

# **Alcohol Related Disease Impact (ARDI) Custom Data User Manual**

Updated: February 5, 2014

# Table of Contents

Introduction.....	3
SECTION I: Data Sets .....	4
Total Deaths .....	4
Prevalence of Alcohol Consumption .....	10
Relative Risks and Direct Alcohol-Attributable Fractions .....	14
SECTION II: Reports .....	15
Alcohol-Attributable Deaths.....	15
Years of Potential Life Lost.....	15
Appendix A: Supporting Documents.....	17
I. Understanding Report Formulas.....	18
II. ARDI ICD Codes .....	20
III. Cause of Death Abbreviations and Age Groups .....	23
IV. Locations for Custom Data File.....	25
V. Indexing Example .....	26
APPENDIX B: Sample SAS Codes.....	27
I. Age Group Codes for SAS XML .....	28
II. Age Group Conversion.....	29
III. Total Deaths by State .....	30
IV. Calculation of Average Daily Alcohol Consumption Using BRFSS .....	43

## 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 alcohol-related harms are available for the United States overall and individually for all 50 states and the District of Columbia. Currently, there are 54 causes of death for which ARDI estimates alcohol-related harms and results are available by sex and age groups.

In addition to the public use site of ARDI found at <https://apps.nccd.cdc.gov/ardi/Hompepage.aspx>, ARDI has a built-in capability to allow users to create their own estimates of alcohol-related harms with the appropriate data sources. Custom data reports can 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.

This user manual is designed to help guide you through the necessary steps to create the data sets that are used to perform custom data analyses in the ARDI application. If you need additional technical assistance, please use the "Contact Us About ARDI" page ([http://apps.nccd.cdc.gov/DACH\\_ARDI/Info/ContactUs.aspx](http://apps.nccd.cdc.gov/DACH_ARDI/Info/ContactUs.aspx)) on any of the ARDI Web site pages and send us your question.

## **SECTION I: Data Sets**

Three types of data sets are used in the ARDI application:

- **Total Deaths**
- **Prevalence of Alcohol Consumption**
- **Relative Risks and Direct Alcohol-Attributable Fractions**

### ***Total Deaths***

The total deaths data set should include those deaths with an underlying cause that corresponds to one of the 54 conditions specified in ARDI application. The location where the death occurred and the year of death should also correspond to the location and time period you to assess using ARDI. For example, if you want to estimate the AAD for Atlanta, Georgia, for 2005-2009 then the death data should be for Atlanta, Georgia, for 2005-2009. However, ARDI does not allow users to enter data on causes of death other than the 54 conditions that are currently included in the application.

ARDI is primarily used to calculate the average number of AAD and YPLL due to excessive alcohol use over multiple years. Calculating average 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. It is preferable to calculate the average number of deaths by condition over a five-year period, but if five years of data are not available, the average number of deaths by condition over three years could be calculated instead. ARDI can also be used to calculate AAD and YPLL over a shorter period of time (e.g., one year), but we would recommend against it for the reasons noted.

### **How to Create a Total Deaths Data Set**

#### **Step 1: Acquire death files**

Death or mortality data are collected by states from death certificates. Annually, this information is sent to the National Center for Health Statistics (NCHS) and is available on the Web site at [http://www.cdc.gov/nchs/data\\_access/Vitalstatsonline.htm](http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm). However, beginning in 2005, users are no longer able to obtain state or local death data from the NCHS Web site. Therefore, to obtain these state or local death data, users will need to contact the government agencies in their state that manages vital statistics and request this information for the years they wish to analyze in ARDI.

ARDI was primarily designed to assess deaths and YPLL due to excessive alcohol consumption in the United States. The application can also be used to estimate AAD and YPLL for locations outside of the United States, but the data required to do these analyses (e.g., death data) will need

to be obtained from other sources. It should be noted that all naming conventions for data files and formats are based on locations within the United States. Any data files created for locations outside of the United States will still need to adhere to the formats and naming conventions within the application.

## **Step 2: Creating total death data files**

Total death data files must contain the following information: cause of death, age of decedent, sex of decedent, and the location where the death was reported.

### **Requirements:**

1. The data must be aggregated to the population level for the location specified (i.e., no individual death data are uploaded into ARDI).
2. The underlying cause of death code ONLY is used.\*
3. The *International Classification of Disease Tenth Edition (ICD-10)* codes are used to identify the 54 causes of death included in ARDI. **See Appendix A II: ARDI ICD Codes.**
4. Each cause of death is represented by a special six-digit code in ARDI. **See Appendix A III: Cause of Death Abbreviations and Age Groups.**
5. The death data must be stratified by sex and five-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. **See Appendix A III: Cause of Death Abbreviations and Age Groups.**
6. The total deaths data file will need a data set name and a location. The data set name can be anything you specify, but the location name must be either the United States, one of the 50 states, or DC. **See Appendix A IV: Locations for Custom Data File.**
7. The data files need to be created in one of two format structures: Extensible Markup Language (XML) or Comma Separated Value (.csv) files.

\*On many death certificates a primary or underlying cause of death is listed in addition to contributing or secondary causes of death. The AAF in ARDI are based upon population estimates of the contribution of alcohol to the primary cause of death. Therefore, ARDI only includes those deaths where the underlying or primary cause of death corresponds to one of the 54 alcohol-related causes included in the application.

## **Step 3: Uploading total death data files into the ARDI application**

### ***OPTION 1: XML instructions***

Use this utility to upload a total deaths data set for a specified location using an Extensible Markup Language (XML) file generated by SAS or other program. Data sets must meet the minimum requirements listed below. Once data sets are in the correct XML format, use the upload utility to import data into ARDI.

**Requirements:**

1. Include only one value for each field in a data set.
2. Use only one data set per custom analyses (one data set can be used to generate multiple reports).
3. The fields shown in the table below must be included exactly as shown in the data set for it to upload correctly.
4. File size must be < 4 MB. File sizes larger than 4 MB will cause the upload to fail. (**Note:** *total deaths* files may be too large).
5. The output must be in XML format.
6. See **Appendix B** for the following SAS conversion programs:
  - a. **Appendix B II: Age Group Conversion:** Converts to the correct age groups (**Note:** Run this program first).
  - b. **Appendix B III: Total Deaths Example:** Generates total deaths data set tagged with XML tags.

Field	Description
ALCXML	Field name required by ARDI at the opening and closing of each record to identify the record as an XML file. Always use "ALCXML"
datasetname	Name of the data set (e.g., county name). There is no restriction on the number of characters used in this field. If you are loading your version of state data, put the state name in the "datasetname" field as well as in the "state" field.
state	State name. For programming reasons, this name must match one of the locations listed in ARDI. <b>See Appendix IV: <u>Locations for Custom Data File</u></b>
agegrp	Age group for each cause of death. Each age group in ARDI is designated with a special letter. <b>See Appendix B I: <u>Age Group Codes for SAS XML</u></b> .
cause	Six-digit code included in ARDI to represent each cause of death.
Male	Deaths for males.
Female	Deaths for females.

**Note:** All fields are required.

**Below are examples of how the XML output file generated by SAS or other programs must look in order to upload correctly into the Custom Data Section of the ARDI application.**

***Example 1: Drownings in Atlanta, GA***

```
<ALCXML>  
<datasetname>Atlanta</datasetname>  
<state>Georgia</state>  
<agegrp>AgeE</agegrp>  
<cause>DROWNI</cause>  
<Male>19</Male>  
<Female>1</Female>  
</ALCXML>
```

**Where:**

- ALCXML marks the start and end of each record in XML format;

- Atlanta is the sub-state level;
- Georgia is the state name;
- DROWNI is the cause of death abbreviation for drowning;
- AgeE is the age group;
- 19 is the number of male deaths; and
- 1 is the number of female deaths.

If a cause of death is sex-specific (e.g., breast cancer, prostate cancer, or spontaneous abortion), place a zero in the field that pertains to the sex that is **not** affected by the cause of death. **Do not leave any fields blank.**

**Note:** Male breast cancer is not included due to limited available science at the time of ARDI development.

***Example 2: Prostate Cancer in Atlanta, GA***

```
<ALCXML>
<datasetname>Atlanta</datasetname>
<state>Georgia</state>
<agegrp>AgeK</agegrp>
<cause>PROCAN</cause>
<Male>3</Male>
<Female>0</Female>
</ALCXML>
```

**Where:**

- ALCXML marks the start and end of each record in XML format;
- Atlanta is the sub-state level;
- Georgia is the state name;
- PROCAN is the cause of death abbreviation for prostate cancer;
- AgeK is the age group;
- 3 is the number of male deaths; and
- 0 is the number of female deaths.

***Uploading Custom Data Files***

**Follow these instructions to upload a file:**

1. Name the XML data set file using the following guide:
  - a. Range of years the data set is covering (example: 2006-2010)
  - b. State for which the data set covers (example: Georgia)
    - i. See **Appendix A IV: Locations for Custom Data File**
  - c. Current year the data set is uploaded (example: 2011)
  - d. The type of data set it is (example: Mortality)

The data set file name should include all the parts listed in #1 with the name concatenated, using an underscore (“\_”) as the separator between parts.

**Examples:** 2006-2010\_Georgia\_2011\_Mortality.xml

2. Enter a **Year** to represent the data set (can be any year).
3. Select the file to upload by clicking the **Browse** button.
4. After selecting the file, click the **Upload** button to upload the data.
5. If the upload is successful, the Custom Data home page will include the new data set in the Total Deaths Data Sets list.

**Notes:**

- If you receive an error message, check your file against the examples and correct any differences.
- File sizes larger than 4 MB will cause the upload to fail. Ensure that file size is limited to <4 MB.
- Make sure to upload a file that already has averaged the deaths for the number of years of data that you are using (e.g. 5 years). ARDI will take the given data in the deaths file and then apply AAF.
- The **Year** entered may not be outwardly visible in your uploaded file or output.
- If you upload multiple versions of the same data set, you may want to delete files that you no longer need. There is no date/time stamp on the uploaded file so it may not be possible to differentiate between files with the same State name. If you need to keep multiple files with the same name, check to assure that you are analyzing the desired file.

### ***OPTION 2: .CSV instructions***

Use this utility to upload total deaths data using a comma-separated-value (.csv) file. Each row must contain the following columns of data (**Note:** No column headings should be included in the data set):

- cause of death abbreviations
- age group
- male deaths
- female deaths

**Notes:**

- Even if there are zero records for age 85+, make sure to keep the age range as 80-85, as opposed to 80-84.
- If you upload multiple versions of the same data set, you may want to delete files that you no longer need. There is no date/time stamp on the uploaded file so it may not be possible to differentiate between files with the same State name. If you need to keep multiple files with the same name, check to assure that you are analyzing the desired file.

### ***Example 1: Drowning***

DROWN1, 15-19, 231, 123

DROWN1, 20-24, 314, 101

#### **Where:**

- DROWN1 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; and
- 123 and 101 are the number of female deaths.

If a cause of death is sex-specific (e.g., breast cancer, prostate cancer, or spontaneous abortion), put zeros in the column that pertains to the sex that is not affected by the cause of death. **Do not leave any columns blank.**

**Note:** Male breast cancer is not included due to limited available science at the time of ARDI development.

### ***Example 2: Prostate Cancer***

PROCAN, 70-75, 89, 0

#### **Where:**

- PROCAN is the cause of death abbreviation for prostate cancer;
- 70-75 is the age group;
- 89 is the number of male deaths; and
- 0 is the number of female deaths.

Upload the files by browsing to the file and hitting the button on the page.

Be sure to name the CSV file using the following guide:

- a. Range of years the data set is covering (example: 2006-2010)
- b. State for which the data set covers (example: Georgia)
  - i. See **Appendix A IV: Locations for Custom Data File**
- c. Current year the data set is uploaded (example: 2011)
- d. The type of data set it is (example: Mortality)

The data set file name should include all the parts listed above with the name concatenated, using an underscore (“\_”) as the separator between parts.

**Examples:** 2006-2010\_Georgia\_2011\_Mortality.csv

In the “Data Set Properties” box enter the following information:

- Name of the data set file which can contain multiple years (e.g. Name = My Data 2006-2010). You may customize the format of the data set file name in this box only.
- Choose the location that the data covers
- Enter a single year to represent the data

## **Troubleshooting .CSV data uploads with ARDI**

- Make sure conditions ALCFAS, ALCMAU, LBWPID, MTVTRA and CHILAB have correct age groups for <1, 1-4, 5-9, 10-14
- If the file is opened in Excel, date formats take over so 1-4 becomes Jan-4, 5-9 becomes May-9, 10-14 becomes Oct 14. To correct these formats, first 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 then manually change the dates back to the age ranges. Then save the file as a .csv file. After completing this DO NOT reopen the file in Excel as the age ranges will default back to dates.
- No column headings should be included in the data set.

### **Step 4: After uploading total death data files into the ARDI application**

- Wait until you receive an email indicating that your upload was processed. Depending on your email settings, the automated message may be sent directly to your Junk/Spam mailbox.
- Go back to the Custom Data homepage. Your data set should appear in the list in the drop down arrow for Total Deaths.
- In the Generate a Report section, select the associated Prevalence of Alcohol Consumption data set and Total Deaths data set.
- Select which report you want to view (AAD or YPLL).
  - Note: The first line of the report titles will be the same as the text selected in the Prevalence of Alcohol Consumption box, rather than the title of the customized uploaded Total Deaths data set.
- Check to be sure that you are viewing results for the desired data set: Scroll to the bottom of the report and click “Data Sets Used to Generate this Report.”

## ***Prevalence of Alcohol Consumption***

When performing custom analyses, you can either use the current default data set included in the ARDI application, or enter your own prevalence data set. ARDI uses the CDC's Behavioral Risk Factor Surveillance System (BRFSS) to obtain estimates of the prevalence of alcohol consumption; however, there are a variety of other datasets that you can use to obtain prevalence estimates as well. Prevalence estimates should be for adults aged 20 years and older, and should be based on the consumption levels (i.e., low, medium, and high) defined in the application. It is also important to be sure that the years for which prevalence data are used match the same years that the total deaths data set covers. Below is information related to how to calculate average daily alcohol consumption using the BRFSS.

### **Calculating Average Daily Alcohol Consumption Using BRFSS Data**

#### **Step 1: Acquisition of BRFSS data files**

The BRFSS is a state-based system of health surveys collecting information on a variety of health risk behaviors including alcohol use. BRFSS was established in 1984 by the CDC and currently includes all 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Guam. Additional information on the BRFSS is available at: [www.cdc.gov/brfss](http://www.cdc.gov/brfss).

From 1993 to 2000 questions on alcohol consumption were asked biannually. However, since 2001, alcohol consumption data have been collected annually. As with the total deaths data set, we suggest using multiple years of data to calculate the average daily alcohol consumption. This will help to take into account any year-to-year fluctuations in the data that may not be related to actual changes in alcohol consumption.

The BRFSS data contain both state data and, for select years, county level or metropolitan area data. If you are estimating alcohol-attributable deaths below the state level, BRFSS data for the corresponding location should be used as well. If this is not possible, then state level estimates may be used to estimate alcohol consumption levels at the local level, although it is important to recognize that local prevalence estimates may be different than those reported for the state.

#### **Step 2: Calculating average daily alcohol consumption**

Three core alcohol questions are used to estimate average daily alcohol consumption.

- Frequency of alcohol use in the past 30 days
- Average quantity of alcohol consumed on days a person reports drinking
- Binge drinking episodes defined as the number of times in the past 30 days a person consumed four or more drinks per occasion for women or five or more drinks per occasion for men.

Before 2006, BRFSS did not have a sex-specific cutpoint and defined binge drinking as consuming five or more drinks per occasion for both sexes.

To estimate average daily alcohol consumption using the information from BRFSS, a procedure called ‘indexing’ incorporates self-reported information on binge drinking episodes into the calculation of average daily alcohol use. The reason for indexing is that many people will report their most common or modal alcohol consumption patterns instead of a true arithmetic mean which would include episodes of binge drinking. Not including binge drinks in calculations of average daily alcohol consumption results in an underestimate of true alcohol use. For further information on how to use indexing with BRFSS data please refer to **Appendix A V: Indexing Example**.

### **Step 3: Determining prevalence of alcohol consumption by specified cutpoints**

After applying the indexing method separately for males and females, average daily alcohol consumption is compared to cutpoints specified by the meta-analyses used to obtain risk estimates for a given cause of death. The cutpoints are used to calculate the proportion of the population that consumes alcohol at low, medium, or high levels.

The following four meta-analyses were used to obtain relative risk estimates for many of the chronic alcohol-related conditions that are included in ARDI. English & Holman and Ridolfo & Stevenson used slightly different cut-points for defining alcohol consumption levels than did Corrao & Bagnardi.

#### ***Cutpoints used by English et al. & Ridolfo and Stevenson***

Alcohol Consumption Level	Average Drinks Per Day	
	Men	Women
Low*	≥0.2	≥0.2
Medium	≥2.9	≥1.4
High	≥4.3	≥2.9

\*Excludes those who were abstinent or had less than 0.2 drinks/day on average within a 30-day time frame.

#### **References:**

- 1) English DR, Holman CDJ, Milne E, Winter MG, Hulse GK, Codde JP, et al. *The Quantification of Drug Caused Morbidity and Mortality in Australia, 1995 edition*. Canberra, Australia: Commonwealth Department of Human Services and Health; 1995
- 2) Ridolfo B, Stevenson, C. *The Quantification of Drug-Caused Morbidity and Mortality in Australia, 1998*. Drug Statistics Series no. 7. AIHW cat. no. PHE 29. Canberra, Australia: AIHW; 2001.

## *Cutpoints used by Corrao et al. & Bagnardi et al.*

Alcohol Consumption Level	Average Drinks Per Day	
	Men	Women
Low*	≥0.1	≥0.1
Medium	≥1.8	≥1.8
High	≥3.6	≥3.6

\*Excludes those who were abstinent or had <0.1 drinks/day on average within a 30-day time frame. For this set of meta-analyses, the cutpoints for males and females are the same at all alcohol consumption levels.

### **References:**

- 1) Corrao G, Bagnardi V, Zambon A, Arico S. Exploring the dose-response relationship between alcohol consumption and the risk of several alcohol-related conditions: a meta-analysis. *Addiction* 1999;94(10):1551–73.
- 2) Bagnardi V, Blangiardo M, La Vecchia C, Corrao G. A meta-analysis of alcohol drinking and cancer risk. *British Journal of Cancer* 2001;85(11):1700–05.

In ARDI, *excessive alcohol use* is defined as either medium or high average daily alcohol consumption. The default estimates of AAD and YPLL in ARDI are for excessive alcohol use, but the application can also be used to assess these outcomes for any alcohol use (i.e., low, medium, or high) as well.

Sample SAS code is available for calculating the prevalence of average daily alcohol consumption using the BRFSS data. See **Appendix B IV: Calculation of Average Daily Alcohol Consumption using BRFSS**. **Note:** Some of the variable names will be different based on the years of data that are used.

### **Step 4: Entering data into ARDI**

To enter prevalence data into ARDI, click on the “Create New Data Set” at the bottom of the prevalence data set for the location you want. You will then need to specify a name for the data set, location, and year for the prevalence estimates. Enter the prevalence estimates for each of the two sets of prevalence cutpoints and three consumption levels (low, medium, and high) by sex, rounding to the nearest tenth, and then save the data you have entered in the Custom Data section of ARDI.

## ***Relative Risks and Direct Alcohol-Attributable Fractions***

Alcohol-attributable fractions are used 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. The different types of AAF are described in more detail in the ARDI Methods.

The Custom Data Section in ARDI allows users to enter their own directly measured AAF for acute causes of death only. This feature was added to ARDI at the request of users who are interested in using Alcohol-Attributable Fractions (AAF) collected from Medical Examiner (ME) data from the state or county for which they are preparing custom reports.

To enter direct AAF into ARDI, click on the “Create New Data Set” at the bottom of the Relative Risks and Direct AAF data set for the location specified. You will need to specify the name of the data set, location, and year to represent the AAF estimates. You will type the AAF estimates for both males and females for all ages for each acute cause of death. For motor-vehicle traffic crashes, you have the ability to enter age-specific AAF for males and females. Press “Save” at the bottom of the page.

Many users may not have access to such location-specific AAF, in which case the default AAF for acute causes of death that have been entered into the ARDI application should be used for state or local analyses as well.

## SECTION II: Reports

### *Alcohol-Attributable Deaths*

In order to produce an Alcohol-Attributable Deaths Report for a custom location, three data sets need to be chosen:

- Total Deaths
- Prevalence of Alcohol Consumption
- Relative Risks and Direct Alcohol-Attributable Fractions

These data sets will be listed at the bottom of the report view.

From the **Alcohol-Attributable Deaths Report** page, you can:

- Download the data
- Print the results
  - **Note: The year field entered during uploading will not appear in the title of the report.**
- Change the report view to show alcohol-attributable deaths for only those under age 21
- Change the report view to show alcohol-attributable deaths for all levels of alcohol consumption
- 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.

### *Years of Potential Life Lost*

In order to produce a Years of Potential Life Lost Report for a custom location, three data sets need to be chosen:

- Total Deaths
- Prevalence of Alcohol Consumption
- Relative Risks and Direct Alcohol-Attributable Fractions

These data sets will be listed at the bottom of the report view.

In addition, the Life Expectancy Data Set is used to calculate YPLL (See ARDI Methods). The current Life Expectancy Data Set included in ARDI is based on the year 2000 standard US population. The Life Expectancy Data Set cannot be altered in the ARDI Custom Data.

From the **Years of Potential Life Lost Report** page, you can:

- Download the data
- Print the results
  - **Note: The year field entered during uploading will not appear in the title of the report.**
- Change the report view to show Years of Potential Life Lost for only those under age 21
- Change the report view to show Years of Potential Life Lost for all levels of alcohol consumption
- Change the report view to show Years of Potential Life Lost by age groups

# **Appendix A: Supporting Documents**

## I. Understanding Report Formulas

ARDI generates each report using a different formula or methodology. This page provides the formulas for the following reports:

### Alcohol-Attributable Deaths Report

ARDI uses the following formula to generate the Alcohol-Attributable Deaths Report.

$$AAF \times Total\ Deaths$$

where

- *AAF* = alcohol-attributable fraction for the cause of death
- *Total Deaths* = total number of deaths by cause for a given demographic group

Calculated estimates of alcohol-attributable deaths have a precision of one decimal place.

### Years of Potential Life Lost Report

ARDI uses the following formula to generate the Years of Potential Life Lost report.

$$AAF \times Total\ Deaths \times YPLR$$

where

- *AAF* = alcohol-attributable fraction for the cause of death
- *Total Deaths* = total number of deaths by cause for a given demographic group
- *YPLR* = years of potential life remaining for a given demographic group

Calculated estimates of years of potential life lost have a precision of one decimal place.

### Alcohol-Attributable Fractions Report

ARDI generates the Alcohol-Attributable Fractions Report for excessive alcohol use or for any alcohol use.

### AAF Report for Any Alcohol Consumption

ARDI uses the following formula to calculate AAF for any level of alcohol consumption:

$$\frac{p_0(rr_0 - 1)}{1 + p_0(rr_0 - 1)} + \frac{p_1(rr_1 - 1)}{1 + p_1(rr_1 - 1)} + \frac{p_2(rr_2 - 1)}{1 + p_2(rr_2 - 1)}$$

where

- $p_0$  = prevalence of low alcohol consumption
- $p_1$  = prevalence of medium alcohol consumption
- $p_2$  = prevalence of high alcohol consumption
- $rr_0$  = relative risk for low alcohol consumption
- $rr_1$  = relative risk for medium alcohol consumption
- $rr_2$  = relative risk for high alcohol consumption

### **AAF Report for Excessive Alcohol Consumption**

ARDI uses the following formula to calculate AAF for excessive alcohol consumption:

$$\frac{p_1(rs_1 - 1)}{1 + p_1(rs_1 - 1)} + \frac{p_2(rs_2 - 1)}{1 + p_2(rs_2 - 1)}$$

where

$$rs_1 = \frac{rr_1}{rr_0}$$

and

$$rs_2 = \frac{rr_2}{rr_0}$$

and

- $p_1$  = prevalence of medium alcohol consumption
- $p_2$  = prevalence of high alcohol consumption
- $rr_0$  = relative risk for low alcohol consumption
- $rr_1$  = relative risk for medium alcohol consumption
- $rr_2$  = relative risk for high alcohol consumption
- $rs_1$  = rescaled relative risk for medium alcohol consumption
- $rs_2$  = rescaled relative risk for high alcohol consumption

## II. ARDI ICD Codes

### Chronic Causes

<b>100% Attributable</b>		
<b>Cause</b>	<b>ICD-9</b>	<b>ICD-10</b>
Alcoholic psychosis	291	F10.3-F10.9
Alcohol abuse	305.0, 303.0	F10.0, F10.1
Alcohol dependence syndrome	303.9	F10.2
Alcohol polyneuropathy	357.5	G62.1
Degeneration of nervous system due to alcohol	*	G31.2
Alcoholic myopathy	*	G72.1
Alcohol cardiomyopathy	425.5	I42.6
Alcoholic gastritis	535.3	K29.2
Alcoholic liver disease	571.0-571.3	K70-K70.4, K70.9
Fetal alcohol syndrome	655.4, 760.71	Q86.0
Fetus and newborn affected by maternal use of alcohol	*	P04.3, O35.4
Alcohol-induced chronic pancreatitis	*	K86.0
<b>Direct AAF Estimate</b>		
<b>Cause</b>	<b>ICD-9</b>	<b>ICD-10</b>
Acute pancreatitis	577.0	K85
Chronic pancreatitis	577.1	K86.1
Epilepsy	345	G40, G41
Esophageal varices	456.0-456.2	I85, I98.2
Gastroesophageal hemorrhage	530.7	K22.6
Liver cirrhosis, unspecified	571.5-571.9	K74.3-K74.6, K76.0, K76.9
Portal hypertension	572.3	K76.6
Spontaneous abortion	634	O03

### Indirect AAF Estimate (English et al and Ridolfo and Stevenson cutpoints)

Cause	ICD-9	ICD-10
Breast cancer, females	174	C50
Cholelithiasis	574	K80
Chronic hepatitis	571.4	K73
Esophageal cancer	150	C15
Hypertension	401-405	I10-I15
Ischemic heart disease	410-414	I20-I25
Laryngeal cancer	161	C32
Liver cancer	155	C22
Low birth weight, prematurity, intrauterine growth retardation or death	656.5, 764, 765	O36.5, O36.4, P05, P07
Oropharyngeal cancer	141, 143-146, 148, 149	C01-C06, C09-C10, C12-C14
Psoriasis	696.1	L40.0-L40.4, L40.8, L40.9
Supraventricular cardiac dysrhythmia	427.0, 427.2, 427.3	I47.1, I47.9, I48

### Indirect AAF Estimate (Corrao et al and Bargnardi et al cutpoints)

Cause	ICD-9	ICD-10
Stroke, ischemic	433-435, 437, 362.34	G45, I63, I65-I67, I69.3
Stroke, hemorrhagic	430-432	I60-I62, I69.0-I69.2
Prostate cancer	185	C61

## Acute Causes

<b>100% Alcohol attributable</b>		
<b>Cause</b>	<b>ICD-9</b>	<b>ICD-10</b>
Alcohol poisoning	980.0, 980.1, E860.0, E860.1, E860.2, E860.9	X45, Y15, T51.0, T51.1, T51.9
Suicide by and exposure to alcohol	*	X65
Excessive blood level of alcohol	790.3	R78.0
<b>Direct AAF Estimate</b>		
<b>Cause</b>	<b>ICD-9</b>	<b>ICD-10</b>
Air-space transport	E840-E845	V95-V97
Aspiration	E911	W78-W79
Child maltreatment	E960-E968	X85-Y09, Y87.1
Drowning injuries	E910	W65-W74
Fall injuries	E880-E888, E848	W00-W19
Fire injuries	E890-E899	X00-X09
Firearms	E922	W32-W34
Homicide	E960-E969	X85-Y09, Y87.1
Hypothermia	E901	X31
Motor-vehicle nontraffic crashes	E820-E825	V02.0, V03.0, V04.0, V09.0, V12-V14(.0-.2), V19.0-V19.3, V20-V28(.0-.2), V29.0-V29.3, V30-V39(.0-.3), V40-V49(.0-.3), V50-V59(.0-.3), V60-V69(.0-.3), V70-V79(.0-.3), V81.0, V82.0, V83-V86(.4-.9), V88.0-V88.8, V89.0
Motor-vehicle traffic crashes	E810-E819	V02(.1, .9), V03(.1, .9), V04(.1, .9), V09.2, V12-V14(.3-.9), V19.4-V19.6, V20-V28(.3-.9), V29.4-V29.9, V30-V39(.4-.9), V40-V49(.4-.9), V50-V59(.4-.9), V60-V69(.4-.9), V70-V79(.4-.9), V80.3-V80.5, V81.1, V82.1, V83-V86(.0-.3), V87.0-V87.8, V89.2
Occupational and machine injuries	E917-E920	W24-W31, W45
Other road vehicle crashes	E800-E807, E826-E829	V01, V05-V06, V09.1, V09.3, V09.9, V10-V11, V15-V18, V19.3, V19.8-V19.9, V80.0-V80.2, V80.6-V80.9, V81.2-V81.9, V82.2-V82.9, V87.9, V88.9, V89.1, V89.3, V89.9
Poisoning (not alcohol)	E850-E869, E924.1	X40-X49 (except X45)
Suicide	E950-E959	X60-X84, (except X65) Y87.0
Water transport	E830-E838	V90-V94

\* No ICD-9 code available, the condition is new to ICD-10.

### III. 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)
<b>Chronic causes</b>		
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, 5-9, 10-14, 15-19, 20-85+
Alcohol-induced chronic pancreatitis	ALCCRP	20-85+
Liver cirrhosis, unspecified	LIVCIR	20-85+
Acute pancreatitis	ACUPAN	20-85+
Chronic pancreatitis	CHRPAN	20-85+
Portal hypertension	PORHYP	20-85+
Gastroesophageal hemorrhage	GASHEM	20-85+
Oropharyngeal cancer	OROCAN	20-85+
Esophageal cancer	ESOCAN	20-85+
Liver cancer	LIVCAN	20-85+
Laryngeal cancer	LARCAN	20-85+
Supraventricular cardiac dysrhythmia	SUCADY	20-85+
Esophageal varices	ESOVAR	20-85+
Stroke, ischemic	STROKI	20-85+
Stroke, hemorrhagic	STROKH	20-85+
Ischemic heart disease	ISCHTD	20-85+
Epilepsy	EPILEP	20-85+

<b>Cause of Death</b>	<b>Cause of Death Abbreviation</b>	<b>Age Range (using 5-year age groupings)</b>
Breast cancer (females only)	BRECAN	20-85+
Hypertension	HYPERT	20-85+
Psoriasis	PSORIA	20-85+
Spontaneous abortion (females only)	SPONAB	20-85+
Cholelithiasis	CHOLEL	20-85+
Low birth weight, prematurity, IUGR or death	LBWPID	<1, 1-4, 5-9, 10-14, 15-19, 20-85+
Chronic hepatitis	CHRHEP	20-85+
Prostate cancer (males only)	PROCAN	20-85+
<b>Acute causes</b>		
Alcohol poisoning	ALCPOI	15-85+
Air-space transport	AIRTRA	15-85+
Aspiration	ASPIRA	15-85+
Child maltreatment	CHILAB	<1, 1-4, 5-9, 10-14
Drowning	DROWN1	15-85+
Excessive blood level of alcohol	ALCBLA	15-85+
Fall injuries	FALLSS	15-85+
Fire injuries	FIREIN	15-85+
Firearm injuries	FIREAR	15-85+
Homicide	HOMICI	15-85+
Hypothermia	HYPOTH	15-85+
Motor-vehicle nontraffic crashes	MTVNTR	15-85+
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)	POISON	15-85+
Suicide	SUICID	15-85+
Suicide by and exposure to alcohol	ALCSUC	15-85+
Water transport	WATTRA	15-85+

## IV. 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

## V. Indexing Example

To calculate average daily alcohol consumption, it is preferable to have three pieces of information related to alcohol use: frequency of alcohol consumption, quantity of alcohol consumption, and frequency of binge drinking (consuming five or more drinks per occasion for males or four or more drinks per occasion for females). An example of how to index binge drinks in the calculation of average daily alcohol consumption (adapted from *Stahre M, Naimi T, Brewer R, Holt J. Measuring average alcohol consumption: the impact of including binge drinks in quantity-frequency calculations. Addiction:101(12):1711-8.*):

1. The total number of binge drinking days (BF) (assuming one binge episode per day) is subtracted from the total number of drinking days (F) to equal the number of drinking days with average levels of daily alcohol consumption (AF).
2. The number of days with average consumption levels (AF) is then multiplied by the average number of drinks reported per day (Q), to equal the total number of drinks on average drinking days ( $Q*AF$ ).
3. The number of binge drinking days (BF) is then multiplied by the average number of binge drinks per episode (BQ) (5 for men or 4 for women) to equal the total number of drinks consumed on binge drinking days ( $BF*BQ$ ).
4. The total number of drinks consumed on average drinking days ( $Q*AF$ ) is then added to the total number of drinks consumed on binge drinking days ( $BF*BQ$ ) to equal the total number of drinks consumed within that time period (e.g., past 30 days).
5. The total number of drinks consumed during the specified time period is then divided by the total number of days to equal the average number of drinks consumed per day.

Example: A 25-year-old male reported drinking that in the past 30 days he drank on 12 of those days, drank about 3 drinks per day, and binge drank three times. To calculate average daily alcohol consumption:

1. Twelve total number of drinking days (F) – 3 binge drinking days (BF) = 9 average drinking days (AF).
2. Nine average drinking days (AF) \* 3 drinks per day (Q) = 27 drinks total on average drinking days ( $Q*AF$ ).
3. Three binge drinking days (BF) \* 5 drinks per binge drinking day (BQ) = 15 drinks per binge drinking day ( $BF*BQ$ ).
4. Twenty-seven drinks on average drinking day ( $Q*AF$ ) + 15 drinks per binge drinking day ( $BF*BQ$ ) = 42 total drinks [ $(Q*AF) + (BF*BQ)$ ].
5. Forty-two drinks divided by 30 days = 1.40 drinks consumed per day on average over a 30-day period.

# **APPENDIX B: Sample SAS Codes**

## I. Age Group Codes for SAS XML

Use the following age group codes to import SAS XML data into a custom data set.

Age Group	SAS XML 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*

\*Be sure to delete age groups that are labeled with “AgeU” before uploading into ARDI

## II. Age Group Conversion

```
/******  
/* SAS Version 8 */  
/* ARDI */  
/* Program: Recode Age Groupings for export to SQL */  
/* Datasets 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;
```

### III. Total Deaths by State

```
/******
/* SAS Version 8 */
/* ARDI CHRONIC and ACUTE CONDITIONS BY DATASETNAME */
/* Program: To Generate Frequencies */
/* Datasets used: Mortality: "NAME OF DATA SET" */
/******

/* CREATE MACROS FOR AGE GROUPS DEFINED IN THE DOCUMENT: "Age Group
Codes for SAS XML Import"*/
```

```
/* 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 CHORNIC CAUSES OF DEATH*/
%macro chroniccond;
(agegrp in ("AgeF", "AgeG", "AgeH", "AgeI", "AgeJ", "AgeK", "AgeL", "AgeM",
  "AgeN", "AgeO", "AgeP", "AgeQ", "AgeR", "AgeS"))
%mend;
```

```
/* AGE 0 TO 14 YEARS USED FOR CHILD MALTREATMENT DEATHS*/
%macro juvenileage;
(agegrp in ("AgeA", "AgeB", "AgeC", "AgeD"))
%mend;
```

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

```
libname name 'specify location of data';
```

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

```

data dataset;
set name.DATAFILENAME;
length cause $6;
/* CREATE YOUR DATASETNAME FIELD*/
/* EXAMPLE datasetname = "New York City"; */
/* DO NOT USE A YEAR OR OTHER NUMBER IN THE DATASETNAME*/
datasetname = 'California';

cause = "UNASIGN";

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

/****** CHRONIC CAUSES OF DEATH *****/
/****** 100% ALCOHOL-ATTRIBUTABLE *****/
/**** To verify ICD-10 Codes: apps.who.int/classifications/icd10/browse/2010/en#*****/

/* 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 K 29.2 */
when ((icd10dx = 'K292') and %chroniccond) cause= "ALCGAS";

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

/* 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, O35.4, age macros not used with this condition \*/

when ((icd10dx = 'P043') or (icd10dx = 'O354')) cause= "ALCMAU";

/\* Alcohol induced chronic pancreatitis K86.0 \*/

when ((icd10dx = 'K860') and %chroniccond) cause= "ALCCRP";

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

/\* Acute pancreatitis K85\*/

when ((icd10dx = 'K85') and %chroniccond) cause="ACUPAN";

/\* Chronic pancreatitis K86.1\*/

when ((icd10dx = 'K861') and %chroniccond) cause= "CHRPAN";

/\* Epilepsy G40,G41 \*/

when ((icd10dx = 'G40') or (icd10dx = 'G41') and %chroniccond)  
cause= "EPILEP";

/\* Esophageal varices I85 I98.2 \*/

when ((icd10dx = 'I85') OR (icd10dx in ('I982'))  
and %chroniccond) cause= "ESOVAR";

/\* Gastroesophageal hemorrhage K22.6 \*/

when ((icd10dx = 'K226') and %chroniccond) cause= "GASHEM";

/\* Liver cirrhosis, unspecified K74.3- K74.6, K76.0, K 76.9\*/

when ((icd10dx in ('K743', 'K744', 'K745', 'K746', 'K760', 'K769'))  
and %chroniccond) cause= "LIVCIR";

/\* Portal Hypertension K76.6 \*/

when ((icd10dx = 'K766') and %chroniccond) cause= "PORHYP";

/\* Spontaneous abortion O03 \*/

when ((icd10dx = 'O03') and %chroniccond) cause= "SPONAB";

/\*\*\*\*\*\* INDIRECT AAF ESTIMATE (ENGLISH ET AL & RIDOLFO & STEVENSON CUTPOINTS \*\*\*\*\*/

/\* Breast Cancer, females C50 \*/

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

/\* Cholelithiasis K80 \*/

when ((icd10dx = 'K80') and %chroniccond) cause= "CHOLEL";

/\* Chronic hepatitis K73 \*/

when ((icd10dx = 'K73') and %chroniccond) cause= "CHRHEP";

```

/* Esophageal cancer C15 */
when ((icd10dx = 'C15')and %chroniccond) cause= "ESOCAN";

/* Hypertension I10- I15 */
when ((icd10dx = 'I10') or (icd10dx =: 'I11') or (icd10dx =: 'I12') or
      (icd10dx =: 'I13') or (icd10dx =: 'I14') or (icd10dx =: 'I15')
      and %chroniccond) cause= "HYPERT";

/* Ischemic 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";

/* Laryngeal cancer C32 */
when ((icd10dx =: 'C32')and %chroniccond) cause= "LARCAN";

/* Liver cancer C22 */
when ((icd10dx =: 'C22')and %chroniccond) cause= "LIVCAN";

/* Low birth weight, prematurity, IUGR or death O36.4,O36.5
   P05, P07, age macros not used with this condition */
when ((icd10dx =: 'P05') or (icd10dx =: 'P07') or
      (icd10dx in ('O364', 'O365'))) cause= "LBWPID";

/* Oropharyngeal cancer 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 = "OROCAN";

/* Psoriasis L40.0 - L40.4, L40.8, L40.9 */
when ((icd10dx in ('L400', 'L401', 'L402', 'L403', 'L404', 'L408',
                  'L409')) and %chroniccond) cause= "PSORIA";

/* Supraventricular cardiac dysrhythmia I47.1 I47.9 I48 */
when ((icd10dx in ('I471', 'I479') or (icd10dx =: 'I48'))
      and %chroniccond) cause= "SUCADY";

/****** INDIRECT AAF ESTIMATE (CORRAO ET AL & BARGNARDI ET AL
CUTPOINTS *****/

/* 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= "STROKI";

```

```
/* 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= "STROKH";
```

```
/* Prostate cancer C61 */
when ((icd10dx = 'C61') and %chroniccond) cause= "PROCAN";
```

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

```
/* Alcohol poisoning X45, Y 15, T 51.0, T 51.1, T51.9 */
when ((icd10dx = 'X45') or (icd10dx = 'Y15') or
      (icd10dx in ('T510', 'T511', 'T519'))
      and %acutecond) cause= "ALCPOI";
```

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

```
/* Excessive blood level of alcohol R78.0 */
when ((icd10dx = 'R780')and %acutecond) cause= "ALCBLA";
```

```
/****** 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="CHILAB";
else cause = "HOMICI";
end;
```

```
/* Drowning W65-W74 */
```

when ((icd10dx = 'W65') or (icd10dx = 'W66') or (icd10dx = 'W67')  
or (icd10dx = 'W68') or (icd10dx = 'W69') or (icd10dx = 'W70')  
or (icd10dx = 'W71') or (icd10dx = 'W72') or (icd10dx = 'W73')  
or (icd10dx = 'W74') and %acutecond) cause= "DROWNI";

/\* Falls W00-W19 \*/

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') and %acutecond)  
cause= "FALLSS";

/\* Fire X00-X09 \*/

when ((icd10dx = 'X00') or (icd10dx = 'X01') or (icd10dx = 'X02')  
or (icd10dx = 'X03') or (icd10dx = 'X04') or (icd10dx = 'X05')  
or (icd10dx = 'X06') or (icd10dx = 'X07') or (icd10dx = 'X08')  
or (icd10dx = 'X09') and %acutecond) cause= "FIREIN";

/\* Firearms W32-W34 \*/

when ((icd10dx = 'W32') or (icd10dx = 'W33') or (icd10dx = 'W34')  
and %acutecond) cause= "FIREAR";

/\* Hypothermia X31 \*/

when ((icd10dx = 'X31') and %acutecond) cause= "HYPOTH";

/\* Motor vehicle nontraffic crashes \*/

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', 'V193'))  
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', 'V838', 'V839'))

```

OR (icd10dx in ('V844', 'V845', 'V846', 'V847', 'V848', 'V849'))
OR (icd10dx in ('V854', 'V855', 'V856', 'V857', 'V858', 'V859'))
OR (icd10dx in ('V864', 'V864', 'V866', 'V867', 'V868', 'V869'))
OR (icd10dx in ('V880', 'V881', 'V882', 'V883', 'V884', 'V885',
    'V886', 'V887', 'V888')) and %acutecond)
cause= "MTVNTR";

```

```

/* Motor Vehicle traffic crashes 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', 'V126', 'V127', 'V128', 'V129'))
OR (icd10dx in ('V133', 'V134', 'V135', 'V136', 'V137', 'V138', 'V139'))
OR (icd10dx in ('V143', 'V144', 'V145', 'V146', 'V147', 'V148', 'V149'))
OR (icd10dx in ('V203', 'V204', 'V205', 'V206', 'V207', 'V208', 'V209'))
OR (icd10dx in ('V213', 'V214', 'V215', 'V216', 'V217', 'V218', 'V219'))
OR (icd10dx in ('V223', 'V224', 'V225', 'V226', 'V227', 'V228', 'V229'))
OR (icd10dx in ('V233', 'V234', 'V235', 'V236', 'V237', 'V238', 'V239'))
OR (icd10dx in ('V243', 'V244', 'V245', 'V246', 'V247', 'V248', 'V249'))
OR (icd10dx in ('V253', 'V254', 'V255', 'V256', 'V257', 'V258', 'V259'))
OR (icd10dx in ('V263', 'V264', 'V265', 'V266', 'V267', 'V268', 'V269'))
OR (icd10dx in ('V273', 'V274', 'V275', 'V276', 'V277', 'V278', 'V279'))
OR (icd10dx in ('V283', 'V284', 'V285', 'V286', 'V287', 'V288', 'V289'))
OR (icd10dx in ('V294', 'V295', 'V296', 'V297', 'V298', 'V299'))
OR (icd10dx in ('V304', 'V305', 'V306', 'V307', 'V308', 'V309'))
OR (icd10dx in ('V314', 'V315', 'V316', 'V317', 'V318', 'V319'))
OR (icd10dx in ('V324', 'V325', 'V326', 'V327', 'V328', 'V329'))
OR (icd10dx in ('V334', 'V335', 'V336', 'V337', 'V338', 'V339'))
OR (icd10dx in ('V344', 'V345', 'V346', 'V347', 'V348', 'V349'))
OR (icd10dx in ('V354', 'V355', 'V356', 'V357', 'V358', 'V359'))
OR (icd10dx in ('V364', 'V365', 'V366', 'V367', 'V368', 'V369'))
OR (icd10dx in ('V374', 'V375', 'V376', 'V377', 'V378', 'V379'))
OR (icd10dx in ('V384', 'V385', 'V386', 'V387', 'V388', 'V389'))
OR (icd10dx in ('V394', 'V395', 'V396', 'V397', 'V398', 'V399'))
OR (icd10dx in ('V404', 'V405', 'V406', 'V407', 'V408', 'V409'))
OR (icd10dx in ('V414', 'V415', 'V416', 'V417', 'V418', 'V419'))
OR (icd10dx in ('V424', 'V425', 'V426', 'V427', 'V428', 'V429'))
OR (icd10dx in ('V434', 'V435', 'V436', 'V437', 'V438', 'V439'))
OR (icd10dx in ('V444', 'V445', 'V446', 'V447', 'V448', 'V449'))
OR (icd10dx in ('V454', 'V455', 'V456', 'V457', 'V458', 'V459'))
OR (icd10dx in ('V464', 'V465', 'V466', 'V467', 'V468', 'V469'))
OR (icd10dx in ('V474', 'V475', 'V476', 'V477', 'V478', 'V479'))
OR (icd10dx in ('V484', 'V485', 'V486', 'V487', 'V488', 'V489'))
OR (icd10dx in ('V494', 'V495', 'V496', 'V497', 'V498', 'V499'))
OR (icd10dx in ('V504', 'V505', 'V506', 'V507', 'V508', 'V509'))
OR (icd10dx in ('V514', 'V515', 'V516', 'V517', 'V518', 'V519'))
OR (icd10dx in ('V524', 'V525', 'V526', 'V527', 'V528', 'V529'))
OR (icd10dx in ('V534', 'V535', 'V536', 'V537', 'V538', 'V539'))

```

```

OR (icd10dx in ('V544', 'V545', 'V546', 'V547', 'V548', 'V549'))
OR (icd10dx in ('V554', 'V555', 'V556', 'V557', 'V558', 'V559'))
OR (icd10dx in ('V564', 'V565', 'V566', 'V567', 'V568', 'V569'))
OR (icd10dx in ('V574', 'V575', 'V576', 'V577', 'V578', 'V579'))
OR (icd10dx in ('V584', 'V585', 'V586', 'V587', 'V588', 'V589'))
OR (icd10dx in ('V594', 'V595', 'V596', 'V597', 'V598', 'V599'))
OR (icd10dx in ('V604', 'V605', 'V606', 'V607', 'V608', 'V609'))
OR (icd10dx in ('V614', 'V615', 'V616', 'V617', 'V618', 'V619'))
OR (icd10dx in ('V624', 'V625', 'V626', 'V627', 'V628', 'V629'))
OR (icd10dx in ('V634', 'V635', 'V636', 'V637', 'V638', 'V639'))
OR (icd10dx in ('V644', 'V645', 'V646', 'V647', 'V648', 'V649'))
OR (icd10dx in ('V654', 'V655', 'V656', 'V657', 'V658', 'V659'))
OR (icd10dx in ('V664', 'V665', 'V666', 'V667', 'V668', 'V669'))
OR (icd10dx in ('V674', 'V675', 'V676', 'V677', 'V678', 'V679'))
OR (icd10dx in ('V684', 'V685', 'V686', 'V687', 'V688', 'V689'))
OR (icd10dx in ('V694', 'V695', 'V696', 'V697', 'V698', 'V699'))
OR (icd10dx in ('V704', 'V705', 'V706', 'V707', 'V708', 'V709'))
OR (icd10dx in ('V714', 'V715', 'V716', 'V717', 'V718', 'V719'))
OR (icd10dx in ('V724', 'V725', 'V726', 'V727', 'V728', 'V729'))
OR (icd10dx in ('V734', 'V735', 'V736', 'V737', 'V738', 'V739'))
OR (icd10dx in ('V744', 'V745', 'V746', 'V747', 'V748', 'V749'))
OR (icd10dx in ('V754', 'V755', 'V756', 'V757', 'V758', 'V759'))
OR (icd10dx in ('V764', 'V765', 'V766', 'V767', 'V768', 'V769'))
OR (icd10dx in ('V774', 'V775', 'V776', 'V777', 'V778', 'V779'))
OR (icd10dx in ('V784', 'V785', 'V786', 'V787', 'V788', 'V789'))
OR (icd10dx in ('V794', 'V795', 'V796', 'V797', '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";

```

```

/* 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 */
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')
and %acutecond) cause= "POISON";
```

```
/* 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 = "UNKNWN";
end;
```

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

```
LENGTH nstate $30 nsex $6;
```

```
if state = . then nstate = 'UNITED STATES';
else if state = 01 then nstate = 'ALABAMA';
else if state = 02 then nstate = 'ALASKA';
else if state = 05 then nstate = 'ARKANSAS';
else if state = 60 then nstate = 'AMERICAN SAMOA';
else if state = 04 then nstate = 'ARIZONA';
else if state = 06 then nstate = 'CALIFORNIA';
else if state = 08 then nstate = 'COLORADO';
else if state = 09 then nstate = 'CONNECTICUT';
else if state = 11 then nstate = 'DISTRICT OF COLUMBIA';
else if state = 10 then nstate = 'DELAWARE';
```

```

else if state = 12 then nstate = 'FLORIDA';
else if state = 13 then nstate = 'GEORGIA';
else if state = 66 then nstate = 'GUAM';
else if state = 15 then nstate = 'HAWAII';
else if state = 19 then nstate = 'IOWA';
else if state = 16 then nstate = 'IDAHO';
else if state = 17 then nstate = 'ILLINOIS';
else if state = 18 then nstate = 'INDIANA';
else if state = 20 then nstate = 'KANSAS';
else if state = 21 then nstate = 'KENTUCKY';
else if state = 22 then nstate = 'LOUISIANA';
else if state = 25 then nstate = 'MASSACHUSETTS';
else if state = 24 then nstate = 'MARYLAND';
else if state = 23 then nstate = 'MAINE';
else if state = 26 then nstate = 'MICHIGAN';
else if state = 27 then nstate = 'MINNESOTA';
else if state = 29 then nstate = 'MISSOURI';
else if state = 28 then nstate = 'MISSISSIPPI';
else if state = 30 then nstate = 'MONTANA';
else if state = 37 then nstate = 'NORTH CAROLINA';
else if state = 38 then nstate = 'NORTH DAKOTA';
else if state = 31 then nstate = 'NEBRASKA';
else if state = 33 then nstate = 'NEW HAMPSHIRE';
else if state = 34 then nstate = 'NEW JERSEY';
else if state = 35 then nstate = 'NEW MEXICO';
else if state = 32 then nstate = 'NEVADA';
else if state = 36 then nstate = 'NEW YORK';
else if state = 39 then nstate = 'OHIO';
else if state = 40 then nstate = 'OKLAHOMA';
else if state = 41 then nstate = 'OREGON';
else if state = 42 then nstate = 'PENNSYLVANIA';
else if state = 72 then nstate = 'PUERTO RICO';
else if state = 44 then nstate = 'RHODE ISLAND';
else if state = 45 then nstate = 'SOUTH CAROLINA';
else if state = 46 then nstate = 'SOUTH DAKOTA';
else if state = 47 then nstate = 'TENNESSEE';
else if state = 48 then nstate = 'TEXAS';
else if state = 49 then nstate = 'UTAH';
else if state = 51 then nstate = 'VIRGINIA';
else if state = 78 then nstate = 'VIRGIN ISLANDS';
else if state = 50 then nstate = 'VERMONT';
else if state = 53 then nstate = 'WASHINGTON';
else if state = 55 then nstate = 'WISCONSIN';
else if state = 54 then nstate = 'WEST VIRGINIA';
else if state = 56 then nstate = 'WYOMING';

```

```

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

```

```

if sex=2 then nsex= '2_Female';

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

/* DELETE ALL RECORDS WHERE THERE IS AN UNKNOWN AGE */
if agegrp = "AgeU" then delete;
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 = dataset;
  tables datasetname*state*agegrp*cause*sex / list missprint nopercnt nocum out=newdata
  noprint;
  run;

/* TRANSPOSE TO HAVE MALE AND FEMALE ESTIMATES OUTPUT IN ROWS */
proc transpose data = newdata out = datasetname2;
  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 datasetname3;
set datasetname2;
drop _name_ _label_;
if _1_Male = . then Male = 0; else Male=_1_Male;
if _2_Fema = . then Female = 0; else Female=_2_Fema;
drop _1_Male _2_Fema;
run;

/*GENERATE XML FILE*/
/*THE "XML" TAG WILL PERFORM THE REGQUIRED TRANSFORMATION FOR THE
SAS DATASET*/
/*THE DATASET NEEDS TO BE NAMED ACCORDING TO THE FOLLOWING NAMING
CONVENTION:
RANGE OF YEARS DATA SET IS COVERING (EX: 2006-2010)_STATE_CURRENT
YEAR_MORTALITY*/
libname target xml 'location to store xml dataset\XML data set name: Year
range_State_Year_mortality.xml';

/*ALCXML is required for the XML file to upload correctly into the ARDI application*/
data target.alcxml;
  set datasetname3;
  run;

```

**Troubleshooting:** In the above SAS code, it is assumed that users are using national formats for the codes; however, some states may use a different coding scheme for their subcodes (e.g., three digits). The above code does take this into consideration but it may not be applied accurately for data sets of other formats.

**Example A:**

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

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

**Example B:**

There are some causes in which we only want specific subcategories within each death, psoriasis is an example:

```
/* Psoriasis L40.0 - L40.4, L40.8, L40.9 */
when ((icd10dx in ('L400', 'L401', 'L402', 'L403', 'L404', 'L408',
'L409')) and %chroniccond) cause= "PSORIA";
```

If your code includes numbers with 3 or more digits:

Make sure that the digits beyond the first 2 are being captured for all causes written with the “=:”. If the code is not capturing all of your data, you may need to add the remaining digits to your code following the equal sign (e.g., C500, C501, C502 ... C509 in Example A above).

## IV. Calculation of Average Daily Alcohol Consumption Using BRFSS

This document contains sample SAS code to calculate the prevalence of alcohol consumption using the Behavioral Risk Factor Surveillance System (BRFSS). The prevalence of alcohol consumption is calculated for two different sets of cutpoints and separately for each sex. The cutpoints reflect the two different meta-analyses used to select the relative risk estimates for specific causes of death. For more information, please refer to the ARDI Methods.

Notes about the sample code:

1. The example below uses variables from the 2005 BRFSS. Please be aware that different years of the BRFSS may have different variable names so it is important for you to check if the year of data that you are using has identical variable names for *each* variable. For example, DRNK2GE was the variable name used for years prior to 2006 but was changed to DRNK3GE in later years.
2. If combining several years of data, please refer to the BRFSS website regarding combining the weighting variables used in the data sets: [www.cdc.gov/brfss](http://www.cdc.gov/brfss).
3. Always use the appropriate population weights – refer to the BRFSS website for information regarding the weighting variables.
4. Only delete those people for whom age information is missing or less than age 20.
5. Do not delete people for whom alcohol information is missing – this will give you an incorrect denominator to calculate the rest of the percentages.
6. The cutpoints used in ARDI specify a minimum level of drinking, thus there are some people that are not included in the lowest drinking category (they are coded 999).

libname 'C: Enter the location of the data set'

**data** brfssmale;

set libraryname.brfssdatasetname;

/\*THIS DATA SET IS FOR MALES ONLY\*/

/\*need to delete females and those people for whom we are missing age information\*/

If sex = **2** then delete;

If age <**20** then delete;

If age in (**7 9**) then delete;

/\*ALCDAY5 refers to the number of days per month or per week a person drank alcohol\*/

/\*AVEDRNK2 refers to the amount of alcohol consumed on drinking days\*/

/\*DRNK3GE5 refers to number of binge episodes (drank 5 or more drinks for men/4 or more drinks for women) in past 30 days\*/

/\*MAXDRNKS refers to the largest number of drinks consumed on any occasion in the past 30 days\*/

/\*DRNKANY5 is a calculated variable to determine if someone drank alcohol in the past 30 days\*/

if AVEDRNK2 ge **77** then AVEDRNK2 =.;

if DRNK3GE5 = **88** then DRNK3GE5 = **0**;

```

if DRNK3GE5 ge 77 then DRNK3GE5 = .;
if MAXDRNKS ge 77 then MAXDRNKS =.;

/*Determine what proportion of days in the past 7 or 30 someone drank alcohol*/
if 101<=ALCDAY5<=107 then do;
days=(ALCDAY5-100)/7;
end;else
if 201<=ALCDAY5<=230 then do;
days=(ALCDAY5-200)/30;
end;else
if ALCDAY5=888 then days=0;else
days=.;
ddays = days * 30; /*calculates how many days in the past month someone drank alcohol*/

/*This section of code applies the indexing technique to the calculation of average daily alcohol
consumption*/

/*If a person does not binge drink and max drinks per episode in the last 30 days was less than 5
on any occasion then you do not use the indexing technique*/
noindex = days*AVEDRNK2;

/*Intensity is used to determine the number of drinks someone consumes when they binge drink
and is derived from the MAXDRNKS variable*/
If DRNK3GE5 >0 then intensity=MAXDRNKS;

/*Use indexing technique to calculate the average number of drinks per day someone consumed
in the past month taking into account the largest number of drinks someone consumed on one
occasion*/
binge = DRNK3GE5/30; /*proportion of days in past 30 someone binge drank*/
dbinge= ((days - binge)*AVEDRNK2); /*subtract the binge drinking proportion of days from
proportion of days someone drinks and multiply the difference by the average number of
drinks*/
bingem= binge*intensity; /*multiplies the proportion of binge drinking days by the max drinks
someone drank*/
indexm= dbinge + bingem; /*sums the number of drinks someone drinks on average and while
binging*/

/*SOME PEOPLE DO NOT REPORT MAX DRINKS BUT DO REPORT AVERAGE
DRINKS AND BINGE DRINKING, index using 5 drinks*/
if MAXDRNKS = . and AVEDRNK2 >0 and DRNK3GE5 >0 then do;
binge = DRNK3GE5/30;
dbinge = ((days-binge)*AVEDRNK2);
bingem=binge*5;
indexm = dbinge + bingem;
end;

/*SOME PEOPLE DO REPORT NUMBER OF DAYS DRANK AND AVERAGE NUMBER
OF DRINKS BUT DO NOT REPORT BINGE DRINKING EPISODES*/

```

```

if DRNK3GE5 = . and AVEDRNK2 >0 AND ALCDAY5 >0 then indexm = noindex;

/*Make sure nondrinkers are not being indexed*/
if ALCDAY5 = 888 then noindex = 0;
if ALCDAY5 = 888 then indexm = 0;

/*Coding missing values for the caluclations of drinking level*/
if DRNKANY5 not in (7 9) and noindex =. then noindex =99;
if DRNKANY5 not in (7 9) and indexm = . then indexm =99;

/*Make sure males who drink but do not binge drink are not indexed*/
if DRNKANY5 = 1 and DRNK3GE5 = 0 then indexm = noindex;

/*This section of code estimates the prevalence of alcohol consumption by the cutpoints
specified by the English et al study (variable name = engm)*/
/*The English et al study specifies a minimum level of drinking and those people drinking below
this threshold are not included in the low drinking category*/
if indexm =0.00000 then engm =0; /*nondrinkers*/
if 0.00001 <= indexm <= 0.19999 then engm =999; /*drinking below minimum threshold for
low drinking category*/
if 0.20000 <= indexm <= 2.89999 then engm =1; /*low drinking category*/
if 2.90000 <= indexm <= 4.29999 then engm =2; /*medium drinking category*/
if 4.30000 <= indexm <= 98 then engm =3; /*high drinking category*/
if indexm =99 then engm =99; /*drinking information is missing*/

/*This section of code estimates the prevalence of alcohol consumption by the cutpoints
specified by the Corroa et al study (variable name = corrm)*/
/*The Corroa et al study specifies a minimum level of drinking and those people drinking below
this threshold are not included in the low drinking category*/
if indexm =0.00000 then corrm =0; /*nondrinkers*/
if 0.00001 <= indexm <= 0.09999 then corrm =999; /*drinking below minimum threshold for
low drinking category*/
if 0.10000 <= indexm <= 1.79999 then corrm =1; /*low drinking category*/
if 1.80000 <= indexm <= 3.59999 then corrm =2; /*medium drinking category*/
if 3.60000 <= indexm <= 98 then corrm =3; /*high drinking category*/
if indexm =99 then corrm =99; /*drinking information is missing*/

run;

PROC FREQ DATA = BRFSSMALE;
TABLES ENGM CORRM;
RUN;

data brfssfemal;
set librname.brfssdatasetname;

/*THIS DATA SET IS FOR FEMALES ONLY*/
/*need to delete males and those people for whom we are missing age information*/

```

If sex = **1** then delete;  
If age <**20** then delete;  
If age in (**7 9**) then delete;

/\*ALCDAY5 refers to the number of days per month or per week a person drank alcohol\*/  
/\*AVEDRNK2 refers to the amount of alcohol consumed on drinking days\*/  
/\*DRNK3GE5 refers to number of binge episodes (drank 5 or more drinks for men/4 or more drinks for women) in past 30 days\*/  
/\*MAXDRNKS refers to the largest number of drinks consumed on any occasion in the past 30 days\*/  
/\*DRNKANY5 is a calculated variable to determine if someone drank alcohol in the past 30 days\*/

if AVEDRNK2 ge **77** then AVEDRNK2 =.;  
if DRNK3GE5 = **88** then DRNK3GE5 = **0**;  
if DRNK3GE5 ge **77** then DRNK3GE5 = .;  
if MAXDRNKS ge **77** then MAXDRNKS =.;

/\*Determine what proportion of days in the past 7 or 30 someone drank alcohol\*/  
if **101**<=ALCDAY5<=**107** then do;  
days=(ALCDAY5-**100**)/**7**;  
end;else  
if **201**<=ALCDAY5<=**230** then do;  
days=(ALCDAY5-**200**)/**30**;  
end;else  
if ALCDAY5=**888** then days=**0**;else  
days=.;  
ddays = days \* **30**; /\*calculates how many days in the past month someone drank alcohol\*/

/\*This section of code applies the indexing technique to the calculation of average daily alcohol consumption\*/

/\*If a person does not binge drink and max drinks per episode in the last 30 days was less than 4 on any occasion then you do not use the indexing technique\*/  
noindex = days\*AVEDRNK2;

/\*Intensity is used to determine the number of drinks someone consumes when they binge drink and is derived from the MAXDRNKS variable\*/  
If DRNK3GE5 >**0** then intensity=MAXDRNKS;

/\*Use indexing technique to calculate the average number of drinks per day someone consumed in the past month taking into account the largest number of drinks someone consumed on one occasion\*/  
binge = DRNK3GE5/**30**; /\*proportion of days in past 30 someone binge drank\*/  
dbinge= ((days - binge)\*AVEDRNK2); /\*subtract the binge drinking proportion of days from proportion of days someone drinks and multiply the difference by the average number of drinks\*/

```
bingef= binge*intensity; /*multiples the proportion of binge drinking days by the max drinks someone drank*/
```

```
indexf= dbinge + bingef; /*sums the number of drinks someone drinks on average and while binging*/
```

```
/*SOME PEOPLE DO NOT REPORT MAX DRINKS BUT DO REPORT AVERAGE DRINKS AND BINGE DRINKING, index using 4 drinks*/
```

```
if MAXDRNKS = . and AVEDRNK2 >0 and DRNK3GE5 >0 then do;
```

```
binge = DRNK3GE5/30;
```

```
dbinge = ((days-binge)*AVEDRNK2);
```

```
bingef=binge*4;
```

```
indexf = dbinge + bingef;
```

```
end;
```

```
/*SOME PEOPLE DO REPORT NUMBER OF DAYS DRANK AND AVERAGE NUMBER OF DRINKS BUT DO NOT REPORT BINGE DRINKING EPISODES*/
```

```
if DRNK3GE5 = . and AVEDRNK2 >0 AND ALCDAY5 >0 then indexf = noindex;
```

```
/*Make sure nondrinkers are not being indexed*/
```

```
if ALCDAY5 = 888 then noindex = 0;
```

```
if ALCDAY5 = 888 then indexf = 0;
```

```
/*Coding missing values for the caluclations of drinking level*/
```

```
if DRNKANY5 not in (7 9) and noindex = . then noindex =99;
```

```
if DRNKANY5 not in (7 9) and indexf = . then indexf =99;
```

```
/*Make sure females who drink but do not binge drink are not indexed*/
```

```
if DRNKANY5 = 1 and DRNK3GE5 = 0 then indexf = noindex;
```

```
/*This section of code estimates the prevalence of alcohol consumption by the cutpoints specified by the English et al study (variable name = engw)*/
```

```
/*The English et al study specifies a minimum level of drinking and those people drinking below this threshold are not included in the low drinking category*/
```

```
if indexf =0.00000 then engw =0; /*nondrinkers*/
```

```
if 0.00001 <= indexf <= 0.19999 then engw =999;
```

```
if 0.20000 <= indexf <= 1.39999 then engw =1; /*low drinking category*/
```

```
if 1.40000 <= indexf <= 2.89999 then engw =2; /*medium drinking category*/
```

```
if 2.90000 <= indexf <= 98 then engw =3; /*high drinking category*/
```

```
if indexf =99 then engw =99; /*drinking information is missing*/
```

```
/*This section of code estimates the prevalence of alcohol consumption by the cutpoints specified by the Corroa et al study (variable name = corrw)*/
```

```
/*The Corroa et al study specifies a minimum level of drinking and those people drinking below this threshold are not included in the low drinking category*/
```

```
if indexf =0.00000 then corrw =0; /*nondrinkers*/
```

```
if 0.00001 <= indexf <= 0.09999 then corrw =999;
```

```
if 0.10000 <= indexf <= 1.79999 then corrw =1; /*low drinking category*/
```

```
if 1.80000 <= indexf <= 3.59999 then corrw =2; /*medium drinking category*/
```

```
if 3.60000 <= indexf <= 98 then corrw =3; /*high drinking category*/  
if indexf =99 then corrw =99; /*drinking information is missing*/
```

```
run;
```

```
PROC FREQ DATA = BRFSFeMALE;  
TABLES ENGW CORRW;  
RUN;
```