dx =: 'W11')
      or (icd10dx =: 'W12') or (icd10dx =: 'W13') or (icd10dx =: 'W14')
      or (icd10dx =: 'W15') or (icd10dx =: 'W16') or (icd10dx =: 'W17')
      or (icd10dx =: 'W18') or (icd10dx =: 'W19') or (icd10dx =: 'Y30'))
      and %fallsage) cause= "FALLIN";
/* FOOTNOTE: Deaths among those aged 15 to 69 */

/* Fire (X00-X06, X08, X09, Y26) */
when (((icd10dx =: 'X00') or (icd10dx =: 'X01') or (icd10dx =: 'X02')
      or (icd10dx =: 'X03') or (icd10dx =: 'X04') or (icd10dx =: 'X05')
      or (icd10dx =: 'X06') or (icd10dx =: 'X08') or (icd10dx =: 'X09')
      or (icd10dx =: 'Y26'))
      and %acutecond) cause= "FIREIN";

/* Firearms (W32-W34,Y22-Y24) */
when (((icd10dx =: 'W32') or (icd10dx =:'W33') or (icd10dx =:'W34'))

```

```

        or (icd10dx =: 'Y22') or (icd10dx =: 'Y23') or (icd10dx =:
'Y24'))
        and %acutecond) cause= "FIREAR";

/* Hypothermia (X31) */
when ((icd10dx =: 'X31') and %acutecond) cause= "HYPOTH";

/* Motor vehicle nontraffic */
when (((icd10dx in ('V020', 'V030', 'V040', 'V090', 'V810', 'V820',
'V890'))
        OR (icd10dx in ('V120', 'V121', 'V122'))
        OR (icd10dx in ('V130', 'V131', 'V132'))
        OR (icd10dx in ('V140', 'V141', 'V142'))
        OR (icd10dx in ('V190', 'V191', 'V192'))
OR (icd10dx in ('V200', 'V201', 'V202'))
OR (icd10dx in ('V210', 'V211', 'V212'))
OR (icd10dx in ('V220', 'V221', 'V222'))
OR (icd10dx in ('V230', 'V231', 'V232'))
OR (icd10dx in ('V240', 'V241', 'V242'))
OR (icd10dx in ('V250', 'V251', 'V252'))
OR (icd10dx in ('V260', 'V261', 'V262'))
OR (icd10dx in ('V270', 'V271', 'V272'))
OR (icd10dx in ('V280', 'V281', 'V282'))
OR (icd10dx in ('V290', 'V291', 'V292', 'V293'))
OR (icd10dx in ('V300', 'V301', 'V302', 'V303'))
OR (icd10dx in ('V310', 'V311', 'V312', 'V313'))
OR (icd10dx in ('V320', 'V321', 'V322', 'V323'))
OR (icd10dx in ('V330', 'V331', 'V332', 'V333'))
OR (icd10dx in ('V340', 'V341', 'V342', 'V343'))
OR (icd10dx in ('V350', 'V351', 'V352', 'V353'))
OR (icd10dx in ('V360', 'V361', 'V362', 'V363'))
OR (icd10dx in ('V370', 'V371', 'V372', 'V373'))
OR (icd10dx in ('V380', 'V381', 'V382', 'V383'))
OR (icd10dx in ('V390', 'V391', 'V392', 'V393'))
OR (icd10dx in ('V400', 'V401', 'V402', 'V403'))
OR (icd10dx in ('V410', 'V411', 'V412', 'V413'))
OR (icd10dx in ('V420', 'V421', 'V422', 'V423'))
OR (icd10dx in ('V430', 'V431', 'V432', 'V433'))
OR (icd10dx in ('V440', 'V441', 'V442', 'V443'))
OR (icd10dx in ('V450', 'V451', 'V452', 'V453'))
OR (icd10dx in ('V460', 'V461', 'V462', 'V463'))
OR (icd10dx in ('V470', 'V471', 'V472', 'V473'))
OR (icd10dx in ('V480', 'V481', 'V482', 'V483'))
OR (icd10dx in ('V490', 'V491', 'V492', 'V493'))
        OR (icd10dx in ('V500', 'V501', 'V502', 'V503'))
OR (icd10dx in ('V510', 'V511', 'V512', 'V513'))
OR (icd10dx in ('V520', 'V521', 'V522', 'V523'))
OR (icd10dx in ('V530', 'V531', 'V532', 'V533'))
OR (icd10dx in ('V540', 'V541', 'V542', 'V543'))
OR (icd10dx in ('V550', 'V551', 'V552', 'V553'))
OR (icd10dx in ('V560', 'V561', 'V562', 'V563'))

```

```

OR (icd10dx in ('V570', 'V571', 'V572', 'V573'))
OR (icd10dx in ('V580', 'V581', 'V582', 'V583'))
OR (icd10dx in ('V590', 'V591', 'V592', 'V593'))
    OR (icd10dx in ('V600', 'V601', 'V602', 'V603'))
OR (icd10dx in ('V610', 'V611', 'V612', 'V613'))
OR (icd10dx in ('V620', 'V621', 'V622', 'V623'))
OR (icd10dx in ('V630', 'V631', 'V632', 'V633'))
OR (icd10dx in ('V640', 'V641', 'V642', 'V643'))
OR (icd10dx in ('V650', 'V651', 'V652', 'V653'))
OR (icd10dx in ('V660', 'V661', 'V662', 'V663'))
OR (icd10dx in ('V670', 'V671', 'V672', 'V673'))
OR (icd10dx in ('V680', 'V681', 'V682', 'V683'))
OR (icd10dx in ('V690', 'V691', 'V692', 'V693'))
OR (icd10dx in ('V700', 'V701', 'V702', 'V703'))
OR (icd10dx in ('V710', 'V711', 'V712', 'V713'))
OR (icd10dx in ('V720', 'V721', 'V722', 'V723'))
OR (icd10dx in ('V730', 'V731', 'V732', 'V733'))
OR (icd10dx in ('V740', 'V741', 'V742', 'V743'))
OR (icd10dx in ('V750', 'V751', 'V752', 'V753'))
OR (icd10dx in ('V760', 'V761', 'V762', 'V763'))
OR (icd10dx in ('V770', 'V771', 'V772', 'V773'))
OR (icd10dx in ('V780', 'V781', 'V782', 'V783'))
OR (icd10dx in ('V790', 'V791', 'V792', 'V793'))
OR (icd10dx in ('V834', 'V835', 'V836', 'V837', 'V839'))
OR (icd10dx in ('V844', 'V845', 'V846', 'V847', 'V849'))
OR (icd10dx in ('V854', 'V855', 'V856', 'V857', 'V859'))
OR (icd10dx in ('V864', 'V864', 'V866', 'V867', 'V869'))
OR (icd10dx in ('V880', 'V881', 'V882', 'V883', 'V884', 'V885',
'V886', 'V887', 'V888'))
    and %acutecond) cause= "MTVNTR";

/* Motor Vehicle traffic age macros not used with this condition */
when ((icd10dx in ('V021', 'V029', 'V031', 'V039', 'V041', 'V049',
'V092', 'V194', 'V195', 'V196', 'V892',
'V811', 'V821'))
OR (icd10dx in ('V123', 'V124', 'V125', 'V129'))
OR (icd10dx in ('V133', 'V134', 'V135', 'V139'))
OR (icd10dx in ('V143', 'V144', 'V145', 'V149'))
OR (icd10dx in ('V203', 'V204', 'V205', 'V209'))
OR (icd10dx in ('V213', 'V214', 'V215', 'V219'))
OR (icd10dx in ('V223', 'V224', 'V225', 'V229'))
OR (icd10dx in ('V233', 'V234', 'V235', 'V239'))
OR (icd10dx in ('V243', 'V244', 'V245', 'V249'))
OR (icd10dx in ('V253', 'V254', 'V255', 'V259'))
OR (icd10dx in ('V263', 'V264', 'V265', 'V269'))
OR (icd10dx in ('V273', 'V274', 'V275', 'V279'))
OR (icd10dx in ('V283', 'V284', 'V285', 'V289'))
OR (icd10dx in ('V294', 'V295', 'V296', 'V298', 'V299'))
OR (icd10dx in ('V304', 'V305', 'V306', 'V307', 'V309'))
OR (icd10dx in ('V314', 'V315', 'V316', 'V317', 'V319'))
OR (icd10dx in ('V324', 'V325', 'V326', 'V327', 'V329'))

```

```

OR (icd10dx in ('V334', 'V335', 'V336', 'V337', 'V339'))
OR (icd10dx in ('V344', 'V345', 'V346', 'V347', 'V349'))
OR (icd10dx in ('V354', 'V355', 'V356', 'V357', 'V359'))
OR (icd10dx in ('V364', 'V365', 'V366', 'V367', 'V369'))
OR (icd10dx in ('V374', 'V375', 'V376', 'V377', 'V379'))
OR (icd10dx in ('V384', 'V385', 'V386', 'V387', 'V389'))
OR (icd10dx in ('V394', 'V395', 'V396', 'V398', 'V399'))
OR (icd10dx in ('V404', 'V405', 'V406', 'V407', 'V408', 'V409'))
OR (icd10dx in ('V414', 'V415', 'V416', 'V417', 'V419'))
OR (icd10dx in ('V424', 'V425', 'V426', 'V427', 'V429'))
OR (icd10dx in ('V434', 'V435', 'V436', 'V437', 'V439'))
OR (icd10dx in ('V444', 'V445', 'V446', 'V447', 'V449'))
OR (icd10dx in ('V454', 'V455', 'V456', 'V457', 'V459'))
OR (icd10dx in ('V464', 'V465', 'V466', 'V467', 'V469'))
OR (icd10dx in ('V474', 'V475', 'V476', 'V477', 'V479'))
OR (icd10dx in ('V484', 'V485', 'V486', 'V487', 'V489'))
OR (icd10dx in ('V494', 'V495', 'V496', 'V498', 'V499'))
OR (icd10dx in ('V504', 'V505', 'V506', 'V507', 'V509'))
OR (icd10dx in ('V514', 'V515', 'V516', 'V517', 'V519'))
OR (icd10dx in ('V524', 'V525', 'V526', 'V527', 'V529'))
OR (icd10dx in ('V534', 'V535', 'V536', 'V537', 'V539'))
OR (icd10dx in ('V544', 'V545', 'V546', 'V547', 'V549'))
OR (icd10dx in ('V554', 'V555', 'V556', 'V557', 'V559'))
OR (icd10dx in ('V564', 'V565', 'V566', 'V567', 'V569'))
OR (icd10dx in ('V574', 'V575', 'V576', 'V577', 'V579'))
OR (icd10dx in ('V584', 'V585', 'V586', 'V587', 'V589'))
OR (icd10dx in ('V594', 'V595', 'V596', 'V598', 'V599'))
OR (icd10dx in ('V604', 'V605', 'V606', 'V607', 'V609'))
OR (icd10dx in ('V614', 'V615', 'V616', 'V617', 'V619'))
OR (icd10dx in ('V624', 'V625', 'V626', 'V627', 'V629'))
OR (icd10dx in ('V634', 'V635', 'V636', 'V637', 'V639'))
OR (icd10dx in ('V644', 'V645', 'V646', 'V647', 'V649'))
OR (icd10dx in ('V654', 'V655', 'V656', 'V657', 'V659'))
OR (icd10dx in ('V664', 'V665', 'V666', 'V667', 'V669'))
OR (icd10dx in ('V674', 'V675', 'V676', 'V677', 'V679'))
OR (icd10dx in ('V684', 'V685', 'V686', 'V687', 'V689'))
OR (icd10dx in ('V694', 'V695', 'V696', 'V698', 'V699'))
OR (icd10dx in ('V704', 'V705', 'V706', 'V707', 'V709'))
OR (icd10dx in ('V714', 'V715', 'V716', 'V717', 'V719'))
OR (icd10dx in ('V724', 'V725', 'V726', 'V727', 'V729'))
OR (icd10dx in ('V734', 'V735', 'V736', 'V737', 'V739'))
OR (icd10dx in ('V744', 'V745', 'V746', 'V747', 'V749'))
OR (icd10dx in ('V754', 'V755', 'V756', 'V757', 'V759'))
OR (icd10dx in ('V764', 'V765', 'V766', 'V767', 'V769'))
OR (icd10dx in ('V774', 'V775', 'V776', 'V777', 'V779'))
OR (icd10dx in ('V784', 'V785', 'V786', 'V787', 'V789'))
OR (icd10dx in ('V794', 'V795', 'V796', 'V798', 'V799'))
OR (icd10dx in ('V803', 'V804', 'V805'))
OR (icd10dx in ('V830', 'V831', 'V832', 'V833'))
OR (icd10dx in ('V840', 'V841', 'V842', 'V843'))
OR (icd10dx in ('V850', 'V851', 'V852', 'V853'))

```

```

    OR (icd10dx in ('V860', 'V861', 'V862', 'V863'))
    OR (icd10dx in ('V870', 'V871', 'V872', 'V873', 'V874', 'V875',
'V876', 'V877', 'V878'))))
    cause= "MTVTRA";
/* FOOTNOTE: Deaths among those of all ages. */

/* Occupational and machine and objects (W24-W31, W45) */
when (((icd10dx = 'W24') or (icd10dx =:'W25') or (icd10dx =:'W26')
    or (icd10dx =:'W27') or (icd10dx =:'W28') or (icd10dx =:'W29')
    or (icd10dx =:'W30') or (icd10dx =:'W31') or (icd10dx =:'W45'))
    and %acutecond) cause= "MACHOB";

/* Other road vehicle accidents */
when (((icd10dx in ('V091', 'V093', 'V099', 'V193', 'V879', 'V889',
'V891', 'V893', 'V899'))
    OR (icd10dx =:'V01') OR (icd10dx =:'V05') OR (icd10dx =:'V06')
    OR (icd10dx =:'V10') or (icd10dx =:'V11')
    OR (icd10dx =:'V15') or (icd10dx =:'V16') or (icd10dx =:'V17')
    OR (icd10dx =:'V18') or (icd10dx in ('V198', 'V199'))
    OR (icd10dx in ('V800', 'V801', 'V802'))
    OR (icd10dx in ('V806', 'V807', 'V808', 'V809'))
    OR (icd10dx in ('V812', 'V813', 'V814', 'V815', 'V816', 'V817',
'V818', 'V819'))
    OR (icd10dx in ('V822', 'V823', 'V824', 'V825', 'V826',
'V827', 'V828', 'V829'))))
    and %acutecond) cause= "ORVACC";

/* Poisoning (not alcohol) (X40-44, X46-49, Y10-Y14, Y16-Y19) */
when (((icd10dx =:'X40') or (icd10dx =:'X41') or (icd10dx =:'X42')
    or (icd10dx =:'X43') or (icd10dx =:'X44') or (icd10dx =:'X46')
    or (icd10dx =:'X47') or (icd10dx =:'X48') or (icd10dx =:'X49')
    or (icd10dx =:'Y10') or (icd10dx =:'Y11') or (icd10dx =:'Y12')
    or (icd10dx =:'Y13') or (icd10dx =:'Y14') or (icd10dx =:'Y16')
    or (icd10dx =:'Y17') or (icd10dx =:'Y18') or (icd10dx =:'Y19'))
    and %acutecond) cause= "OTHPOI";

/* Suicide (X60-64, X66-84, Y87.0) */
when (((icd10dx =:'X60') or (icd10dx =:'X61') or (icd10dx =:'X62')
    or (icd10dx =:'X63') or (icd10dx =:'X64') or (icd10dx =:'X66')
    or (icd10dx =:'X67') or (icd10dx =:'X68') or (icd10dx =:'X69')
    or (icd10dx =:'X70') or (icd10dx =:'X71') or (icd10dx =:'X72')
    or (icd10dx =:'X73') or (icd10dx =:'X74') or (icd10dx =:'X75')
    or (icd10dx =:'X76') or (icd10dx =:'X77') or (icd10dx =:'X78')
    or (icd10dx =:'X79') or (icd10dx =:'X80') or (icd10dx =:'X81')
    or (icd10dx =:'X82') or (icd10dx =:'X83') or (icd10dx =:'X84')
    or (icd10dx =:'Y870')) and %acutecond) cause= "SUICID";

/* Water Transport (V90-V94) */
when (((icd10dx =:'V90') or (icd10dx =:'V91') or (icd10dx =:'V92')
    or (icd10dx =:'V93') or (icd10dx =:'V94')) and %acutecond)
    cause= "WATTRA";

```

```

/*LABEL ALL OTHER CAUSES OF DEATH UNKNOWN*/
otherwise cause = "UNKWN";
end;

if cause NE "UNKWN"; /*DELETE UNKNOWN CAUSES OF DEATH */

LENGTH nstate $30  nsex $6;

if stfips_res = ' ' then nstate = 'UNITED STATES';
else if stfips_res = 'AL' then nstate = 'ALABAMA';
else if stfips_res = 'AK' then nstate = 'ALASKA';
else if stfips_res = 'AR' then nstate = 'ARKANSAS';
else if stfips_res = 'AS' then nstate = 'AMERICAN SAMOA';
else if stfips_res = 'AZ' then nstate = 'ARIZONA';
else if stfips_res = 'CA' then nstate = 'CALIFORNIA';
else if stfips_res = 'CO' then nstate = 'COLORADO';
else if stfips_res = 'CT' then nstate = 'CONNECTICUT';
else if stfips_res = 'DC' then nstate = 'DISTRICT OF COLUMBIA';
else if stfips_res = 'DE' then nstate = 'DELAWARE';
else if stfips_res = 'FL' then nstate = 'FLORIDA';
else if stfips_res = 'GA' then nstate = 'GEORGIA';
else if stfips_res = 'GU' then nstate = 'GUAM';
else if stfips_res = 'HI' then nstate = 'HAWAII';
else if stfips_res = 'IA' then nstate = 'IOWA';
else if stfips_res = 'ID' then nstate = 'IDAHO';
else if stfips_res = 'IL' then nstate = 'ILLINOIS';
else if stfips_res = 'IN' then nstate = 'INDIANA';
else if stfips_res = 'KS' then nstate = 'KANSAS';
else if stfips_res = 'KY' then nstate = 'KENTUCKY';
else if stfips_res = 'LA' then nstate = 'LOUISIANA';
else if stfips_res = 'MA' then nstate = 'MASSACHUSETTS';
else if stfips_res = 'MD' then nstate = 'MARYLAND';
else if stfips_res = 'ME' then nstate = 'MAINE';
else if stfips_res = 'MI' then nstate = 'MICHIGAN';
else if stfips_res = 'MN' then nstate = 'MINNESOTA';
else if stfips_res = 'MO' then nstate = 'MISSOURI';
else if stfips_res = 'MS' then nstate = 'MISSISSIPPI';
else if stfips_res = 'MT' then nstate = 'MONTANA';
else if stfips_res = 'NC' then nstate = 'NORTH CAROLINA';
else if stfips_res = 'ND' then nstate = 'NORTH DAKOTA';
else if stfips_res = 'NE' then nstate = 'NEBRASKA';
else if stfips_res = 'NH' then nstate = 'NEW HAMPSHIRE';
else if stfips_res = 'NJ' then nstate = 'NEW JERSEY';
else if stfips_res = 'NM' then nstate = 'NEW MEXICO';
else if stfips_res = 'NV' then nstate = 'NEVADA';
else if stfips_res = 'NY' then nstate = 'NEW YORK';
else if stfips_res = 'OH' then nstate = 'OHIO';
else if stfips_res = 'OK' then nstate = 'OKLAHOMA';
else if stfips_res = 'OR' then nstate = 'OREGON';
else if stfips_res = 'PA' then nstate = 'PENNSYLVANIA';

```

```

else if stfips_res = 'PR' then nstate = 'PUERTO RICO';
else if stfips_res = 'RI' then nstate = 'RHODE ISLAND';
else if stfips_res = 'SC' then nstate = 'SOUTH CAROLINA';
else if stfips_res = 'SD' then nstate = 'SOUTH DAKOTA';
else if stfips_res = 'TN' then nstate = 'TENNESSEE';
else if stfips_res = 'TX' then nstate = 'TEXAS';
else if stfips_res = 'UT' then nstate = 'UTAH';
else if stfips_res = 'VA' then nstate = 'VIRGINIA';
else if stfips_res = 'VI' then nstate = 'VIRGIN ISLANDS';
else if stfips_res = 'VT' then nstate = 'VERMONT';
else if stfips_res = 'WA' then nstate = 'WASHINGTON';
else if stfips_res = 'WI' then nstate = 'WISCONSIN';
else if stfips_res = 'WV' then nstate = 'WEST VIRGINIA';
else if stfips_res = 'WY' then nstate = 'WYOMING';

/*MALE NEEDS TO SHOW UP IN THE DATASET BEFORE FEMALE*/
if sex='M' then nsex= '1_Male';
if sex='F' then nsex= '2_Female';

drop stfips_res sex;
rename nstate=state nsex=sex;

run;

/*SAMPLE CODE TO CREATE TABLE OF FREQUENCY OF CAUSE OF EACH DEATH BY
LOCATION (STATE), AGE GROUP, AND SEX;
OTHER CODES INCLUDING PROC SQL CAN BE USED*/
proc freq data = mortc;
  tables datasetname*state*agegrp*cause*sex / list missprint nopercnt
  nocum out=numsex noprint;
run;

/* TRANSPOSE TO HAVE MALE AND FEMALE ESTIMATES OUTPUT IN ROWS */
proc transpose data = numsex out = numsex2;
  by datasetname state agegrp cause;
  id sex;
  var Count;
run;

/* NO MISSING DATA IS ALLOWED, IF THERE ARE NO CAUSES OF DEATH FOR A
PARTICULAR CONDITION THEN "0" MUST BE
INCLUDED */
data numsex3;
set numsex2;
drop _name_ _label_;
if _1_Male = . then Male = 0; else Male=_1_Male;
if _2_Fema = . then Female = 0; else Female=_2_Fema;
/* This code captures the proportion of esophageal cancer deaths due to
squamous cell carcinoma, which are the proportion of */
/* esophageal cancers related to alcohol based on the Surveillance,
Epidemiology, and End Results data (SEER18). */

```

```

if cause= "CANESO" then Male=Male*0.274;
if cause= "CANESO" then Female=Female*0.532;
/* end edit */
drop _1_Male _2_Fema;
Male=round(Male/5,0.1);
Female=round(Female/5,0.1);
run;

*****
*****;
*summarize by whole US;
proc freq data = mortc;
    tables datasetname*agegrp*cause*sex / list missprint nopercnt nocum
out=numsex_us noprint;
    run;

/* TRANSPOSE TO HAVE MALE AND FEMALE ESTIMATES OUTPUT IN ROWS */
proc transpose data = numsex_us out = numsex2_us;
    by datasetname agegrp cause;
    id sex;
    var Count;
    run;

/* NO MISSING DATA IS ALLOWED, IF THERE ARE NO CAUSES OF DEATH FOR A
PARTICULAR CONDITION THEN "0" MUST BE
INCLUDED */
data numsex3_us;
set numsex2_us;
drop _name_ _label_;
if _1_Male = . then Male = 0; else Male=_1_Male;
if _2_Fema = . then Female = 0; else Female=_2_Fema;
/* This code captures the proportion of esophageal cancer deaths due to
squamous cell carcinoma, which are the proportion of */
/* esophageal cancers related to alcohol
*/
if cause= "CANESO" then Male=Male*0.274;
if cause= "CANESO" then Female=Female*0.532;
/* end edit */
drop _1_Male _2_Fema;
Male=round(Male/5,0.1);
Female=round(Female/5,0.1);
run;
*****
*****;

data numsex4;
    set numsex3_us numsex3;

    newage=agegrp;
    drop agegrp;

```



```

rename newage = agegrp;
  if state = " " then state="United States";
run;

%macro prt(state);
ods csv file="\FILELOCATION.csv";
options nonumber nodate ls=70 missing=" " ;
proc print data=numsex4 noobs;
title "Frequency of mortality from 2015-2019 in &state";
where state="&state";
run;
ods csv close;
%mend;

%prt(state=United States);
%prt(state=ALABAMA);
  %prt(state=ALASKA);
  %prt(state=ARKANSAS);
  %prt(state=ARIZONA);
  %prt(state=CALIFORNIA);
  %prt(state=COLORADO);
  %prt(state=CONNECTICUT);
  %prt(state=DISTRICT OF COLUMBIA);
  %prt(state=DELAWARE);
  %prt(state=FLORIDA);
  %prt(state=GEORGIA);
  %prt(state=HAWAII);
  %prt(state=IOWA);
  %prt(state=IDAHO);
  %prt(state=ILLINOIS);
  %prt(state=INDIANA);
  %prt(state=KANSAS);
  %prt(state=KENTUCKY);
  %prt(state=LOUISIANA);
  %prt(state=MASSACHUSETTS);
  %prt(state=MARYLAND);
  %prt(state=MAINE);
  %prt(state=MICHIGAN);
  %prt(state=MINNESOTA);
  %prt(state=MISSOURI);
  %prt(state=MISSISSIPPI);
  %prt(state=MONTANA);
  %prt(state=NORTH CAROLINA);
  %prt(state=NORTH DAKOTA);
  %prt(state=NEBRASKA);
  %prt(state=NEW HAMPSHIRE);
  %prt(state=NEW JERSEY);
  %prt(state=NEW MEXICO);
  %prt(state=NEVADA);
  %prt(state=NEW YORK);
  %prt(state=OHIO);

```

```

%prt (state=OKLAHOMA);
%prt (state=OREGON);
%prt (state=PENNSYLVANIA);
%prt (state=RHODE ISLAND);
%prt (state=SOUTH CAROLINA);
%prt (state=SOUTH DAKOTA);
%prt (state=TENNESSEE);
%prt (state=TEXAS);
%prt (state=UTAH);
%prt (state=VIRGINIA);
%prt (state=VERMONT);
%prt (state=WASHINGTON);
%prt (state=WISCONSIN);
%prt (state=WEST VIRGINIA);
%prt (state=WYOMING);

```

Troubleshooting

Example A: Coding for inclusion of all subcodes

```

when ((icd10dx =: 'C50') and %chroniccond) cause= "CANBRE";

/*The '=' sign with a ':' should tell SAS to pull all the codes that have
"C50" in them.*/

```

Example B: Coding to inclusion of only specific subcodes

There are some causes in which we only want specific subcategories within each death, hemorrhagic stroke is an example (e.g., I69.0-I69.2):

```

/* Stroke, hemorrhagic (I60-I62, I69.0-I69.2) */
when (((icd10dx =: 'I60') OR (icd10dx =: 'I61') or (icd10dx =: 'I62')
or (icd10dx in ('I690', 'I691', 'I692'))
and %chroniccond) cause= "STRHEM";

```

If your code includes numbers with 3 or more digits, make sure that the digits beyond the first two are being captured for all causes written with the “=:”. If the code is not capturing all data, you may need to add the remaining digits to your code following the equal sign (e.g., C500, C501, C502 ... C509).

4. Prevalence of Average Daily Alcohol Consumption Calculation and Sample SAS code

This section includes the sample SAS code to calculate the prevalence of alcohol consumption using the [BRFSS](#), adjusted to per capita alcohol sales.

The six steps show you how to use the sample code to calculate prevalence of average daily alcohol consumption:

1. **Adapt the codes to specify the state of interest of your prevalence analyses.**
2. **Check if the year of data that you are using has identical variable names for each variable.**
 - For example, AVEDRNK2 was the variable name used for several years prior to 2019 but was later changed to AVEDRNK3.
 - The example below uses variables from the 2015–2019 BRFSS. Different years of the BRFSS may have different variable.
3. **If combining several years of data, refer to the [BRFSS website](#) regarding combining the weighting variables used in the data sets. Always use the appropriate population weights.**
4. **Check whether all desired years of data are available for your location.** See the sample SAS code provided to address this issue if data are missing.
5. **Do not delete people for whom alcohol information is missing.** This would produce an incorrect denominator for the remaining calculations.
6. **Understand code 999:** Some of the [cut points](#) used in ARDI specify a minimum level of drinking, thus there are some people who are not included in the lowest drinking category (they are coded 999).

This sample code uses variables from the 2015–2019 BRFSS.

Step 1. Merge multiple years of data from the BRFSS (e.g., 2015–2019)

```
/* *****  
***** */  
/* STATE-SPECIFIC SAS CODE, */  
/* This program reads the one-year BRFSS data files, uses the alcohol core  
question data, and then combines the multiple years of data */  
/* NOTE: There is nothing state-specific in this Step 1. All states are  
carried to Step 2, where the specific state is selected. */  
/* *****  
***** */  
  
/* As a first step, download and save the five BRFSS public files (in ASCII  
format) corresponding to the years of interest (e.g., 2015 to 2019) */  
/* For each year, also download the SASOUTYY_LLCP.SAS program, which is then  
run to convert the above ASCII format to a SAS dataset */
```

```

/* In the SASOUTYY program, fill in the FILENAME DATAIN and LIBNAME DATAOUT
V7 statements with the appropriate filepaths */
/* But you do not have to fill in the LIBNAME LIBRARY and FILENAME FORMATIN
statements: that is for another input path */
libname In2015 "U:\FILEPATH";
libname In2016 "U:\ FILEPATH";
libname In2017 "U:\ FILEPATH";
libname In2018 "U:\ FILEPATH";
libname In2019 "U:\ FILEPATH";

/* Now, we read in each year and keep the necessary BRFSS variables */
/* 2015 */
data BRFSS2015;
    set In2015.sasdata;
    keep _STATE SEX AGE _AGE80 _AGEG5YR ALCDAY5 AVEDRNK2 DRNK3GE5 MAXDRNKS
    _LLCPWT DRNKANY5 DROCDY3 _ _RFBING5 _DRNKDY4 _DRNKWEK _DRNKMO4 _RFDRHV4
    _RFDRHV5;
run;

proc sql;
    create table Alcohol_2015 as
    select "2015" as Year, _STATE, SEX as Sex, _AGE80 as Age80, DRNKANY5,
    ALCDAY5, AVEDRNK2, DRNK3GE5, MAXDRNKS, _LLCPWT
    from BRFSS2015
    ;
quit;

/* 2016 */
data BRFSS2016;
    set In2016.sasdata;
    keep _STATE SEX AGE _AGE80 _AGEG5YR ALCDAY5 AVEDRNK2 DRNK3GE5 MAXDRNKS
    _LLCPWT DRNKANY5 DROCDY3 _ _RFBING5 _DRNKDY4 _DRNKWEK _DRNKMO4 _RFDRHV4
    _RFDRHV5;
run;

proc sql;
    create table Alcohol_2016 as
    select "2016 " as Year, _STATE, SEX as Sex, _AGE80 as Age80, DRNKANY5,
    ALCDAY5, AVEDRNK2, DRNK3GE5, MAXDRNKS, _LLCPWT
    from BRFSS2016
    ;
quit;

/* 2017 */
data BRFSS2017;
    set In2017.sasdata;
    keep _STATE SEX AGE _AGE80 _AGEG5YR ALCDAY5 AVEDRNK2 DRNK3GE5 MAXDRNKS
    _LLCPWT DRNKANY5 DROCDY3 _ _RFBING5 _DRNKDY4 _DRNKWEK _DRNKMO4 _RFDRHV4
    _RFDRHV5;
run;

proc sql;
    create table Alcohol_2017 as
    select "2017" as Year, _STATE, SEX as Sex, _AGE80 as Age80, DRNKANY5,
    ALCDAY5, AVEDRNK2, DRNK3GE5, MAXDRNKS, _LLCPWT
    from BRFSS2017
    ;
quit;

```

```

/* 2018 */
data BRFSS2018;
  set In2018.sasdata;
  keep _STATE SEX1 AGE _AGE80 _AGEG5YR ALCDAY5 AVEDRNK2 DRNK3GE5 MAXDRNKS
  _LLCPWT DRNKANY5 DROCDY3 _RFBING5 _DRNKDY4 _DRNKWEK _DRNKMO4 _RFDRHV4
  _RFDRHV5;
run;

proc sql;
  create table Alcohol_2018 as
  select "2018" as Year, _STATE, SEX1 as Sex, _AGE80 as Age80, DRNKANY5,
  ALCDAY5, AVEDRNK2, DRNK3GE5, MAXDRNKS, _LLCPWT
  from BRFSS2018
  ;
quit;

/* 2019 */
/* In 2019, the variable AVEDRNK2 changed to AVEDRNK3. Rename to match the
other years for the purpose of this analysis */
data BRFSS2019;
  set In2019.sasdata;
  keep _STATE SEXVAR AGE _AGE80 _AGEG5YR ALCDAY5 AVEDRNK3 DRNK3GE5 MAXDRNKS
  _LLCPWT DRNKANY5 DROCDY3 _RFBING5 _DRNKDY4 _DRNKWEK _DRNKMO4 _RFDRHV4
  _RFDRHV5;
run;

proc sql;
  create table Alcohol_2019 as
  select "2019" as Year, _STATE, SEXVAR as Sex, _AGE80 as Age80, DRNKANY5,
  ALCDAY5, AVEDRNK3 as AVEDRNK2, DRNK3GE5, MAXDRNKS, _LLCPWT
  from BRFSS2019
  ;
quit;

/* Merge the five years together into one datafile and output for further
analysis */
data BRFSS_Alcohol_2015_to_2019;
  set Alcohol_2015 Alcohol_2016 Alcohol_2017 Alcohol_2018 Alcohol_2019;
run;

libname out "U:\FILEPATH";
data out.BRFSS_Alcohol_2015_to_2019;
  set BRFSS_Alcohol_2015_to_2019;
run;

```

Step 2. Calculate standard drinks per day adjusting to 73% of per capita alcohol sales

```

/*****
*****/
/* This step will create a final standard drinks per day file by applying the
ARDI method of adjusting to 73% of per capita alcohol sales) */
/*****
*****/
libname ardiin 'U:\FILEPATH';

```

```

/* CHOOSE YOUR STATE FOR THE STATE-SPECIFIC ANALYSIS */
%let STATE = ALABAMA;

/* Read in the combined BRFSS 2015-2019 file created in Step 1 */
/* This is a stacked datafile including the responses from all five years */

/* This step removes Guam(66) and PR(72) and Virgin Island (78) & further
names the states so it does not have to be done later */
data ARDI;
    set ardiin.brfss_alcohol_2015_to_2019;
    keep year State _state sex age80 ALCDAY5 AVEDRNK2 _LLCPWT;
    where (_state not in (66 72 78));

    LENGTH state $30;
    if _state = 00 then state = 'zUNITED STATES';
    else if _state = 01 then state = 'ALABAMA';
    else if _state = 02 then state = 'ALASKA';
    else if _state = 05 then state = 'ARKANSAS';
    else if _state = 60 then state = 'AMERICAN SAMOA';
    else if _state = 04 then state = 'ARIZONA';
    else if _state = 06 then state = 'CALIFORNIA';
    else if _state = 08 then state = 'COLORADO';
    else if _state = 09 then state = 'CONNECTICUT';
    else if _state = 11 then state = 'DISTRICT OF COLUMBIA';
    else if _state = 10 then state = 'DELAWARE';
    else if _state = 12 then state = 'FLORIDA';
    else if _state = 13 then state = 'GEORGIA';
    else if _state = 66 then state = 'GUAM';
    else if _state = 15 then state = 'HAWAII';
    else if _state = 19 then state = 'IOWA';
    else if _state = 16 then state = 'IDAHO';
    else if _state = 17 then state = 'ILLINOIS';
    else if _state = 18 then state = 'INDIANA';
    else if _state = 20 then state = 'KANSAS';
    else if _state = 21 then state = 'KENTUCKY';
    else if _state = 22 then state = 'LOUISIANA';
    else if _state = 25 then state = 'MASSACHUSETTS';
    else if _state = 24 then state = 'MARYLAND';
    else if _state = 23 then state = 'MAINE';
    else if _state = 26 then state = 'MICHIGAN';
    else if _state = 27 then state = 'MINNESOTA';
    else if _state = 29 then state = 'MISSOURI';
    else if _state = 28 then state = 'MISSISSIPPI';
    else if _state = 30 then state = 'MONTANA';
    else if _state = 37 then state = 'NORTH CAROLINA';
    else if _state = 38 then state = 'NORTH DAKOTA';
    else if _state = 31 then state = 'NEBRASKA';
    else if _state = 33 then state = 'NEW HAMPSHIRE';
    else if _state = 34 then state = 'NEW JERSEY';
    else if _state = 35 then state = 'NEW MEXICO';
    else if _state = 32 then state = 'NEVADA';
    else if _state = 36 then state = 'NEW YORK';
    else if _state = 39 then state = 'OHIO';
    else if _state = 40 then state = 'OKLAHOMA';
    else if _state = 41 then state = 'OREGON';
    else if _state = 42 then state = 'PENNSYLVANIA';
    else if _state = 72 then state = 'PUERTO RICO';
    else if _state = 44 then state = 'RHODE ISLAND';

```

```

else if _state = 45 then state = 'SOUTH CAROLINA';
else if _state = 46 then state = 'SOUTH DAKOTA';
else if _state = 47 then state = 'TENNESSEE';
else if _state = 48 then state = 'TEXAS';
else if _state = 49 then state = 'UTAH';
else if _state = 51 then state = 'VIRGINIA';
else if _state = 78 then state = 'VIRGIN ISLANDS';
else if _state = 50 then state = 'VERMONT';
else if _state = 53 then state = 'WASHINGTON';
else if _state = 55 then state = 'WISCONSIN';
else if _state = 54 then state = 'WEST VIRGINIA';
else if _state = 56 then state = 'WYOMING';
run;

data ARDI_State;
  set ARDI;
  if state = "&STATE.";
run;

/*****
*****/
/* New Jersey was unable to collect enough BRFSS data in 2019 to meet the
minimum requirements for inclusion in the 2019 annual aggregate data set. The
following steps address this issue by combining the other four years and
reweighting by the NJ 2019 population aged 18 years and older*/
/* This section of the SAS program only needs to be run if you are completing
a New Jersey-based analysis. However, it can be adapted if a year of data is
missing for your state analysis. */
/*****
*****/;
/*data NJ;*/
/*  set ARDI_State;*/
/*  if State="NEW JERSEY";*/
/*run;*/
/**/
/*proc freq data = NJ;*/
/*  tables Year / out=NJ_Year;*/
/*  weight _LLCPWT;*/
/*run;*/
/**/
/*/* From the US Census Bureau, the estimate for the 18+ population of New
Jersey in 2019 is 6,945,873 */*/
/*/* This sql re-weights the NJ2015-18 sample to impute a sample for 2019
*/*/
/*proc sql;*/
/*  create table NJ2 as*/
/*  select "2019" as Year, *, _LLCPWT * 6945873/sum(_LLCPWT) as New_Weight*/
/*  from NJ*/
/*  ;*/
/*quit;*/
/**/
/*/* Now rename and add this imputed dataset to the five year dataset */*/
/*data NJ3;*/
/*  set NJ2;*/
/*  drop _LLCPWT;*/
/*  rename New_Weight=_LLCPWT;*/
/*run;*/
/**/

```

```

/*data ARDI_State;*/
/* set ARDI_State NJ3;*/
/*run;*/
/* End New Jersey specific coding */

/*****
*****/
/* Create dataset in order to calculate adjusted drinks per day

*/
/* These analytic steps will remove any respondents who did not answer the
necessary alcohol core questions.
*/
/* Two variables are used:

*/
/*          ALCDAY5: how many days per week or per month did you have at
least one drink
*/
/*          AVEDRNK2 (AVEDRNK3 in 2019): During the past 30 days, on the
days you drank, about how many drinks did you drink on average? */
/*
/*****
*****/

/* A. This data step resets the information in the two alcohol variables, so
it will be easier to process and apply inclusion criteria */
data ARDI_Labeled;
  set ARDI_State;

  *Calculate the proportion of days that are drinking days (Prop_Drink_Days);
  if 101<=ALCDAY5<=107 then do;
    Prop_Drink_Days=(ALCDAY5-100)/7;
  end;
  else if 201<=ALCDAY5<=230 then do;
    Prop_Drink_Days=(ALCDAY5-200)/30;
  end;
  else if ALCDAY5=888 then Prop_Drink_Days=0;
  else Prop_Drink_Days=.;

  *Average quantity per DRINKING DAY (Avg_Quantity) in the past 30 days;
  if 1<=AVEDRNK2<=30 then Avg_Quantity=AVEDRNK2;
  else if 31<=AVEDRNK2<=76 then Avg_Quantity=30; *as directed by CDC, this
variable is capped at 30;
  else if ALCDAY5=888 then Avg_Quantity=0;
  else Avg_Quantity=.;
run;

/* B. Respondents must provide a certain amount of information, otherwise
their consumption cannot be calculated */
/* Here we remove respondents without sufficient information */

data ARDI_Inclusions;
  set ARDI_Labeled;

  if Sex not in (1 2) then delete; *ARDI prevalence is estimated by sex;

```



```

    *Respondent must have a valid answer for both Quantity (AVEDRNK2, or
    AVEDRNK3 in 2019) and Frequency (ALCDAY5) questions;
    if Prop_Drink_Days = . then delete; *if the respondent refuses to answer
    the necessary alcohol core questions, they are not included;

    if Avg_Quantity = . then delete;
run;

/*****
*****/
/* C. The following three data steps will calculate the average Drinks Per
Day for each included respondent
*/
/* This will be done using the ARDI methodology of adjusting to 73% of
national per capita alcohol sales */
/* Defined variables and definitions:*/
/* DPD = Standard drinks per day */
/* DPD_QF - drinks per day, using the survey responses for quantity and
frequency, as is */
/* DPD_73DS - drinks per day, using the adjustment to 73% of national per
capita alcohol sales */
/*****
*****/
data ARDI_DPD;
    set ARDI_Inclusions;

    DPD_QF=Prop_Drink_Days*Avg_Quantity; *QF drinks per day is average quantity
x prop of drinking days with answers as given in the survey;
run;

*this is a reordering for readability and quality assurance;
proc sql;
    create table ARDI_DPD_Reorder as
    select Year, _State, State, Sex, Age80, _LLCPWT, ALCDAY5, AVEDRNK2,
        Avg_Quantity, Prop_Drink_Days, DPD_QF
    from ARDI_DPD
    ;
quit;

/*****
*****/
/* D. Apply the individual-level adjustment factor of 2.3166, which will
adjust the BRFSS to 73% of national per capita alcohol sales
*/
data ARDI_DPD_Final;
    set ARDI_DPD_Reorder;
    DPD_73DS = DPD_QF * 2.3166;
run;

/* At this stage we have a full five-year BRFSS data set with the Drinks Per
Day variable for output */
libname ardiout 'U:\FILEPATH';
data ardiout.&STATE.;
    set ARDI_DPD_Final;
run;

```

Step 3. Create prevalence estimates by level of average daily alcohol use for ARDI

```

/*****
*****/
/* This program will create final ARDI prevalence estimates for the state
chosen in the Step 2 */
/*****
*****/

/*****
*****/
/* Create prevalence estimates for Low, Medium and High alcohol consumption
categories for males and females */
/*****
*****/

/*****
*****/
/* 1. Standardized cut-points */
/* Men: Low >0 to <=2; Medium >2 to <=4; High >4 */
/* Women: Low >0 to <=1; Medium >1 to <=2; High >2 */
/*****
*****/
/* 2. Corrao et al. (1999,2004) cut-points */
/* Men: Low >=0.1 to <1.8; Medium >=1.8 to <3.6; High >=3.6 */
/* Women: Low >=0.1 to <1.8; Medium >=1.8 to <3.6; High >=3.6 */
/*****
*****/
/* 3. English and Holman 1995 cut-points */
/* Men: Low >=0.2 to <2.9; Medium >=2.9 to <4.3; High >=4.3 */
/* Women: Low >=0.2 to <1.4; Medium >=1.4 to <2.9; High >=2.9 */
/*****
*****/
/* 4. Zhao et al. (2017) cut-points */
/* Men: Low >=0.09 to <1.8; Medium >=1.8 to <3.2; High >=3.2 */
/* Women: Low >=0.09 to <1.8; Medium >=1.8 to <3.2; High >=3.2 */
/*****
*****/

*Choose your state, as in Step 2;
%let STATE = ALABAMA;

libname ARDIIN "U:\FILEPATH";

data ARDI_DPD_Final;
    set ARDIIN.&STATE;
run;

/* Delineate which category( Abstainer, Low, Medium, High) a respondent is
in, for each set of cut-points */
data ARDI_Prevalence1;
    set ARDI_DPD_Final;

    /* Standardized cut-points */
    if Sex=1 then do;
        if DPD_73DS=0 then Standard=0;
        else if 0<DPD_73DS<=2 then Standard=1;
        else if 2<DPD_73DS<=4 then Standard=2;
        else if 4<DPD_73DS<=70 then Standard=3;
        else Standard=9;
    end;
    else if Sex=2 then do;
        if DPD_73DS=0 then Standard=0;

```

```

        else if 0<DPD_73DS<=1 then Standard=1;
        else if 1<DPD_73DS<=2 then Standard=2;
        else if 2<DPD_73DS<=70 then Standard=3;
        else Standard=9;
    end;
    /* Corrao cut-points */
    if 0<=DPD_73DS<0.1 then Corrao=0;
    else if 0.1<=DPD_73DS<1.8 then Corrao=1;
    else if 1.8<=DPD_73DS<3.6 then Corrao=2;
    else if 3.6<=DPD_73DS<=70 then Corrao=3;
    else Corrao=9;
    /* English and Holman cut-points */
    if Sex=1 then do;
        if 0<=DPD_73DS<0.2 then English=0;
        else if 0.2<=DPD_73DS<2.9 then English=1;
        else if 2.9<=DPD_73DS<4.3 then English=2;
        else if 4.3<=DPD_73DS<=70 then English=3;
        else English=9;
    end;
    else if Sex=2 then do;
        if 0<=DPD_73DS<0.2 then English=0;
        else if 0.2<=DPD_73DS<1.4 then English=1;
        else if 1.4<=DPD_73DS<2.9 then English=2;
        else if 2.9<=DPD_73DS<=70 then English=3;
        else English=9;
    end;
    /* Zhao cut-points */
    if 0<=DPD_73DS<0.09 then Zhao=0;
    else if 0.09<=DPD_73DS<1.8 then Zhao=1;
    else if 1.8<=DPD_73DS<3.2 then Zhao=2;
    else if 3.2<=DPD_73DS<=70 then Zhao=3;
    else Zhao=9;
run;

/* Sort for the subsequent proc freqs */
proc sort data=ARDI_Prevalence1;
    by State _State;
run;

/* Generate weighted prevalence estimates for Abstainers, L, M, H by state
for each sex and cut-point grouping */
/* A. Standard cut-points, men, 20+, for most chronic conditions */
proc freq data=ARDI_Prevalence1;
    by state _state;
    tables Standard/out=Standard_Men_20up_State;
    weight _LLCPWT;
    where sex=1 and Age80>=20;
run;
data Standard_Men_20up;
    retain state _state Drinking_Group Standard;
    set Standard_Men_20up_State;
    if Standard=0 then Drinking_Group="Abstainers";
    else if Standard=1 then Drinking_Group="Low";
    else if Standard=2 then Drinking_Group="Medium";
    else if Standard=3 then Drinking_Group="High";
    drop COUNT;
    rename Percent=Standard_Men_20up Standard=Group;
run;

```

```

/* B. Standard cut-points, women, 20+, for most chronic conditions */
proc freq data=ARDI_Prevalence1;
  by state _state;
  tables Standard/out=Standard_Women_20up_State;
  weight _LLCPWT;
  where sex=2 and Age80>=20;
run;
data Standard_Women_20up;
  retain state _state Drinking_Group Standard;
  set Standard_Women_20up_State;
  if Standard=0 then Drinking_Group="Abstainers";
  else if Standard=1 then Drinking_Group="Low";
  else if Standard=2 then Drinking_Group="Medium";
  else if Standard=3 then Drinking_Group="High";
  drop COUNT;
  rename Percent=Standard_Women_20up Standard=Group;
run;

/* C. Standard cut-points, men, 20-64, for pneumonia ONLY */
proc freq data=ARDI_Prevalence1;
  by state _state;
  tables Standard/out=Standard_Men_2064_State;
  weight _LLCPWT;
  where sex=1 and 20<=Age80<=64;
run;
data Standard_Men_2064;
  retain state _state Drinking_Group Standard;
  set Standard_Men_2064_State;
  if Standard=0 then Drinking_Group="Abstainers";
  else if Standard=1 then Drinking_Group="Low";
  else if Standard=2 then Drinking_Group="Medium";
  else if Standard=3 then Drinking_Group="High";
  drop COUNT;
  rename Percent=Standard_Men_2064 Standard=Group;
run;

/* D. Standard cut-points, women, 20-64, for pneumonia ONLY */
proc freq data=ARDI_Prevalence1;
  by state _state;
  tables Standard/out=Standard_Women_2064_State;
  weight _LLCPWT;
  where sex=2 and 20<=Age80<=64;
run;
data Standard_Women_2064;
  retain state _state Drinking_Group Standard;
  set Standard_Women_2064_State;
  if Standard=0 then Drinking_Group="Abstainers";
  else if Standard=1 then Drinking_Group="Low";
  else if Standard=2 then Drinking_Group="Medium";
  else if Standard=3 then Drinking_Group="High";
  drop COUNT;
  rename Percent=Standard_Women_2064 Standard=Group;
run;

/* E. Standard cut-points, women, 18-44, for infant death conditions (INFLBW,
INFPRE, INFSGA) ONLY */
proc freq data=ARDI_Prevalence1;

```

```

by state _state;
tables Standard/out=Standard_Women_1844_State;
weight _LLCPWT;
where sex=2 and 18<=Age80<=44;
run;
data Standard_Women_1844;
retain state _state Drinking_Group Standard;
set Standard_Women_1844_State;
if Standard=0 then Drinking_Group="Abstainers";
else if Standard=1 then Drinking_Group="Low";
else if Standard=2 then Drinking_Group="Medium";
else if Standard=3 then Drinking_Group="High";
drop COUNT;
rename Percent=Standard_Women_1844 Standard=Group;
run;

/* F. Corrao cut-points, men, 20+, for chronic pancreatitis and chronic
hepatitis */
proc freq data=ARDI_Prevalence1;
by state _state;
tables Corrao/out=Corrao_Men_State;
weight _LLCPWT;
where sex=1 and Age80>=20;
run;
data Corrao_Men;
retain state _state Drinking_Group Corrao;
set Corrao_Men_State;
if Corrao=0 then Drinking_Group="Abstainers";
else if Corrao=1 then Drinking_Group="Low";
else if Corrao=2 then Drinking_Group="Medium";
else if Corrao=3 then Drinking_Group="High";
drop COUNT;
rename Percent=Corrao_Men Corrao=Group;
run;

/* G. Corrao cut-points, women, 20+, for chronic pancreatitis and chronic
hepatitis */
proc freq data=ARDI_Prevalence1;
by state _state;
tables Corrao/out=Corrao_Women_State;
weight _LLCPWT;
where sex=2 and Age80>=20;
run;
data Corrao_Women;
retain state _state Drinking_Group Corrao;
set Corrao_Women_State;
if Corrao=0 then Drinking_Group="Abstainers";
else if Corrao=1 then Drinking_Group="Low";
else if Corrao=2 then Drinking_Group="Medium";
else if Corrao=3 then Drinking_Group="High";
drop COUNT;
rename Percent=Corrao_Women Corrao=Group;
run;

/* H. English and Holman cut-points, men, 20+, for gallbladder disease */
proc freq data=ARDI_Prevalence1;
by state _state;
tables English/out=English_Men_State;

```

```

weight _LLCPWT;
where sex=1 and Age80>=20;
run;
data English_Men;
retain state _state Drinking_Group English;
set English_Men_State;
if English=0 then Drinking_Group="Abstainers";
else if English=1 then Drinking_Group="Low";
else if English=2 then Drinking_Group="Medium";
else if English=3 then Drinking_Group="High";
drop COUNT;
rename Percent=English_Men English=Group;
run;

/* I. English and Holman cut-points, women, 20+, for gallbladder disease */
proc freq data=ARDI_Prevalence1;
by state _state;
tables English/out=English_Women_State;
weight _LLCPWT;
where sex=2 and Age80>=20;
run;
data English_Women;
retain state _state Drinking_Group English;
set English_Women_State;
if English=0 then Drinking_Group="Abstainers";
else if English=1 then Drinking_Group="Low";
else if English=2 then Drinking_Group="Medium";
else if English=3 then Drinking_Group="High";
drop COUNT;
rename Percent=English_Women English=Group;
run;

/* J. Zhao cut-points, men, 20+, for coronary heart disease */
proc freq data=ARDI_Prevalence1;
by state _state;
tables Zhao/out=Zhao_Men_State;
weight _LLCPWT;
where sex=1 and Age80>=20;
run;
data Zhao_Men;
retain state _state Drinking_Group Zhao;
set Zhao_Men_State;
if Zhao=0 then Drinking_Group="Abstainers";
else if Zhao=1 then Drinking_Group="Low";
else if Zhao=2 then Drinking_Group="Medium";
else if Zhao=3 then Drinking_Group="High";
drop COUNT;
rename Percent=Zhao_Men Zhao=Group;
run;

/* K. Zhao cut-points, women, 20+, for coronary heart disease */
proc freq data=ARDI_Prevalence1;
by state _state;
tables Zhao/out=Zhao_Women_State;
weight _LLCPWT;
where sex=2 and Age80>=20;
run;
data Zhao_Women;

```

```

retain state _state Drinking_Group Zhao;
set Zhao_Women_State;
if Zhao=0 then Drinking_Group="Abstainers";
else if Zhao=1 then Drinking_Group="Low";
else if Zhao=2 then Drinking_Group="Medium";
else if Zhao=3 then Drinking_Group="High";
drop COUNT;
rename Percent=Zhao_Women Zhao=Group;
run;

/* Merge these 11 files into one master prevalence file and convert into
prevalence estimates instead of percentages */
data ARDI_prevalence;
merge Standard_Men_20up Standard_Women_20up
      Standard_Men_2064 Standard_Women_2064
      Standard_Women_1844
      Corrao_Men Corrao_Women
      English_Men English_Women
      Zhao_Men Zhao_Women;
by _state Group;
Standard_Men_20up=Standard_Men_20up/100;
Standard_Women_20up=Standard_Women_20up/100;
Standard_Men_2064=Standard_Men_2064/100;
Standard_Women_2064=Standard_Women_2064/100;
Standard_Women_1844=Standard_Women_1844/100;
Corrao_Men=Corrao_Men/100;
Corrao_Women=Corrao_Women/100;
English_Men=English_Men/100;
English_Women=English_Women/100;
Zhao_Men=Zhao_Men/100;
Zhao_Women=Zhao_Women/100;
run;

*Reorient the file so all prevalences for one state are on the same
line;%macro Transpose_Prevalences(PVarName,Sex,Cutpoint);
data Prev_temp1;
set ARDI_prevalence;
keep State _State Drinking_Group &PVarName.;
run;
proc sort data=Prev_temp1; by State _State Drinking_Group; run;
proc transpose data=Prev_temp1 out=Prev_temp2(drop=_NAME_ _LABEL_);
by State _State;
var &PVarName.;
id Drinking_Group;
run;
proc sql;
create table P_&Sex._&Cutpoint. as
select State, _State,
/*      "&Sex." as Sex, */
      Abstainers as P_Abs_&Sex._&Cutpoint.,
      Low as P_Low_&Sex._&Cutpoint., Medium
as P_Med_&Sex._&Cutpoint., High as
      P_Hi_&Sex._&Cutpoint.
from Prev_temp2
group by _State
order by _State
;
quit;

```

```

proc datasets library=work; delete Prev_temp;; run; quit;
%mend;
%Transpose_Prevalences(Standard_Men_20up, Men, St20p);
%Transpose_Prevalences(Standard_Women_20up, Women, St20p);
%Transpose_Prevalences(Standard_Men_2064, Men, St2064);
%Transpose_Prevalences(Standard_Women_2064, Women, St2064);
%Transpose_Prevalences(Standard_Women_1844, Women, St1844);
%Transpose_Prevalences(Corrao_Men, Men, Corrao);
%Transpose_Prevalences(Corrao_Women, Women, Corrao);
%Transpose_Prevalences(English_Men, Men, English);
%Transpose_Prevalences(English_Women, Women, English);
%Transpose_Prevalences(Zhao_Men, Men, Zhao);
%Transpose_Prevalences(Zhao_Women, Women, Zhao);

data Standardized;
  merge P_men_st20p P_women_st20p;
  by _state;
  Prevalence="Standardized";
  rename P_Low_Men_St20p=P_Low_Men;
  rename P_Med_Men_St20p=P_Med_Men;
  rename P_Hi_Men_St20p=P_Hi_Men;
  rename P_Low_Women_St20p=P_Low_Women;
  rename P_Med_Women_St20p=P_Med_Women;
  rename P_Hi_Women_St20p=P_Hi_Women;
run;

data Standardized2064;
  merge P_men_st2064 P_women_st2064;
  by _state;
  Prevalence="Standardized2064";
  rename P_Low_Men_St2064=P_Low_Men;
  rename P_Med_Men_St2064=P_Med_Men;
  rename P_Hi_Men_St2064=P_Hi_Men;
  rename P_Low_Women_St2064=P_Low_Women;
  rename P_Med_Women_St2064=P_Med_Women;
  rename P_Hi_Women_St2064=P_Hi_Women;

data temp;
  set P_women_st1844;

data Corrao;
  merge P_men_corrao P_women_corrao;
  by _state;
  Prevalence="Corrao";
  rename P_Low_Men_Corrao=P_Low_Men;
  rename P_Med_Men_Corrao=P_Med_Men;
  rename P_Hi_Men_Corrao=P_Hi_Men;
  rename P_Low_Women_Corrao=P_Low_Women;
  rename P_Med_Women_Corrao=P_Med_Women;
  rename P_Hi_Women_Corrao=P_Hi_Women;

data English;
  merge P_men_english P_women_english;
  by _state;
  Prevalence="English";
  rename P_Low_Men_English=P_Low_Men;
  rename P_Med_Men_English=P_Med_Men;
  rename P_Hi_Men_English=P_Hi_Men;

```



```

rename P_Low_Women_English=P_Low_Women;
rename P_Med_Women_English=P_Med_Women;
rename P_Hi_Women_English=P_Hi_Women;

data Zhao;
merge P_men_zhao P_women_zhao;
by _state;
Prevalence="Zhao";
rename P_Low_Men_Zhao=P_Low_Men;
rename P_Med_Men_Zhao=P_Med_Men;
rename P_Hi_Men_Zhao=P_Hi_Men;
rename P_Low_Women_Zhao=P_Low_Women;
rename P_Med_Women_Zhao=P_Med_Women;
rename P_Hi_Women_Zhao=P_Hi_Women;
run;

proc sql;
create table Standardized1844 as
select state, _state, "Standardized1844" as Prevalence,
       P_Low_Women_St1844 as P_Low_Men,
       P_Med_Women_St1844 as P_Med_Men,
       P_Hi_Women_St1844 as P_Hi_Men,
       P_Low_Women_St1844 as P_Low_Women,
       P_Med_Women_St1844 as P_Med_Women,
       P_Hi_Women_St1844 as P_Hi_Women
from temp;
quit;

data Prev_All;
retain State _State Prevalence P_Low_Men P_Med_Men P_Hi_Men P_Low_Women
P_Med_Women P_Hi_Women;
set Standardized1844 Standardized Standardized2064 Corrao English Zhao;
keep State _State Prevalence P_Low_Men P_Med_Men P_Hi_Men P_Low_Women
P_Med_Women P_Hi_Women;
run;

proc sort data=Prev_All;
by _State Prevalence;
run;

proc sql;
create table final_prevalence as
select State,
       "2015-2019" as Year,
       Prevalence,
       P_Low_Men as P_Low_Men format 4.2,
       P_Med_Men as P_Med_Men format 4.2,
       P_Hi_Men as P_Hi_Men format 4.2,
       P_Low_Women as P_Low_Women format 4.2,
       P_Med_Women as P_Med_Women format 4.2,
       P_Hi_Women as P_Hi_Women format 4.2
from Prev_All;
quit;

/* Code to output the final state-specific prevalence file as either an CSV
file or a TXT file */

/**/

```

```
/*proc export data=final_prevalence*/  
/*outfile="U:\FILEPATH.csv"*/  
/*          replace;*/  
/*run;*/  
/**/  
/*data _NULL_;*/  
/*  set final_prevalence;*/  
/*  file "U:\FILEPATH.txt" dlm=",";*/  
/*  put (_All_) (~);*/  
/*run;*/  
/**/  
/**/
```