Instruction: estimating the size of the population at high risk of coronavirus

Data source: 2018 Medicaid claims, MEPS, ACS

Program: SAS, Stata

NYU Wagner | March 2020

**Medicaid**

Two SAS programs, ICD 10/CCS crosswalks.

Step1. To run the Medicaid analysis, you will need to create a diagnostic database using the claims. The diagnostic database should have a diagnosis in a row, including both primary and secondary diagnoses, with individual identifiers in a separate column like the example below.

|  |  |
| --- | --- |
| Mbr1 | ICD1 |
| Mbr1 | ICD2 |
| Mbr1 | ICD3 |
| Mbr2 | ICD1 |
| Mbr2 | ICD2 |
| Mbr3 | ICD1 |
| Mbr3 | ICD2 |

Step2. Using sql to tag along the CCS categories and DX\_cancer from the cross walk “CCS ICD10 with CCI Chronic Flag WITH DX CATS V4.xlsx”

Step3. Run “Coronavirus High Risk Population - Medicaid - Building Database.sas” to create an individual database for coronavirus risk analysis

Step4. Run “Coronavirus High Risk Population - Medicaid - Making Tables.sas” to create the table 1-6

**MEPS**

1. Obtain the full year consolidated MEPS datasets for 2014-2017 from here. <https://www.meps.ahrq.gov/mepsweb/data_stats/download_data_files_results.jsp?cboDataYear=All&cboDataTypeY=1%2CHousehold+Full+Year+File&buttonYearandDataType=Search&cboPufNumber=All&SearchTitle=Consolidated+Data>. Follow MEPS instructions on how to load said data into Stata.
2. Do the same for the MEPS medical conditions datasets from 2014-2017. <https://www.meps.ahrq.gov/mepsweb/data_stats/download_data_files_results.jsp?cboDataYear=All&cboDataTypeY=1%2CHousehold+Full+Year+File&buttonYearandDataType=Search&cboPufNumber=All&SearchTitle=Medical+Conditions>
3. The file “AppendYearsAndMakeHierarchy.do” will append the full year consolidated files, separate out those in the western census region, and generate the following insurance hierarchy
   1. Any Other Insurance
   2. Private Insurance
   3. Medicaid
   4. Medicare
   5. Medicaid+Medicare Duals
   6. Uninsured
4. The file “AppendYearsAndCleanConditionsForHIV.do” appends the medical condition files and keeps observations that record either the clinical classification code for HIV infection for 2014 and 2015, or the ICD10 code in 2016 and 2017.
5. The file “CleanMEPS2014-2017” merges the HIV status from the condition files, defines age categories, creates dummy variables for the conditions deemed as high risk, and generates a variable for the number of conditions a person has.
6. The file “MakeTables.do” makes the left part, without the subheading of MEPS Ratio, of the MEPS Chronic Disease Prevalence by Insurance Status, MEPS Disease Prevalence by Insurance – One Condition, and MEPS Disease Prevalence by Insurance – One or More Conditions.
7. To make the right part with the subheading of MEPS Ratio, take the value for the insurance status and disease prevalence of interest in the 50-64 age category and divide it by the corresponding Medicaid value. For example, the MEPS ratio in the Chronic Disease Prevalence by Insurance Status for uninsured diabetics is 0.4905. This is obtained by dividing the disease prevalence of uninsured diabetics, 9.34%, by the disease prevalence of Medicaid diabetics, 19.05%. We use the 50-64 age category because of low cell counts in other insurance and age categories; very few people age 75 and up are uninsured for the full year.

**ACS.**

We use the ACS to calculate the number of NYC residents who fall into each of the age/gender/coverage groups. These values are located in the population row of each insurance category. We then multiply each cell by the Medicaid prevalence estimates from Table 3. For groups not covered by Medicaid, and for Medicaid/Medicare duals, we then adjust the figures based on the factors in table 7a and 7b.