**Generating Tables for the Web Tool**

This process describes how to construct tables for the web-based tool for cancer reporting by zones. A template for the tool is available on GitHub at <https://github.com/NCI-NAACCR-Zone-Design/Cancer-Map-Template>. If desired, Westat can assist in developing an initial deployment of the web tool.

For this process, we assume files reside in a Windows folder called C:\Work\WebToolTables. This folder can be located anywhere, included on a network drive. The SEER\*Stat output files contain tables with small cells that will ultimately be suppressed based on your desired suppression rules. Thus, access to the folder where the SEER\*Stat output files are kept should be limited to authorized users.

In this process, we refer to variable parameters using square brackets. For example, [StAbbr] refers to the state abbreviation for your registry.

* Create a copy of the WebToolTables distribution folder in the desired location (we assume this is at “C:\Work\WebToolTables” for this process). Copy the final zoned tracts and zone list Excel files provided with the zone design results (“ZonedTracts\_[RunNum]\_final.xlsx” and “ZoneList\_[RunNum]\_final.xlsx”) to this folder.
* This process assumes that cancer incidence rates for zones and the state as a whole have been calculated using the Zone Rate Calculations process. Copy the two resulting exported SEER\*Stat matrix files (“[RunNum]zone\_RateCalcs.txt” and “[StAbbr]state\_RateCalcs.txt”) to the WebToolTables folder. If county rates were also calculated, copy the resulting exported SEER\*Stat matrix file (“[StAbbr]county\_RateCalcs.txt”) to the WebToolTables folder as well.
* Before starting to generate tables, decide what data you want to include in the web tool. Options include:
  + What cancer sites to include.
  + Years of cancer data to be presented: latest data year and year ranges (1-year, 5-year, 10-year)
  + Which race/ethnic groups to include
  + What demographic variables to include and how to align the demographic data years with the cancer data years
* Create a version of the CancerSiteTable\_[StAbbr] Excel file for your registry in the WebToolTables folder. A sample version of this file is provided in the distribution folder. The first worksheet called “SEER\_Stat” is used in SAS Step 51 to process the SEER\*Stat zone rate tables. The rest of the worksheets are used in the web tool as indicated below. See the web tool ReadMe file for details.
  + The “SEER\_Stat” worksheet: should include a row for each of the cancer sites included in the SEER\*Stat output from Step 3b ([RunNum]zone\_RateCalcs.xlsx)
  + The “WebTool CANCERSITE” worksheet: include only the cancer sites that you wish to include in the web tool. In the sample worksheet, sites are in order by descending rates at the national level. Rarer cancer sites can be deleted from the end of the worksheet if desired. The contents should match the web tool table SEARCHOPTIONS\_CANCERSITE in index.js.
  + The “WebTool SEX” worksheet: used by the web tool to label the sex values; it is unlikely that this worksheet would need to be changed. The contents should match the web tool table SEARCHOPTIONS\_SEX in index.js.
  + The “WebTool CANCERSEX” worksheet: used by the web tool to identify cancer sites that are reported for just one sex. Entries for the cancer sites (“key” variable) should match the entries for the sex-specific cancer sites in the “WebTool CANCERSITE” worksheet (“value” variable). Should match the web tool table CANCER\_SEXES in index.js.
  + The “WebTool TIME” worksheet: should have a row for group of reporting years to include in the web tool. The contents should match the web tool table SEARCHOPTIONS\_TIME in index.js.
  + The “WebTool RACE” worksheet: should have a row for race/ethnicity to include in the web tool. The contents should match the web tool table SEARCHOPTIONS\_RACE in index.js.
* Run the 51\_ZoneWebToolRates\_public\_county SAS program to create the cancer incidence rate tables for the web tool based on the SEER\*Stat results
  + Set the %let parameters at the beginning of the code to specify:
    1. The state abbreviation and final run number
    2. The latest year and year groupings of the data
    3. Whether or not to include national rates and, if so, the name of the SEER\*Stat matrix output dataset with the appropriate national cancer rates. National rates are available based on USCS data in the “National\_Cancer\_Rates” folder. Use the dataset with the ending year you need. For example, “AllUS\_Combined\_2021subm\_to2017.txt” has 1-year, 5-year, and 10-year data ending in 2017.
    4. Whether or not to include county rates.
    5. The data path if different from “C:\Work\WebToolTables”
  + Investigate any warning messages generated by the program as these usually mean something is not working as expected
  + Cells with 15 or fewer cases are suppressed in the DATA step where the “RateTable\_wSuppr” dataset is created (the section of code labeled “Suppress counts and rates if 15 or fewer cases”). This code may need to be modified if you have different suppression rules.
  + The program creates a table of cancer incidence rates needed for the web tool with all geographies combined (e.g., zone, state, national, county), as well as a summary report.
* Create a version of the DemogVarTable\_[StAbbr] Excel file for your registry in the WebToolTables folder. A sample version of this file is provided in the distribution folder. The two worksheets are used in the web tool as indicated below. See the web tool ReadMe file for details.
  + The “Webtool DEMOG\_TABLES” worksheet: should include a row for each demographic variable to be displayed in the web tool. The contents should match the web tool table DEMOGRAPHIC\_TABLES in index.js.
  + The “Webtool Tooltips” worksheet: should include a row for each tooltip identified in the “Webtool DEMOG\_TABLES” worksheet. The “tooltip contents” column should include an explanation of the variable and a description of the source of the data. The contents should match the web tool tooltips as defined in index.html.
* Run the 52\_ZoneWebToolDemog\_public\_county SAS program to create the demographic data tables for the web tool
  + Set the %let parameters at the beginning of the code to specify:
    1. The state, state abbreviation, and final run number
    2. Whether or not to include national data
    3. Whether or not to include county data
    4. Demographic variables to include from Census 2010 and ACS

The program currently assumes that any needed Census 2010 variables are included in the ZoneList dataset imported from the ZoneList\_[RunNum]\_final.xlsx file. ACS variables can be drawn from Westat’s tract-level ACS data extract (copies and data dictionaries are available in the “Census\_Data\_Tables” folder under the “materials” parent folder). These datasets are available for ACS 2011-2022, where the ACS 2011-2019 datasets are based on Census 2010 geographies and the ACS 2020-2022 datasets are based on Census 2020 geographies. Note that zones are currently defined based on Census 2010 tracts.

For ACS variables, there is an option to include two data periods to calculate the demographic statistic across 10 years to match the 10-year cancer data period. For the ACS parameters, set:

* + - * Whether to use one ACS period or two
        + If presenting 10-year cancer rates, you can either use ACS variables based on a single 5-year ACS period or use the average values across two 5-year ACS periods. To use a second ACS period, set oneACSPeriod=no and specify the ACS\_endyr2 parameter (see below). Otherwise, set oneACSPeriod=yes.
      * Years associated with the ACS period(s)
        + For ACS\_endyr1, set this to be the end year of the 5-year ACS period you want to use. You may want to match the 5-year ACS period to the years of diagnosis for your cancer rates. For example, if the 5-year cancer rates you are presenting are for 2015-2019, set ACS\_endyr1=2019. If you are presenting 10-year cancer rates, you will set this to the end year of the 10-year period.
        + For ACS\_endyr2, you only need to set this if oneACSPeriod=no (i.e., if presenting 10-year cancer rates). Setting the second 5-year ACS period will average the ACS data across two 5-year periods to estimate the ACS demographic statistics across the 10-year period. Set ACS\_endyr2 to be the end year of the 5-year ACS period you want to use (i.e., the 5-year period not covered in ACS\_endyr1). For example, if the 10-year cancer rates you are presenting are for 2010-2019, set ACS\_endyr2=2014.
      * What ACS variables to include (keepACS for period 1, keepACS2 for period 2). Normally these will be the same unless a particular variable is not available in both periods. The keepACS2 parameter values are only used if oneACSPeriod=no.
        + Variable names can be found in the ACS data dictionaries mentioned above.
    1. The data path if different from “C:\Work\WebToolTables”
  + Investigate any warning messages generated by the program as these usually mean something is not working as expected
  + The program creates a table of demographic data needed for the web tool with all geographies combined (e.g., zone, state, national, county), as well as a summary report.