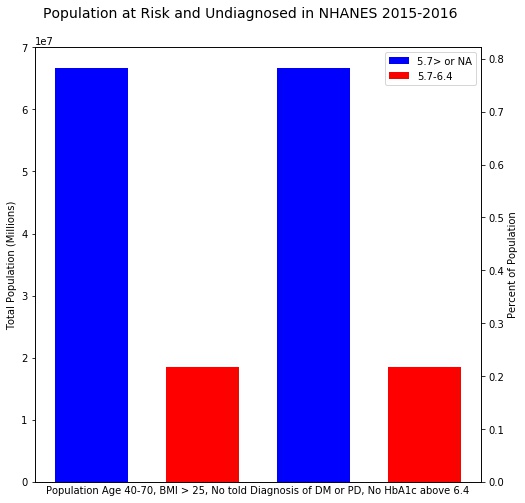
**COST IMPLICATIONS OF UNDIAGNOSED AND UNTREATED DIABETES**

1) Use the 2015-2016 NHANES and determine the proportion of individuals with undiagnosed prediabetes among individuals 40-70 who are overweight or obese (USPSTF guideline for screening). We will use HbA1c 5.7%-6.4% as the indicator of prediabetes.

Eliminate anyone with diagnosed diabetes or diagnosed prediabetes.





Data Log

All files are located in a private repository located at

<https://github.com/andrewcistola/PreDM.git>

8/20

Using SAS:

Followed directions on NHANES Data tutorial

(https://www.cdc.gov/nchs/tutorials/NHANES/preparing/download/intro.htm)

Downloaded 2015-2016 data in SAS format

Collected DEMO, DIQ, GHB, GLU, HIQ, HUQ, INS, OGTT variables

Followed directions to convert to SAS files

Merged into single SAS dataset,

saved code as NHANES\_DOWNLOAD in NHANES/CODE

saved data as MERGE in NHANES/DATA

exported csv file to PreDM/\_data for staging in python

8/21

Using Excel:

nhanes\_merged\_raw.xlsx file staged in excel for analysis in python

All columns converted to number

File converted to csv

file saved as nhanes\_merged\_staged.csv

Using Python:

Nhanes\_merged\_raw.xlsx file read into Python

Created new column UndgDM based on

DIQ10 not being 1 (no diabetes diagnosis)

DIQ010 not being 1 (no prediabetes diagnosis)

LBGXGH over 6.4

1 = undiagnosed dm, 2 = no

Created new column based on

DIQ10 not being 1 (no diabetes diagnosis)

DIQ010 not being 1 (no prediabetes diagnosis)

UndgDM not being 1 (not undiagnosed DM)

LBGXGH between 5.7 and 6.4

1 = undiagnosed pd, 2 = no

“Prediabetes” or “Healthy or Unknown”

Wrote to csv with added columns nhanes\_merged\_undg\_staged.csv

Subset by RIDAGEYR 40-70

BMXBMI over 25

Removed all extra columns

Grouped by WTMECYR (MEC weights)

Exported table and created plots