

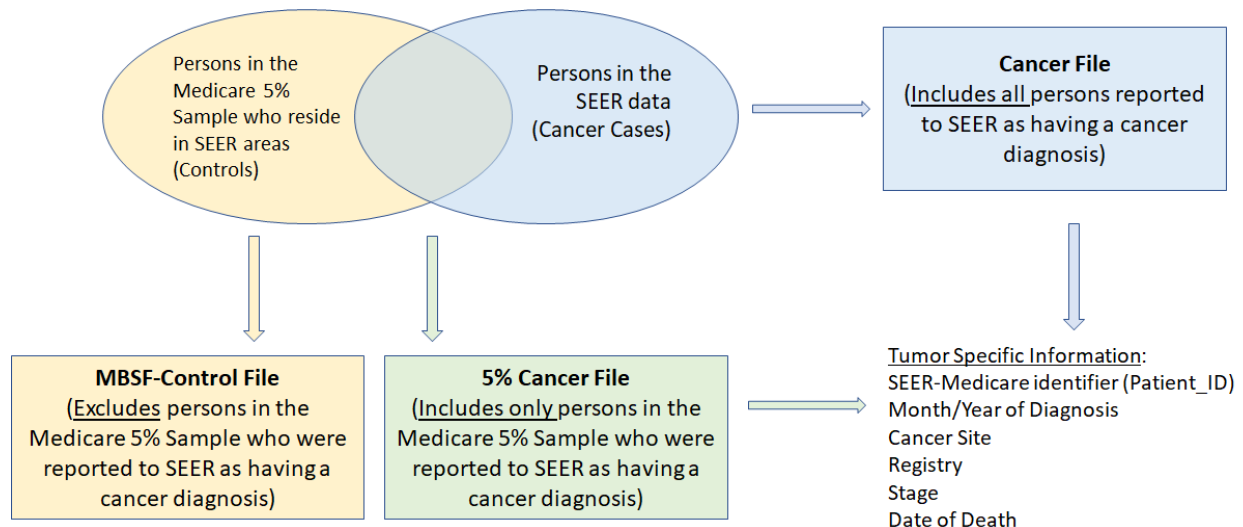
## SEER-Medicare 5% Cancer File

### What is the SEER-Medicare 5% Cancer File?

The SEER-Medicare 5% Cancer File includes individuals who resided in a SEER area and were in the Medicare 5% sample (i.e., included in the MBSF-Control File) but because they were reported to a SEER registry as having had a cancer diagnosis they were then moved over into the cancer cases file (i.e., included Cancer File). The information provided in the SEER-Medicare 5% Cancer File allows the ability to include all individuals who were in the Medicare 5% sample who developed cancer. Therefore, a “true” 5% sample can be recreated at any given time by adding persons included in SEER-Medicare 5% Cancer File with the MBSF-Control File. Information on the date of cancer diagnosis and select tumor characteristics (e.g., cancer site and stage) is available for the included individuals in the SEER-Medicare 5% Cancer File.

### Who is Included in the SEER-Medicare 5% Cancer File?

The SEER-Medicare 5% Cancer File includes all individuals who were included in the SEER-Medicare 5% control sample (i.e., individuals who resided in a SEER area and were included in the 5% Medicare sample and thus in the MBSF-Control File) who have been diagnosed with ANY cancer, as reported to a SEER registry. The file was created by linking the MBSF-Control File with the Cancer File to identify individuals in the SEER-Medicare 5% sample who were ever diagnosed with a cancer according to SEER.



## What Information is Included in the SEER-Medicare 5% Cancer File?

The SEER-Medicare 5% Cancer File includes select cancer characteristics from individuals diagnosed with cancer who were included in the SEER-Medicare 5% control sample (i.e., the MBSF-Control File) including:

- SEER-Medicare Identifier (Patient\_ID)
- Month of diagnosis
- Year of diagnosis
- Cancer site of diagnosis
- Registry of diagnosis
- Historical Stage
- AJCC Stage
- Date of Death

## How Many People are in the SEER-Medicare 5% Cancer File Diagnosis File?

The following table provides an estimate of the number of new cancer diagnoses, by cancer type and year, in the SEER-Medicare 5% Cancer File. Note this table only presents a subset of diagnosis years; the file includes cancer diagnoses back to 1999, depending on when a registry joined the SEER Program.

Cancer Type	Year of Diagnosis														Total
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	
<b>Breast</b>	1039	1060	1054	989	1006	936	863	911	859	862	893	867	955	887	13181
<b>Colorectal</b>	977	947	860	772	755	677	601	589	587	511	530	498	512	458	9274
<b>Lung</b>	1175	1173	1136	1056	1096	1013	931	968	924	887	906	930	908	825	13928
<b>Prostate</b>	1240	1129	1250	1243	1064	1014	1030	967	823	810	758	830	930	888	13976
<b>Other Cancers</b>	3519	3449	3489	3335	3181	3214	3130	3132	3219	3122	3125	3125	3256	3164	45460
<b>Total (Any Cancer)</b>	7950	7758	7789	7395	7102	6854	6555	6567	6412	6192	6212	6250	6561	6222	95819

Table Footnote: Includes individuals 65+ at diagnosis with their first or only cancer diagnosis. Individuals must have been enrolled in Medicare Part A + B and not enrolled in Medicare Advantage in the month of diagnosis.

## How Can Researchers Use the File?

Common uses for the SEER-Medicare 5% Cancer File include:

- 1) **Identifying Controls for a Case-Control Study:** Researchers often use the MBSF-Control File to identify appropriate controls for a study. Individuals are not included in the MBSF-Control File if they have been diagnosed with cancer. The SEER-Medicare 5% Cancer File allows researchers to add individuals who are part of the Medicare 5% sample who have a cancer diagnosis with individuals included in the Medicare 5%

sample who have not been diagnosed with cancer (i.e., individuals included in the MBSF-Control File) to recreate a “true” 5% sample. Knowledge of the cancer type, dates of diagnosis and other select cancer characteristics is available to determine control eligibility.

- 2) **Examining Cancer as an Outcome:** Researchers may be interested in identifying a cancer diagnosis (in general, not a specific cancer type) as a primary study outcome. SEER-Medicare policy does not allow researchers to request the entire data set. Using the SEER-Medicare 5% Cancer File allows researchers to examine the relationship between non-cancer exposures (i.e., cancer screening, Parkinson’s, Stroke) and cancer development within the 5% Medicare sample.

### **Why Would I Want to Use the SEER-Medicare 5% Cancer File?**

- 1) **Identify all Cancers in a Representative Sample of Medicare Beneficiaries:** SEER-Medicare policy does not allow researchers to request the entire data set (i.e., all cancers included in the Cancer File). Using the SEER-Medicare 5% Cancer File allows researchers to identify all individuals who were included in the SEER-Medicare 5% sample who were diagnosed with cancer, as reported to a SEER registry, as well as identify select cancer characteristics for these individuals.
- 2) **Manageable File Sizes:** SEER-Medicare data files can be extremely large and challenging to analyze without large computing capabilities (e.g., servers with large processing memory and storage space). Using a subset of cancers diagnosed in the SEER-Medicare 5% sample allows for more manageable analytic file sizes- particularly for questions evaluating cancer in general as an outcome.
- 3) **Comparable Populations for Comparing Prevalence of Cancer to Other Outcomes:** The Cancer File contain 100% of cancers diagnosed within a SEER area. The MBSF-Control File contains 5% of individuals who reside in SEER areas who are not known to have cancer. Using the SEER-Medicare 5% Cancer File in combination with the MBSF-Control File allows for evaluating comparable relative burdens of disease when comparing cancer prevalence to other disease prevalence (i.e., heart disease) in a population.