National Program of Cancer Registries (NPCR) and Surveillance, Epidemiology & End Results (SEER)

NPCR and SEER Incidence—U.S. Cancer Statistics 2005–2016 Public Use Database Data Standards and Data Dictionary

November, 2018 Submission
Diagnosis Years 2005–2016

U.S. Department of Health and Human Services
Centers for Disease Control and Prevention

NIH NATIONAL CANCER INSTITUTE

CS304502-A

Table of Contents

Message to Data Users

June 3, 2019

We are happy to share the 2019-release of the U.S. Cancer Statistics public use dataset from CDC's National Program of Cancer Registries (NPCR) and the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. This database provides population-based cancer statistics on the *entire* United States population.

The NPCR and SEER Program are comprehensive cancer surveillance systems that work collaboratively with partners to collect, compile, and disseminate information on more than 1.8 million cancer cases annually. U.S. Cancer Statistics data products, like this public use database, are made possible by the dedicated efforts of reporting facilities, cancer registrars, central cancer registries, and CDC NPCR and NCI SEER staff and contractors. I thank everyone for their contributions in collecting these important and high quality data.

Cancer registry data provide a foundation of cancer surveillance activities that are used to measure progress and target cancer prevention and control activities. We encourage researchers to use our data to inform scientific inquiries, programs, and policies. Through use of this U.S. Cancer Statistics public use data source, researchers can positively impact comprehensive cancer prevention and control as well as the care and quality of lives for those diagnosed with cancer.

Sincerely,

Vicki Benard, PhD
Branch Chief, Cancer Surveillance Branch
Division of Cancer Prevention and Control
National Center for Chronic Disease Prevention and Health Promotion
Centers for Disease Control and Prevention

Overview of CDC's National Program of Cancer Registries and NCI's Surveillance, Epidemiology, and End Results Program

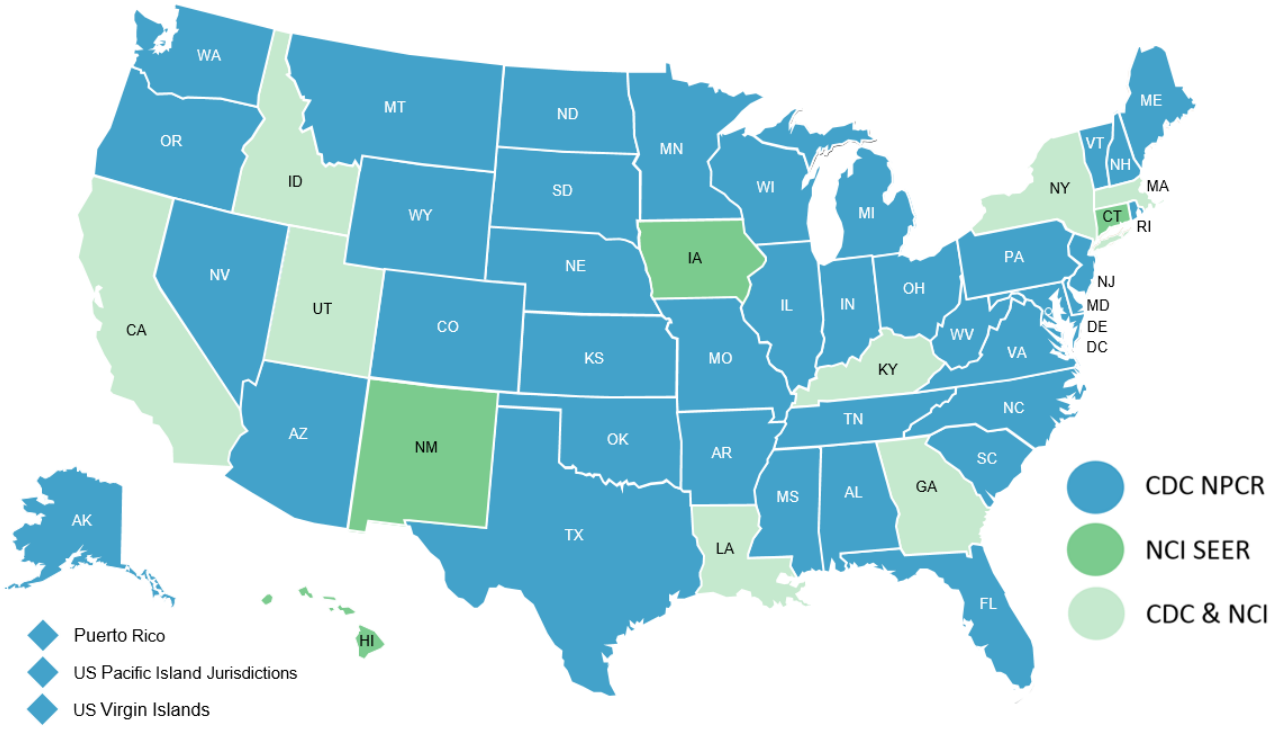
The National Program of Cancer Registries (NPCR), administered by the Centers for Disease Control and Prevention (CDC), was established by Congress in 1992. Through cooperative agreements, NPCR supports central cancer registries in 46 states, the District of Columbia, Puerto Rico, U.S. Pacific Island Jurisdictions, and the U.S. Virgin Islands (see Figure 1 below).

The Surveillance, Epidemiology, and End Results (SEER) Program, administered by the National Cancer Institute (NCI), has been funded since 1973 as a result of the National Cancer Act of 1971. SEER collects reportable cancer cases from 19 U.S. geographic areas, including 12 states (see Figure 1 below). Together, CDC's NPCR and NCI's SEER Program cover the entire United States population. These combined data are the official source of federal statistics on cancer incidence and are referred to as the U.S. Cancer Statistics.

The cancer registries funded by CDC and NCI routinely collect data on patient demographics, primary tumor site, tumor morphology, stage at diagnosis, first course of treatment, and outcomes. Medical facilities such as hospitals, doctor's offices, pathology laboratories, and other treatment centers send demographic and clinical information related to people with cancer to a central cancer registry, where the information is consolidated and goes through rigorous checks for quality and completeness. All hospitals are required by state law to report cancer cases to the central cancer registry in their respective states. On an annual basis, the central cancer registries submit demographic and clinical information about each cancer case to CDC and/or NCI.

This national coverage of cancer data from CDC's NPCR and NCI's SEER Program enables researchers, clinicians, policy makers, public health professionals, and members of the public to monitor the burden of cancer, evaluate the success of programs, and identify additional needs for cancer prevention and control efforts at national, state, and local levels.

Figure 1. Central cancer registry programs submitting data to NPCR and SEER in 2018



Data Collection, Submission, Quality, and Coverage

The most current data come from the November 2018 NPCR and SEER submissions, which cover cancer cases diagnosed from January 1, 2005 through December 31, 2016. Each year, NPCR- and SEER-funded central cancer registries submit data on cancer diagnosed during the most recent year to the respective program. In addition, data from previous years are updated with information from the newly submitted records to ensure case completeness and high quality. NPCR and SEER allow an interval of 23 months after the close of the diagnosis year for submission (for the 2016 data, NPCR required submission by November 30, 2018 and SEER required submission by November 1, 2018).

CDC and NCI support the data collection and quality standards in the North American Association of Central Cancer Registries (NAACCR) consensus documents. During data collection, CDC and NCI also apply additional rigorous quality control edits, data completeness evaluations, and data quality assessments on all NPCR and SEER data. For a registry's data to be included in the U.S. Cancer Statistics public research data file, they must have met the following quality and completeness criteria for publication[1]—

- 5% or fewer cases are ascertained solely on the basis of a death certificate.
- 3% or fewer cases are missing information on sex.
- 3% or fewer cases are missing information on age.
- 5% or fewer cases are missing information on race.
- 97% or more of the registry's records passed a set of single-field and interfield computerized edits.

NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 Public Use Database

Two NPCR and SEER Incidence – U.S. Cancer Statistics public use databases are available for researchers: the 2001–2016 database and the 2005–2016 database. **This data standards document is specific to the 2005–2016 database**.

The 2001–2016 database includes race and ethnicity variables, while the 2005–2016 database does not. The 2005–2016 database includes Puerto Rico data, while the 2001–2016 database does not.

- The 2001–2016 database's population denominators are race-specific, ethnicity-specific, and sex-specific county population estimates from the U.S. Census (July 1, 2010–2017 bridged–race vintage 2017 population estimates), modified by SEER and aggregated to the state and national levels.

- The 2005–2016 database's population denominators are sex-specific, are from the 2010 U.S. Census, and are not available by race or ethnicity.

Table 1 lists the population coverage by diagnosis year for the NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 public research data. In the 2019 release of the public use database there is 100% population coverage for all 50 states, the District of Columbia, and Puerto Rico for cases diagnosed from 2005 through 2016.

[1] Additional information is available at https://www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm

Table 1. U.S. population coverage [a], NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 Public Use Research Database.

Diagnosis year(s)	Percentage of U.S. population covered in database
2005	100%
2006	100%
2007	100%
2008	100%
2009	100%
2010	100%
2011	100%
2012	100%
2013	100%
2014	100%
2015	100%
2016	100%

[a] The NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 Public Use Research Database includes data submitted by all 50 states, the District of Columbia, and Puerto Rico. U.S. Pacific Island Jurisdiction and U.S. Virgin Island data are not included in the database.

Variable List

Table 3 shows all of the variables available in the 2005-2016 public use research database.

Table 3. Variables in the NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 Public Use Research Database

SEER*Stat Category	SEER*Stat Variable Name	Restrictions
Age at Diagnosis	Age recode with <1 year olds	
Race, Sex, Year Dx, Registry, County	Sex	
	Year of diagnosis	
	Addr at DX – state	
	Program	
Site and Morphology	Primary site – labeled	
	Histologic type ICD-O-3	
	Grade	
	Diagnostic confirmation	
	ICD-O-3 hist/behavior, labeled	
	Site recode ICD-O-3/WHO 2008	
	ICCC site recode ICD-O-3/WHO 2008	
	ICCC site rec extended ICD-O-3/WHO 2008	
	AYA site recode/WHO 2008	
	Lymphoma subtype recode/WHO 2008	
	Behavior recode for analysis derived/WHO2008	
Stage – LRD [Summary and Historic]	Merged summary stage 2000	
Therapy	RX summ – surg prim site	Female breast only
Extent of Disease – CS	CS site-specific factor 1	Restricted to 2 groups: - Female breast - Brain and diagnosis years ≥2011
	CS site-specific factor 2	Female breast only
	CS site-specific factor 15	Female breast only and diagnosis years ≥ 2010
	Laterality	
Multiple Primary Fields	Sequence number – central	
Dates	Year of birth	
	Month of diagnosis	
Other	Type of reporting source	
Merged System-Supplied	Alcohol-related cancers	
	HPV-related cancers	
	Obesity-related cancers	
	Physical inactivity-related cancers	
	Tobacco-related cancers	
	State race eth suppress	

Abbreviations used in the variable names –

Addr Address
AYA Adolescent and young adult
CS Collaborative stage

Dx	Diagnosis
Hisp	Hispanic
ICCC	International Classification of Childhood Cancer
ICD-O-3	*International Classification of Diseases for Oncology*, Third Edition
LRD	Local, regional, distant
NHIA	NAACCR Hispanic identification algorithm
USCS	U.S. Cancer Statistics
WHO	World Health Organization

Data Citations

Please use these standard citations for tables and figures when presented in presentations or publications.

- **For population coverage[2]:** Data are from population-based registries that participate in CDC's National Program of Cancer Registries and/or NCI's Surveillance, Epidemiology, and End Results Program and meet high-quality data criteria. These registries cover approximately [XX]% of the U.S. population.
- **For age-adjusted rates:** Rates are per 100,000 persons and are age-adjusted to the 2000 U.S. standard population (19 age groups – Census P25–1130).
- **For the 2005–2016 database:** National Program of Cancer Registries and Surveillance, Epidemiology, and End Results SEER*Stat Database: NPCR and SEER Incidence – U.S. Cancer Statistics Public Use Database, with Puerto Rico, Nov 2018 submission (2005-2016), United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. Released June 2019, based on November 2018 submissions. Available at www.cdc.gov/cancer/public-use.

[2] See Table 1 for percentage population coverage applicable to years being analyzed.

Cautionary Notes

Before using the database, analysts should read and understand the following section. If you have questions regarding these notes, please contact CDC at uscsdata@cdc.gov.

Case Inclusions and Exclusions

NPCR- and SEER-supported cancer registries report all incident cases coded as *in situ* (non malignant) and invasive (malignant; primary site only) according to the *International Classification of Diseases for Oncology, Third Edition* (ICD-O-3), with the following exceptions—

- *In situ* cancers of the cervix are not reported.
- Basal and squamous cell carcinomas of the skin are not reported, except when these occur on the skin of the genital organs.
- Non-malignant (including borderline and benign) central nervous system tumors are reported.
- *In situ* cancers of the urinary bladder are re-coded as invasive behavior because the information needed to distinguish between *in situ* and invasive bladder cancers is not always available or reliable. Stage for these cases remains coded as *in situ*.[1]

Additionally, in this public use database –

- Cancer cases that were identified only through death certificate or autopsy reports have been excluded.
- Cases with an unknown age or with sex other than male or female have been excluded from the database. The frequency counts presented in this document will not change based on whether *Known Age* or *Male or Female Sex* is checked on the SEER*Stat Selection tab.
- *Malignant Behavior* is a default selection for this database, as this restriction is used by CDC's NPCR and NCI's SEER Program for generating most official cancer statistics.

  This database includes *in situ* and nonmalignant central nervous system (CNS) cases. These nonmalignant cases can be analyzed by unselecting the *Malignant Behavior* check box on the SEER*Stat Selection tab.

Suppression Rules[2-3]

Complementary Cell Suppression

When analyzing data at the state or regional levels, counts for national and regional data must be suppressed if a single state in a region or division is suppressed. This practice is referred to as *complementary cell suppression* and is necessary to prevent users from subtracting to find suppressed counts. Rates, confidence intervals (CIs), and populations can be shown at the national and regional levels. This suppression should occur when a single or multiple years of data are being presented.

Suppressing fewer than 16 cases

The suppression rule is fewer than 16 cases for the time period based on rate stability. This suppression rule is enforced automatically in this database.

When the numbers of cases used to compute the incidence rates are small, those rates tend to have poor reliability. Therefore, to discourage misinterpretation and misuse of counts, rates, and trends that are unstable because of the small number of cases, these statistics are not shown in tables and figures if the counts are fewer than 16 for the time period. A count of fewer than about 16 in a numerator results in a standard error of the rate that is about 25% or more as large as the rate itself. Equivalently, a count of fewer than about 16 results in the width of the 95% confidence interval around the rate being at least as large as

the rate itself. These relationships were derived under the assumption of a Poisson process and with the standard population age distribution close to the observed population age distribution.

Another important reason for employing a cell suppression threshold value is to protect the confidentiality of patients whose data are included in a report by reducing or eliminating the risk of identity disclosure. The cell suppression threshold value of 16 is recommended to protect patient confidentiality given the low level of geographic and clinical detail provided.

Case Level Data

As a further mechanism to protect data confidentiality and due to data sharing agreements with some of the states providing data for this database, the case listing function in SEER*Stat has been disabled for this database.

Benign Central Nervous System (CNS) Tumors

Cancer registries began collecting information on nonmalignant brain and other central nervous system tumors with cases diagnosed in 2004. Collection of these tumors is in accordance with Public Law 107-260, the Benign Brain Tumor Cancer Registries Amendment Act, which mandates that NPCR registries collect data on all brain and other central nervous system tumors with a behavior code of 0 (benign) and those with a behavior code of 1 (borderline), in addition to *in situ* and malignant tumors. Data for nonmalignant brain and other nervous system tumors were available from all registries contributing to this report.

Primary Site Variables[4]

Beginning in diagnosis year 2010, some of the lymphoma and leukemia ICD-O-3 codes were updated based on changes from the World Health Organization. The appropriate site recode variables to use to include these updates are *Site recode ICD-O-3/WHO 2008* for all ages and *International Classification of Childhood Cancer (ICCC) site recode ICD-O-3/WHO 2008* and *ICCC site rec extended ICD-O-3/WHO 2008* for the childhood cancer recodes.

Consider reviewing the variable *Site recode ICD-O-3/WHO 2008* before using the directly coded primary site. For more information on the SEER primary site recodes, see http://seer.cancer.gov/siterecode/.

Histologic Type ICD-O-3[5-8]

Beginning with 2010 diagnoses, this item also includes histology codes as specified in the 2008 World Health Organization (WHO) hematopoietic/lymphoid publication, which are listed on pages 3–5 of the NAACCR *2010 Implementation Guidelines and Recommendations*, available at https://www.facs.org/~/media/files/quality%20programs/cancer/coc/2010implementationguidelines.ashx.

Stage

A merged variable, *Merged Summary Stage 2000*, has been created to span three time periods when two different staging schemes were used. Stage at diagnosis is recorded using *SEER Summary Stage 2000* for diagnosis years 2001–2003, *Derived SEER Summary Stage 2000* for diagnosis years 2004–2015, and *SEER Summary Stage 2000* for diagnosis year 2016.

The coding logic for this merged variable is:

- If a case was diagnosed between 2001 and 2003, stage at diagnosis is recorded using the *SEER Summary Stage 2000* variable value.
- If a case was diagnosed between 2004 and 2015, then the stage at diagnosis is recorded using the *Derived SEER Summary Stage 2000* variable value.
- If a case was diagnosed in 2016, then the stage at diagnosis is recorded using the *SEER Summary Stage 2000*.
- If the *Derived SEER Summary Stage 2000* variable is blank and a valid value is available for the *SEER Summary Stage 2000* variable, that value is used to populate the merged variable. For example, if a case was diagnosed in 2015 and *Derived SEER Summary Stage* was blank, but *SEER Summary Stage*

had a value of “local,” then the merged variable was coded as local stage. Otherwise, the merged variable left blank.

This is the stage variable included in both the NPCR 2005–2016 and 2001–2016 public use databases.

Reporting Delay[9]

NPCR and SEER registries annually submit all eligible years of data to CDC and NCI, respectively. As a result, cases submitted in previous years may be deleted, and new cases diagnosed in previous years may be added. The addition of new cases is called a *reporting delay.* For example, reporting of melanoma cases diagnosed in an outpatient facility may be delayed. As a result, the trend in incident melanoma cases might superficially appear to have dropped in the most recent year.

References

1. Young JL Jr, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA (eds). SEER Summary Staging Manual - 2000: Codes and Coding Instructions, National Cancer Institute, NIH Pub. No. 01-4969, Bethesda, MD, 2001.

2. Federal Committee on Statistical Methodology. *Report on Statistical Disclosure Limitations Methodology (Statistical Working Paper 22).* Washington, DC: Office of Management and Budget; 2005. Available at https://fcsm.sites.usa.gov/files/2014/04/spwp22.pdf.

3. Doyle P, Lane JI, Theeuwes JM, Zayatz LM. *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies.* Amsterdam: Elsevier Science; 2001.

4. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin D, et al., editors. *International Classification of Diseases for Oncology.* Third Edition. Geneva: World Health Organization; 2000.

5. *International Classification of Diseases for Oncology.* Third Edition, First Revision. Geneva: World Health Organization, 2013.

6. Ruhl J, Adamo M, Dickie L. (January 2015). *Hematopoietic and Lymphoid Neoplasm Coding Manual*. National Cancer Institute, Bethesda, MD 20850-9765.

7. Surveillance, Epidemiology, and End Results Program. *2007 Multiple Primary and Histology Coding Rules*. Bethesda, MD: US Department of Health and Human Services, National Cancer Institute; Revised August 24, 2012; Accessed January 25, 2017. https://seer.cancer.gov/tools/mphrules/.

8. Surveillance, Epidemiology, and End Results Program. *Hematopoietic and Lymphoid Neoplasm Database*. Bethesda, MD: US Department of Health and Human Services, National Cancer Institute; 2016. https://seer.cancer.gov/seertools/hemelymph.

9. Clegg LX, Feuer EJ, Midthune DN, Fay MP, Hankey BF. Impact of reporting delay and reporting error on cancer incidence rates and trends. *Journal of the National Cancer Institute* 2002;94(20):1537–1545.

Checklist for a NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 Public Use Data Analysis

- ☐ If a user-defined primary site variable was created (rather than using the *Site recode ICD-O-3/WHO 2008* variable):
  - Did you exclude leukemias and lymphomas (9590–9992)?
  - Did you consider excluding Kaposi sarcoma (9140) and mesothelioma (9050–9055)?[1]
- ☐ If your analysis includes histology, and if appropriate for the cancer site, did you use the *Diagnostic Confirmation* variable to specify the analysis be limited to *Microscopically confirmed* cases?[2]
- ☐ If you are analyzing sex-specific cancers (such as prostate cancer or female breast cancer), did you limit the analysis to the appropriate sex to get the correct population denominator?[3]
- ☐ When reporting rates, have you included the label "per 100,000 persons," "per 100,000 women," or "per 100,000 men"?
- ☐ Have you included citations for the:
  - Percentage of United States population coverage provided by the database?
  - NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2016 Public Use Research Database?[4]

---

[1]See Cautionary Notes section entitled *Primary Site Variables*.

[2]See *Diagnostic Confirmation* variable descriptions.

[3]See *Sex* variable description.

[4]See *Data Citation* section.

NPCR and SEER Incidence – U.S. Cancer Statistics 2005 – 2016 Public Use Database: Variable Definition and Frequency

Please note that the frequencies presented in this document were created with *Malignant Behavior* unselected on the SEER*Stat Selection tab.

SEER*Stat Item Name: **Age recode with <1 year olds**
Source of Standard: NAACCR
Source Item Name: Derived from *Age at diagnosis*
Source Item Number: 230

Description

This variable indicates age range at diagnosis by grouping patient into one of 19 categories (0, 1–4, 5–9, …, 75–79, 80–84, ≥85 years). Derived from the NAACCR variable *Age at diagnosis [230]*, which is the age (in years) of the patient at diagnosis.

Considerations for use

Different primary tumors for the same patient may have different values. Records for persons with multiple primary cancers cannot be identified in this database.

Values	Frequency	Percentage
00 years	14,360	0.1%
01–04 years	46,718	0.2%
05–09 years	35,495	0.2%
10–14 years	41,715	0.2%
15–19 years	70,915	0.3%
20–24 years	113,544	0.5%
25–29 years	181,669	0.9%
30–34 years	273,046	1.3%
35–39 years	409,943	2.0%
40–44 years	698,502	3.3%
45–49 years	1,142,964	5.5%
50–54 years	1,752,330	8.4%
55–59 years	2,303,832	11.0%
60–64 years	2,716,255	13.0%
65–69 years	2,958,943	14.1%
70–74 years	2,636,089	12.6%
75–79 years	2,275,749	10.9%
80–84 years	1,750,295	8.4%
85+ years	1,499,149	7.2%

SEER*Stat Item Name: **Sex**

Source of Standard: NAACCR
Source Item Name: Sex
Source Item Number: 220

Description

This variable indicates the sex of the patient.

Considerations for use

- To get the correct population denominator, “female” must be selected when analyzing female-specific cancers (such as ovarian cancer or female breast cancer) and “male” for male-specific cancers (such as prostate cancer).
- Due to small case counts and the lack of an associated population file, cases for sex other than male or female are excluded from this database.

Values	Frequency	Percentage
Male	10,326,654	49.4%
Female	10,594,859	50.6%

SEER*Stat Name: **Year of diagnosis**
Source of Standard: NAACCR
Source Item Name: Derived from *Date of initial diagnosis (CoC)*
Source Item Number: 390

Description

This variable indicates the year of initial diagnosis by a recognized medical practitioner for the cancer being reported, whether clinically or microscopically confirmed. Derived from *Date of initial diagnosis (CoC) [390]*.

Considerations for use

- As an additional confidentiality measure, date of diagnosis is not provided.
- For more information, please see
  - NAACCR data dictionary https://www.naaccr.org/data-standards-data-dictionary
  - FORDS www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals
  - SEER coding manuals http://seer.cancer.gov/tools/codingmanuals/historical.html

Values	Frequency	Percentage
2005	1,585,491	7.6%
2006	1,626,121	7.8%
2007	1,679,978	8.0%
2008	1,704,926	8.1%
2009	1,732,200	8.3%
2010	1,723,979	8.2%
2011	1,770,447	8.5%
2012	1,766,066	8.4%
2013	1,799,794	8.6%
2014	1,824,194	8.7%
2015	1,862,796	8.9%
2016	1,845,521	8.8%

SEER*Stat Item Name: **Addr at DX – State**
Source of Standard: NAACCR
Source Item Name: State at diagnosis (CoC)
Source Item Number: 80

Description

This variable indicates the U.S. state in which the patient lived at the time the reportable tumor was diagnosed.

Considerations for use

- If a patient has multiple tumors, the state of residence at the time of diagnosis may differ for each.
- Multiple records for an individual with more than one primary cancer cannot be linked in this database.
- For more information, please see
  - NAACCR data dictionary https://www.naaccr.org/data-standards-data-dictionary
  - FORDS variable “state at diagnosis” at www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals

Values	Frequency	Percentage
Alaska	35,463	0.2%
Alabama	327,818	1.6%
Arkansas	198,296	0.9%
Arizona	371,375	1.8%
California	2,141,801	10.2%
Colorado	286,308	1.4%
Connecticut	278,061	1.3%
District of Columbia	37,417	0.2%
Delaware	72,795	0.3%
Florida	1,531,036	7.3%
Georgia	595,719	2.8%
Hawaii	90,016	0.4%
Idaho	97,437	0.5%
Illinois	865,684	4.1%
Indiana	430,726	2.1%
Iowa	227,420	1.1%
Kansas	192,470	0.9%
Kentucky	336,526	1.6%
Louisiana	307,796	1.5%
Massachusetts	477,033	2.3%
Maryland	379,834	1.8%
Maine	111,477	0.5%
Michigan	709,540	3.4%
Minnesota	360,015	1.7%
Missouri	411,035	2.0%
Mississippi	196,467	0.9%
Montana	74,164	0.4%
North Carolina	665,087	3.2%
North Dakota	46,528	0.2%
Nebraska	121,905	0.6%

Values	Frequency	Percentage
New Hampshire	104,339	0.5%
New Jersey	665,979	3.2%
New Mexico	115,560	0.6%
Nevada	150,998	0.7%
New York	1,453,603	6.9%
Ohio	808,241	3.9%
Oklahoma	243,471	1.2%
Oregon	265,497	1.3%
Pennsylvania	1,022,337	4.9%
Puerto Rico	80,759	0.4%
Rhode Island	179,130	0.9%
South Carolina	327,215	1.6%
South Dakota	56,377	0.3%
Tennessee	437,479	2.1%
Texas	1,313,364	6.3%
Utah	131,187	0.6%
Virginia	493,489	2.4%
Vermont	48,975	0.2%
Washington	463,050	2.2%
Wisconsin	400,953	1.9%
West Virginia	148,230	0.7%
Wyoming	34,031	0.2%

SEER*Stat Item Name: **Program**

Source of Standard: NPCR
Source Item Name: Not applicable
Source Item Number: Not applicable

Description

This variable indicates whether a state is funded by CDC's NPCR or NCI's SEER Program.

Considerations for use

Central cancer registries that received funding from NPCR and submitted any 2001–2016 diagnosis years data (i.e., Alabama, Alaska, Arizona, Arkansas, California, Colorado, Delaware, District of Columbia, Florida, Georgia, Idaho, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, 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, and Wyoming) are categorized as "NPCR" states.

"SEER" refers to central cancer registries receiving funding only from SEER during the 2001–2016 diagnosis years (i.e., Connecticut, Hawaii, Iowa, and New Mexico).

Values	Frequency	Percentage
NPCR	20,210,456	96.6%
SEER	711,057	3.4%

SEER*Stat Item Name: **Primary Site – labeled**
Source of Standard: NAACCR
Source Item Name: Derived from *Primary site*
Source Item Number: 400

Description

This variable indicates the topography code from *The International Classification of Diseases for Oncology*, Third Edition (ICD-O-3) for the primary site of the tumor being reported.

Considerations for use

- Beginning in diagnosis year 2010, there were updates to some of the lymphoma and leukemia codes. To include these updates, the appropriate primary site variables to use are *Site recode ICD-O-3/WHO 2008* for all ages, and *ICCC site recode ICD-O-3/WHO 2008* for the childhood cancer recodes.
- For more information, please see SEER coding manuals http://seer.cancer.gov/tools/codingmanuals/historical.html.

Values	Frequency	Percentage
C00.0-C77.9 (specific site)	20,596,616	98.4%
C80.9 (Unknown primary site)	324,897	1.6%

SEER*Stat Item Name: **Histologic Type ICD-O-3**
Source of Standard: NAACCR
Source Item Name: Histologic Type ICD-O-3
Source Item Number: 522

Description

This variable indicates the morphology code from *The International Classification of Diseases for Oncology*, Third Edition (ICD-O-3) that describes the histologic type (the microscopic composition of cells and tissue for a specific primary tumor) of the primary tumor being reported.

Considerations for use

- This data item is required for cancer cases diagnosed on or after January 1, 2001.
- The histology codes for some tumors may be based on clinical diagnoses, not pathologic confirmation. When analyzing a specific histology, we suggest using the *Diagnostic Confirmation* variable in conjunction with this variable. Beginning with 2010 diagnoses, this item also includes histology codes as specified in the *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008)*, which are listed on pages 4–6 of the NAACCR 2010 Implementation Guidelines https://www.naaccr.org/wp-content/uploads/2016/11/2010-Implementation-Guidelines-and-Recommendations_Revised-June-2010.pdf.
- For more Information, please see:
  - SEER 2007 Multiple Primary and Histology Coding Rules: https://seer.cancer.gov/tools/mphrules/mphrules_instructions.pdf
  - SEER Hematopoietic Project: https://seer.cancer.gov/tools/heme/
  - ICD-O-3 SEER site/Histology validation list: https://seer.cancer.gov/icd-o-3
  - Surveillance, Epidemiology, and End Results Program. *Hematopoietic and Lymphoid Neoplasm Database*. Bethesda, MD: US Department of Health and Human Services, National Cancer Institute; 2016. https://seer.cancer.gov/seertools/hemelymph
  - Ruhl J, Adamo M, Dickie L. (January 2015). *Hematopoietic and Lymphoid Neoplasm Coding Manual*. National Cancer Institute, Bethesda, MD 20850-9765.
  - *International Classification of Diseases for Oncology*, Third Edition, First Revision. Geneva: World Health Organization, 2013: https://codes.iarc.fr/

Values	Frequency	Percentage
8000–9992	20,921,513	100.0%

SEER*Stat Item Name: **Grade**
Source of Standard: NAACCR
Source Item Name: Grade
Source Item Number: 440

Description

This variable indicates the grade or degree of differentiation of the primary tumor being reported. For lymphomas and leukemias, this field also is used to indicate T-, B-, Null-, or NK-cell origin.

Considerations for use

- The practice of grading varies greatly among pathologists throughout the world, and many malignant tumors are not graded routinely. Since different grading systems may be used, review the site-specific modules available at https://training.seer.cancer.gov/modules_site_spec.html and the most current FORDS manual (www.facs.org/cancer/coc/fordsmanual.html).

  Each module has an abstracting, coding, and staging section, which has a morphology and grading subsection. Some modules, but not all, contain notes about the grading system that may have been used. Currently, there is no variable to differentiate a specific grading system from another one if more than two grading systems are mentioned.
- Diagnostic practices also influence coding practices. For example, preliminary analysis of tumor grade for prostate cancer showed an increase in the proportion of higher grades from 2002 to 2003. Additional review showed this increase to be artificial, as the International Society of Urologic Pathologists, in conjunction with WHO, had made a series of recommendations for modification of the Gleason grading system to reflect contemporary knowledge, alleviate uncertainty, and promote uniformity in its application. One recommendation was for pathologists to report all higher tertiary grade components of the tumor as part of the Gleason score. Another recommendation was made for reporting of any higher-grade cancer, no matter how small quantitatively.

  The percentage of cases with a known grade varies by primary cancer site. Rules for coding the tumor grade differ for some primary sites. As a result, it may be appropriate to have a tumor grade coded as “9 – unknown.”
- For brain tumor cases diagnosed in 2011 and later, cancer registries were required to report the World Health Organization (WHO) Grade Classifcation. Please see the variable description *CS Site-Specific Factor 1* for more information on this brain-specific grade classification.

Values	Frequency	Percentage
Well differentiated; Grade I	1,934,785	9.2%
Moderately differentiated; Grade II	4,954,743	23.7%
Poorly differentiated; Grade III	4,089,830	19.5%
Undifferentiated; anaplastic; Grade IV	671,335	3.2%
T-cell	76,771	0.4%
B-cell; pre-B; B-precursor	1,118,644	5.3%
Null cell; non T-non B	1,790	0.0%
NK cell; natural killer cell (1995+)	3,435	0.0%
Unknown	8,070,180	38.6%

SEER*Stat Item Name: **Diagnostic confirmation**

Source of Standard: NAACCR
Source Item Name: Diagnostic confirmation
Source Item Number: 490

Description

This variable records the best method of diagnostic confirmation of the cancer being reported at any time in the patient's history. The rules for coding differ between solid tumors and hematopoietic and lymphoid neoplasms.

Considerations for use

- For analyses that include histology, it is recommended to use the following selection statement in the SEER*Stat Selection tab: "Diagnostic confirmation is = to Microscopically confirmed".
- Diagnostic confirmation is useful to calculate rates based on microscopically confirmed cancers. Full incidence calculations must also include cases that are only confirmed clinically. The percentage of cases that are "clinically diagnosed only" is an indication of whether case finding includes sources outside of pathology reports.
- The microscopically confirmed method has the highest priority for diagnostic confirmation. The remaining values were assigned when the presence of cancer was confirmed with multiple diagnostic methods.
- "Positive histology AND immunophenotyping AND/OR positive genetic studies" (used only for hematopoietic and lymphoid neoplasms M-9590/3-9992/3) was adopted for use beginning with 2010 diagnoses.
- For more information, please see:
  - FORDS at www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals
  - SEER coding manuals https://seer.cancer.gov/tools/codingmanuals/historical.html

Values	Frequency	Percentage
Microscopically confirmed (total)	19,585,018	93.6%
Positive histology	18,727,822	89.5%
Positive exfoliative cytology, no positive histology	618,881	3.0%
Positive histology AND immunophenotyping AND/OR positive genetic studies	216,402	1.0%
Positive microscopic confirm, method not specified	21,913	0.1%
Positive laboratory test/marker study	87,939	0.4%
Direct visualization without microscopic confirmation	23,718	0.1%
Radiography without microscopic confirm	814,122	3.9%
Clinical diagnosis only	128,846	0.6%
Unknown	281,870	1.3%

SEER*Stat Item Name: **ICD-O-3 Hist/behavior, labeled**

Source of Standard: SEER*Stat recode
Source Item Name: ICD-O-3 Hist/behavior, labeled
Source Item Number: Not applicable

Description

This variable includes each ICD-O-3 histology code and behavior code and the respective name of that histology and behavior.

Considerations for use

- For more information, please see:
  - *International Classification of Diseases for Oncology.* Third Edition, First Revision. Geneva: World Health Organization, 2013.
  - SEER ICD-O-3 Coding Materials https://seer.cancer.gov/icd-o-3

ICD–O–3 Code	Label	Frequency	Percentage
800	Neoplasms, NOS	485,040	2.3%
801–804	Epithelial Neoplasms, NOS	1,137,523	5.4%
805–808	Squamous Cell Neoplasms	1,581,351	7.6%
809–811	Basal Cell Neoplasms	5,902	0.0%
812–813	Transitional Cell Papillomas and Caricnomas	896,979	4.3%
814–838	Adneomas and Adenocarcinomas	8,226,492	39.3%
839–842	Adnexal and Skin Appendage Neoplasms	24,395	0.1%
843	Mucoepidermoid Neoplasms	17,805	0.1%
844–849	Cystic, Mucinous and Serous Neoplasms	538,368	2.6%
850–854	Ductal and Lobular Neoplasms	3,132,296	15.0%
855	Acinar Cell Neoplasms	39,705	0.2%
856–857	Complex Epithelial Neoplasms	71,025	0.3%
858	Thymic Epithelial Neoplasms	9,848	0.0%
859–867	specialized Gonadal Neoplasms	6,091	0.0%
868–871	Paragangliomas and Glomus Tumors	3,546	0.0%
872–879	Nevi and Melanomas	1,468,102	7.0%
880	Soft Tissue Tumors and Sarcomas, NOS	46,358	0.2%
881–883	Fibromatous Neoplasms	42,245	0.2%
884	Myxomatous Neoplasms	1,115	0.0%
885–888	Lipomatous Neoplasms	32,023	0.2%
889–892	Myomatous Neoplasms	48,435	0.2%
893–899	Complex Mixed and Stromal Neoplasms	104,601	0.5%
900–903	Fibroepithelial Neoplasms	5,825	0.0%
904	Synovial-Like Neoplasms	7,366	0.0%
905	Mesothelial Neoplasms	38,214	0.2%
906–909	Germ Cell Neoplasms	109,563	0.5%
910	Trophoblastic Neoplasms	4,673	0.0%
911	Mesonephromas	244	0.0%
912–916	Blood Vessel Tumors	49,174	0.2%
917	Lymphatic Vessel Tumors	238	0.0%
918–924	Osseous and Chondromatous Neoplasms	24,800	0.1%
925	Giant Cell Tumors	1,026	0.0%

ICD–O–3 Code	Label	Frequency	Percentage
926	Miscellaneous Bone Tumors	6,521	0.0%
927–934	Odotogenic Tumors	704	0.0%
935–937	Miscellaneous Tumors	14,093	0.1%
938–948	Gliomas	250,602	1.2%
949–952	Neuroepitheliomatous Neoplasms	22,029	0.1%
953	Meningiomas	336,918	1.6%
954–957	Nerve Sheath Tumors	84,354	0.4%
958	Granular Cell Tumors and Alveolar Soft Part Sarcomas	964	0.0%
959–972	Hodgkin and Non-Hodgkin Lymphomas	862,569	4.1%
973	Plasma Cell Tumors	264,947	1.3%
974	Mast Cell Tumors	2,211	0.0%
975	Neoplasms of Histiocytes and Accessory Lymphoid Cells	6,201	0.0%
976	Immunoproliferative Disease	14,895	0.1%
980–994	Leukemias	567,934	2.7%
995–996	Chronic Myeloproliferative Disorders	128,414	0.6%
997	Other Hematologic Disorders	15,747	0.1%
998–999	Myelodysplastic Syndromes	182,042	0.9%

SEER*Stat Item Name: **Site recode ICD-O-3/WHO 2008**
Source of Standard: NAACCR
Source Item Name: Derived from *Primary site* and *Histologic code ICD-O-3*
Source Item Number: 400 (*Primary site*) and 522 (*Histologic code ICD-O-3*)

Description

This recode variable is defined by the SEER Program. The values of the site recode ICD-O-3/WHO 2008 variable are based on NAACCR variables for ICD-O-3, the primary site and histology code of the primary tumor being reported, with updated information for Hematopoietic codes based on *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008)*. The site recode variables define the major cancer sites that are commonly used in the reporting of cancer incidence data.

Considerations for use

- This is the recommended variable for analyses by primary cancer site.
- More information, including the site recode number, is available at https://seer.cancer.gov/siterecode.

Values	Frequency	Percentage
All Sites (total)	**20,921,513**	100.0%
Oral Cavity and Pharynx	490,723	2.3%
Lip	27,315	0.1%
Tongue	147,652	0.7%
Salivary Gland	51,350	0.2%
Floor of Mouth	26,679	0.1%
Gum and Other Mouth	66,807	0.3%
Nasopharynx	21,864	0.1%
Tonsil	86,931	0.4%
Oropharynx	22,034	0.1%
Hypopharynx	28,415	0.1%
Other Oral Cavity and Pharynx	11,676	0.1%
Digestive System	3,442,845	16.5%
Esophagus	198,542	0.9%
Stomach	273,714	1.3%
Small Intestine	94,587	0.5%
Colon and Rectum	1,813,989	8.7%
Colon excluding Rectum	1,297,827	6.2%
Cecum	276,834	1.3%
Appendix	40,331	0.2%
Ascending Colon	254,582	1.2%
Hepatic Flexure	60,938	0.3%
Transverse Colon	119,965	0.6%
Splenic Flexure	38,637	0.2%
Descending Colon	77,325	0.4%
Sigmoid Colon	345,067	1.6%
Large Intestine, NOS	84,148	0.4%
Rectum and Rectosigmoid Junction	516,162	2.5%
Rectosigmoid Junction	129,090	0.6%
Rectum	387,072	1.9%
Anus, Anal Canal and Anorectum	85,474	0.4%
Liver and Intrahepatic Bile Duct	307,244	1.5%
Liver	271,600	1.3%

Values	Frequency	Percentage
Intrahepatic Bile Duct	35,644	0.2%
Gallbladder	47,117	0.2%
Other Biliary	70,953	0.3%
Pancreas	492,688	2.4%
Retroperitoneum	15,203	0.1%
Peritoneum, Omentum and Mesentery	23,448	0.1%
Other Digestive Organs	19,886	0.1%
Respiratory System	2,725,788	13.0%
Nose, Nasal Cavity and Middle Ear	28,450	0.1%
Larynx	162,457	0.8%
Lung and Bronchus	2,526,433	12.1%
Pleura	1,153	0.0%
Trachea, Mediastinum and Other Respiratory Organs	7,295	0.0%
Bones and Joints	36,164	0.2%
Soft Tissue including Heart	129,886	0.6%
Skin excluding Basal and Squamous	1,497,774	7.2%
Melanoma of the Skin	1,430,193	6.8%
Other Non-Epithelial Skin	67,581	0.3%
Breast (female and male combined)	3,378,602	16.1%
Female Genital System	1,116,539	5.3%
Cervix Uteri	153,523	0.7%
Corpus and Uterus, NOS	582,926	2.8%
Corpus Uteri	566,053	2.7%
Uterus, NOS	16,873	0.1%
Ovary	256,754	1.2%
Vagina	18,002	0.1%
Vulva	79,970	0.4%
Other Female Genital Organs	25,364	0.1%
Male Genital System	2,631,858	12.6%
Prostate	2,502,149	12.0%
Testis	101,077	0.5%
Penis	23,775	0.1%
Other Male Genital Organs	4,857	0.0%
Urinary System	1,568,772	7.5%
Urinary Bladder	845,841	4.0%
Kidney and Renal Pelvis	672,799	3.2%
Ureter	34,170	0.2%
Other Urinary Organs	15,962	0.1%
Eye and Orbit	38,827	0.2%
Brain and Other Nervous System	724,002	3.5%
Brain	288,834	1.4%
Cranial Nerves Other Nervous System	435,168	2.1%
Endocrine System	709,297	3.4%
Thyroid	523,962	2.5%
Other Endocrine including Thymus	185,335	0.9%
Lymphoma	890,786	4.3%
Hodgkin Lymphoma	104,365	0.5%
Hodgkin – Nodal	101,733	0.5%
Hodgkin – Extranodal	2,632	0.0%

Values	Frequency	Percentage
Non-Hodgkin Lymphoma	786,421	3.8%
NHL – Nodal	531,238	2.5%
NHL – Extranodal	255,183	1.2%
Myeloma	262,554	1.3%
Leukemia	543,948	2.6%
Lymphocytic Leukemia	271,568	1.3%
Acute Lymphocytic Leukemia	59,164	0.3%
Chronic Lymphocytic Leukemia	195,347	0.9%
Other Lymphocytic Leukemia	17,057	0.1%
Myeloid and Monocytic Leukemia	244,082	1.2%
Acute Myeloid Leukemia	157,379	0.8%
Acute Monocytic Leukemia	9,070	0.0%
Chronic Myeloid Leukemia	70,694	0.3%
Other Myeloid/Monocytic Leukemia	6,939	0.0%
Other Leukemia	28,298	0.1%
Other Acute Leukemia	9,453	0.0%
Aleukemic, Subleukemic and NOS	18,845	0.1%
Mesothelioma	38,214	0.2%
Kaposi Sarcoma	15,021	0.1%
Miscellaneous	679,913	3.2%

SEER*Stat Item Name: **ICCC site recode ICD-O-3/WHO 2008**
Source of Standard: SEER
Source Item Name: Derived from NAACCR *Primary site*, *Histologic code ICD-O-3*, and *Behavior code ICD-O-3*
Source Item Number: NAACCR 400 (*Primary site*), 522 (*Histologic code ICD-O-3*), and 523 (*Behavior code ICD-O-3*)

Description

This variable indicates the classification of childhood cancer, which is based on tumor morphology and primary site, and emphasizes morphology rather than primary site, as is done for adults.

Considerations for use

- This recode is defined by the SEER Program, which was based on definitions presented in Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. *International Classification of Childhood Cancer*, Third Edition.
- For comparison of *ICCC site recode extended ICD-O-3/WHO 2008* and this variable see https://seer.cancer.gov/iccc/iccc-who2008.html.
- Additional information is available at https://seer.cancer.gov/iccc/iccc3.html and https://seer.cancer.gov/iccc/.

Note: This frequency table is restricted to individuals 19 years old or younger.

Values	Frequency	Percentage
I Leukemias, myeloproliferative & myelodysplastic diseases	47,786	22.8%
I(a) Lymphoid leukemias	34,055	16.3%
I(b) Acute myeloid leukemias	8,178	3.9%
I(c) Chronic myeloproliferative diseases	2,508	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,527	0.7%
I(e) Unspecified and other specified leukemias	1,518	0.7%
II Lymphomas and reticuloendothelial neoplasms	27,919	13.3%
II(a) Hodgkin lymphomas	12,558	6.0%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	9,271	4.4%
II(c) Burkitt lymphoma	2,445	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	3,290	1.6%
II(e) Unspecified lymphomas	355	0.2%
III CNS and misc intracranial and intraspinal neoplasms	49,197	23.5%
III(a) Ependymomas and choroid plexus tumor	3,922	1.9%
III(b) Astrocytomas	17,157	8.2%
III(c) Intracranial and intraspinal embryonal tumors	6,069	2.9%
III(d) Other gliomas	5,814	2.8%
III(e) Other specified intracranial/intraspinal neoplasms	14,287	6.8%
III(f) Unspecified intracranial and intraspinal neoplasms	1,948	0.9%
IV Neuroblastoma and other peripheral nervous cell tumors	8,745	4.2%
IV(a) Neuroblastoma and ganglioneuroblastoma	8,498	4.1%
IV(b) Other peripheral nervous cell tumors	247	0.1%
V Retinoblastoma	3,266	1.6%
VI Renal tumors	7,093	3.4%
VI(a) Nephroblastoma and other nonepithelial renal tumors	6,347	3.0%
VI(b) Renal carcinomas	722	0.3%
VI(c) Unspecified malignant renal tumors	24	0.0%

Values	Frequency	Percentage
VII Hepatic tumors	2,434	1.2%
VII(a) Hepatoblastoma	1,817	0.9%
VII(b) Hepatic carcinomas	595	0.3%
VII(c) Unspecified malignant hepatic tumors	22	0.0%
VIII Malignant bone tumors	9,094	4.3%
VIII(a) Osteosarcomas	5,084	2.4%
VIII(b) Chondrosarcomas	332	0.2%
VII(c) Ewing tumor and related sarcomas of bone	3,016	1.4%
VIII(d) Other specified malignant bone tumors	472	0.2%
VIII(e) Unspecified malignant bone tumors	190	0.1%
IX Soft tissue and other extraosseous sarcomas	12,287	5.9%
IX(a) Rhabdomyosarcomas	4,784	2.3%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	1,305	0.6%
IX(c) Kaposi sarcoma	52	0.0%
IX(d) Other specified soft tissue sarcomas	4,788	2.3%
IX(e) Unspecified soft tissue sarcomas	1,358	0.6%
X Germ cell & trophoblastic tumors & neoplasms of gonads	11,984	5.7%
X(a) Intracranial & intraspinal germ cell tumors	2,228	1.1%
X(b) Extracranial & extragonadal germ cell tumors	1,466	0.7%
X(c) Malignant gonadal germ cell tumors	7,496	3.6%
X(d) Gonadal carcinomas	460	0.2%
X(e) Other and unspecified malignant gonadal tumors	334	0.2%
XI Other malignant epithelial neoplasms and melanomas	20,069	9.6%
XI(a) Adrenocortical carcinomas	209	0.1%
XI(b) Thyroid carcinomas	9,125	4.4%
XI(c) Nasopharyngeal carcinomas	548	0.3%
XI(d) Malignant melanomas	4,695	2.2%
XI(e) Skin carcinomas	85	0.0%
XI(f) Other and unspecified carcinomas	5,407	2.6%
XII Other and unspecified malignant neoplasms	772	0.4%
XII(a) Other specified malignant tumors	417	0.2%
XII(b) Other unspecified malignant tumors	355	0.2%
Not classified by ICCC or *in situ*	8,557	4.1%

SEER*Stat Item Name: **ICCC site rec extended ICD-O-3/WHO 2008**
Source of Standard: SEER
Source Item Name: Derived from NAACCR *Primary site*, *Histologic code ICD-O-3*, and *Behavior code ICD-O-3*
Source Item Number: NAACCR 400 (*Primary site*), 522 (*Histologic code ICD-O-3*), and 523 (*Behavior code ICD-O-3*)

Description

This variable indicates the classification of childhood cancer, which is based on tumor morphology and primary site, and emphasizes morphology rather than primary site, as is done for adults. This variable contains extended classification codes of childhood cancer, based on definitions presented in *International Classification of Childhood Cancer, Third Edition (ICCC, 3rd Edition)* and *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008)*.

Considerations for use

- For comparison of *ICCC site recode ICD-O-3/WHO 2008* and this variable see https://seer.cancer.gov/iccc/iccc-who2008.html.
- Additional information is available at http://seer.cancer.gov/iccc/iccc3_ext.html and https://seer.cancer.gov/iccc/.

Note: This frequency table is restricted to individuals 19 years old or younger.

Values	Frequency	Percentage
I Leukemias, myeloproliferative & myelodysplastic diseases	47,786	22.8%
I(a) Lymphoid leukemias	34,055	16.3%
I(a.1) Precursor cell leukemias	32,913	15.7%
I(a.2) Mature B-cell leukemias	900	0.4%
I(a.3) Mature T-cell and NK cell leukemias	119	0.1%
I(a.4) Lymphoid leukemia, NOS	123	0.1%
I(b) Acute myeloid leukemias	8,178	3.9%
I(c) Chronic myeloproliferative diseases	2,508	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,527	0.7%
I(e) Unspecified and other specified leukemias	1,518	0.7%
II Lymphomas and reticuloendothelial neoplasms	27,919	13.3%
II(a) Hodgkin lymphomas	12,558	6.0%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	9,271	4.4%
II(b.1) Precursor cell lymphomas	2,730	1.3%
II(b.2) Mature B-cell lymphomas except Burkitt lymphoma	3,731	1.8%
II(b.3) Mature T-cell and NK-cell lymphomas	2,323	1.1%
II(b.4) Non-Hodgkin lymphomas, NOS	487	0.2%
II(c) Burkitt lymphoma	2,445	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	3,290	1.6%
II(e) Unspecified lymphomas	355	0.2%
III CNS and misc intracranial and intraspinal neoplasms	49,197	23.5%
III(a) Ependymomas and choroid plexus tumor	3,922	1.9%
III(a.1) Ependymomas	2,905	1.4%
III(a.2) Choroid plexus tumor	1,017	0.5%
III(b) Astrocytomas	17,157	8.2%
III(c) Intracranial and intraspinal embryonal tumors	6,069	2.9%

Values	Frequency	Percentage
III(c.1) Medulloblastomas	4,017	1.9%
III(c.2) PNET	1,110	0.5%
III(c.3) Medulloepithelioma	64	0.0%
III(c.4) Atypical teratoid/rhabdoid tumor	878	0.4%
III(d) Other gliomas	5,814	2.8%
III(d.1) Oligodendrogliomas	621	0.3%
III(d.2) Mixed and unspecified gliomas	5,075	2.4%
III(d.3) Neuroepithelial glial tumors of uncertain orig	118	0.1%
III(e) Other specified intracranial/intraspinal neoplasms	14,287	6.8%
III(e.1) Pituitary adenomas and carcinomas	6,142	2.9%
III(e.2) Tumors of sellar region (craniopharyngiomas)	2,053	1.0%
III(e.3) Pineal parenchymal tumors	480	0.2%
III(e.4) Neuronal and mixed neuronal-glial tumors	4,082	2.0%
III(e.5) Meningiomas	1,530	0.7%
III(f) Unspecified intracranial and intraspinal neoplasms	1,948	0.9%
IV Neuroblastoma and other peripheral nervous cell tumors	8,745	4.2%
IV(a) Neuroblastoma and ganglioneuroblastoma	8,498	4.1%
IV(b) Other peripheral nervous cell tumors	247	0.1%
V Retinoblastoma	3,266	1.6%
VI Renal tumors	^1	^1
VI(a) Nephroblastoma and other nonepithelial renal tumors	^1	^1
VI(a.1) Nephroblastoma	5,966	2.9%
VI(a.2) Rhabdoid renal tumor	166	0.1%
VI(a.3) Kidney sarcomas	207	0.1%
VI(a.4) pPNET of kidney	^2	^2
VI(b) Renal carcinomas	722	0.3%
VI(c) Unspecified malignant renal tumors	24	0.0%
VII Hepatic tumors	2,434	1.2%
VII(a) Hepatoblastoma	1,817	0.9%
VII(b) Hepatic carcinomas	595	0.3%
VII(c) Unspecified malignant hepatic tumors	22	0.0%
VIII Malignant bone tumors	9,094	4.3%
VIII(a) Osteosarcomas	5,084	2.4%
VIII(b) Chondrosarcomas	332	0.2%
VIII(c) Ewing tumor and related sarcomas of bone	3,016	1.4%
VIII(c.1) Ewing tumor and Askin tumor of bone	2,893	1.4%
VIII(c.2) pPNET of bone	123	0.1%
VIII(d) Other specified malignant bone tumors	472	0.2%
VIII(d.1) Malignant fibrous neoplasms of bone	42	0.0%
VIII(d.2) Malignant chordomas	223	0.1%
VIII(d.3) Odontogenic malignant tumors	60	0.0%
VIII(d.4) Miscellaneous malignant bone tumors	147	0.1%
VIII(e) Unspecified malignant bone tumors	190	0.1%
IX Soft tissue and other extraosseous sarcomas	^1	^1
IX(a) Rhabdomyosarcomas	4,784	2.3%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	^1	^1
IX(b.1) Fibroblastic and myofibroblastic tumors	694	0.3%

Values	Frequency	Percentage
IX(b.2) Nerve sheath tumors	594	0.3%
IX(b.3) Other fibromatous neoplasms	^2	^2
IX(c) Kaposi sarcoma	52	0.0%
IX(d) Other specified soft tissue sarcomas	4,788	2.3%
IX(d.1) Ewing tumor and Askin tumor of soft tissue	650	0.3%
IX(d.2) pPNET of soft tissue	270	0.1%
IX(d.3) Extrarenal rhabdoid tumor	287	0.1%
IX(d.4) Liposarcomas	290	0.1%
IX(d.5) Fibrohistiocytic tumors	1,118	0.5%
IX(d.6) Leiomyosarcomas	185	0.1%
IX(d.7) Synovial sarcomas	1,057	0.5%
IX(d.8) Blood vessel tumors	200	0.1%
IX(d.9) Osseous & chondromatous neoplasms of soft tissue	96	0.0%
IX(d.10) Alveolar soft parts sarcoma	167	0.1%
IX(d.11) Miscellaneous soft tissue sarcomas	468	0.2%
IX(e) Unspecified soft tissue sarcomas	1,358	0.6%
X Germ cell & trophoblastic tumors & neoplasms of gonads	^1	^1
X(a) Intracranial & intraspinal germ cell tumors	2,228	1.1%
X(a.1) Intracranial & intraspinal germinomas	1,301	0.6%
X(a.2) Intracranial & intraspinal teratomas	634	0.3%
X(a.3) Intracranial & intraspinal embryonal carcinomas	34	0.0%
X(a.4) Intracranial & intraspinal yolk sac tumor	31	0.0%
X(a.5) Intracranial & intraspinal choriocarcinoma	24	0.0%
X(a.6) Intracranial & intraspinal tumors of mixed forms	204	0.1%
X(b) Extracranial & extragonadal germ cell tumors	^1	^1
X(b.1) Germinomas: extracranial/extragonadal	158	0.1%
X(b.2) Malignant teratomas: extracranial/extragonadal	556	0.3%
X(b.3) Embryonal carcinomas: extracranial/extragonadal	^2	^2
X(b.4) Yolk sac tumor: extracranial/extragonadal	331	0.2%
X(b.5) Choriocarcinomas: extracranial/extragonadal	157	0.1%
X(b.6) Other mixed germ cell: extracranial/extragonadal	248	0.1%
X(c) Malignant gonadal germ cell tumors	^1	^1
X(c.1) Malignant gonadal germinomas	1,575	0.8%
X(c.2) Malignant gonadal teratomas	1,254	0.6%
X(c.3) Gonadal embryonal carcinomas	697	0.3%
X(c.4) Gonadal yolk sac tumor	695	0.3%
X(c.5) Gonadal choriocarcinoma	73	0.0%
X(c.6) Malignant gonadal tumors of mixed forms	3,201	1.5%
X(c.7) Malignant gonadal gonadoblastoma	^2	^2
X(d) Gonadal carcinomas	460	0.2%
X(e) Other and unspecified malignant gonadal tumors	334	0.2%
XI Other malignant epithelial neoplasms and melanomas	20,069	9.6%
XI(a) Adrenocortical carcinomas	209	0.1%
XI(b) Thyroid carcinomas	9,125	4.4%
XI(c) Nasopharyngeal carcinomas	548	0.3%
XI(d) Malignant melanomas	4,695	2.2%
XI(e) Skin carcinomas	85	0.0%

Values	Frequency	Percentage
XI(f) Other and unspecified carcinomas	5,407	2.6%
XI(f.1) Carcinomas of salivary glands	934	0.4%
XI(f.2) Carcinomas of colon and rectum	600	0.3%
XI(f.3) Carcinomas of appendix	1,359	0.6%
XI(f.4) Carcinomas of lung	461	0.2%
XI(f.5) Carcinomas of thymus	77	0.0%
XI(f.6) Carcinomas of breast	178	0.1%
XI(f.7) Carcinomas of cervix uteri	129	0.1%
XI(f.8) Carcinomas of bladder	288	0.1%
XI(f.9) Carcinomas of eye	27	0.0%
XI(f.10) Carcinomas of other specified sites	1,214	0.6%
XI(f.11) Carcinomas of unspecified site	140	0.1%
XII Other and unspecified malignant neoplasms	^[1]	^[1]
XII(a) Other specified malignant tumors	^[1]	^[1]
XII(a.1) Gastrointestinal stromal tumor	88	0.0%
XII(a.2) Pancreatoblastoma	38	0.0%
XII(a.3) Pulmonary blastoma and pleuropulmonary blastoma	213	0.1%
XII(a.4) Other complex mixed and stromal neoplasms	42	0.0%
XII(a.5) Mesothelioma	36	0.0%
XII(a.6) Other specified malignant tumors	^[2]	^[2]
XII(b) Other unspecified malignant tumors	355	0.2%
Not classified by ICCC or *in situ*	8,557	4.1%

[1] Values are not reported due to the need for complementary cell suppression.

[2] Counts of fewer than 16 cases and the corresponding percentages are suppressed.

SEER*Stat Item Name: **AYA site recode ICD-O-3/WHO 2008**
Source of Standard: SEER
Source Item Name: Derived from NAACCR *Primary site*, *Histologic code ICD-O-3*, and *Behavior code ICD-O-3*
Source Item Number: NAACCR 400 (*Primary site*), 522 (*Histologic code ICD-O-3*), and 523 (*Behavior code ICD-O-3*)

Description

This variable was based on ICD-O-3, updated for Hematopoietic codes based on *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008)*. It was developed to better define the major cancer sites that affect adolescents or young adults between 15 and 29 years of age.

Considerations for use

This recode variable is defined by the SEER Program. It was based on definitions proposed by RD Barr and colleagues in Barr RD, Holowaty EJ, Birch JM. Classification Scheme for tumors diagnosed in adolescents and young adults. *Cancer* 2006;106(7):1425–30. For more information see https://seer.cancer.gov/ayarecode.

Note: This frequency table is restricted to individuals 15 to 29 years old.

Values	Frequency	Percentage
1 Leukemias	22,304	6.1%
1.1 Acute lymphoid leukemia	8,625	2.4%
1.2 Acute myeloid leukemia	8,319	2.3%
1.3 Chronic myeloid leukemia	3,609	1.0%
1.4 Other and unspecified leukemia	1,751	0.5%
2 Lymphomas	50,847	13.9%
2.1 Non-Hodgkin lymphoma	20,056	5.5%
2.2 Hodgkin lymphoma	30,791	8.4%
3 CNS and Oth Intracranial and Intraspinal Neo (all behav)	34,238	9.4%
3.1. Astrocytoma	10,573	2.9%
3.1.1 Specified low-grade astrocytic tumors	4,084	1.1%
3.1.2 Glioblastoma and anaplastic astrocytoma	4,063	1.1%
3.1.3 Astrocytoma, NOS	2,426	0.7%
3.2 Other glioma	5,619	1.5%
3.3 Ependymoma	2,214	0.6%
3.4. Medulloblastoma and other PNET	1,674	0.5%
3.4.1 Medulloblastoma	1,018	0.3%
3.4.2 Supratentorial PNET	656	0.2%
3.5 Other specified intracranial and intraspinal neoplasms	12,122	3.3%
3.6 Unspecified intracranial and intraspinal neoplasms	2,036	0.6%
3.6.1 Unspec malignant intracranial and intraspinal neo	301	0.1%
3.6.2 Unspec ben/border intracran. and intraspin neo	1,735	0.5%
4 Osseous & Chondromatous Neoplasms	8,834	2.4%
4.1 Osteosarcoma	3,769	1.0%
4.2 Chondrosarcoma	1,177	0.3%
4.3 Ewing tumor	3,036	0.8%
4.4 Other specified and unspecified bone tumors	852	0.2%
5 Soft Tissue Sarcomas	14,853	4.1%

Values	Frequency	Percentage
5.1 Fibromatous neoplasms	3,523	1.0%
5.2 Rhabdomyosarcoma	1,663	0.5%
5.3 Other soft tissue sarcoma	9,667	2.6%
5.3.1 Specified soft tissue sarcoma	7,426	2.0%
5.3.1.1 Specified (excluding Kaposi sarcoma)	5,864	1.6%
5.3.1.2 Kaposi sarcoma	1,562	0.4%
5.3.2 Unspecified soft tissue sarcoma	2,241	0.6%
6 Germ Cell and Trophoblastic Neoplasms	42,115	11.5%
6.1 Germ cell and trophoblastic neoplasms of gonads	38,480	10.5%
6.2 Germ cell and trophoblastic neo of nongonadal sites	3,635	1.0%
6.2.1 Intracranial (all behaviors)	1,405	0.4%
6.2.2 Other nongonadal	2,230	0.6%
7 Melanoma and Skin Carcinomas	30,547	8.3%
7.1 Melanoma	30,263	8.3%
7.2 Skin carcinomas	284	0.1%
8 Carcinomas	109,450	29.9%
8.1 Thyroid carcinoma	48,466	13.2%
8.2 Other carcinoma of head and neck	5,846	1.6%
8.2.1 Nasopharyngeal carcinoma	1,002	0.3%
8.2.2 Other sites in lip, oral cavity and pharynx	4,295	1.2%
8.2.3 Nasal cav, mid ear, sinus, larynx, ill-def head/neck	549	0.1%
8.3 Carcinoma of trachea, bronchus, and lung	2,468	0.7%
8.4 Carcinoma of breast	13,444	3.7%
8.5 Carcinoma of genitourinary tract	20,871	5.7%
8.5.1 Carcinoma of kidney	4,694	1.3%
8.5.2 Carcinoma of bladder	1,730	0.5%
8.5.3 Carcinoma of gonads	3,232	0.9%
8.5.4 Carcinoma of cervix and uterus	10,667	2.9%
8.5.5 Carc of oth and ill-defined sites	548	0.1%
8.6 Carcinoma of gastrointestinal tract	16,521	4.5%
8.6.1 Carcinoma of colon and rectum	11,236	3.1%
8.6.2 Carcinoma of stomach	1,534	0.4%
8.6.3 Carcinoma of liver and intrahepatic bile ducts	1,435	0.4%
8.6.4 Carcinoma of pancreas	1,216	0.3%
8.6.5 Carc oth and ill-def sites, gastrointestinal tract	1,100	0.3%
8.7 Carcinoma of other and ill-defined sites	1,834	0.5%
8.7.1 Adrenocortical carcinoma	312	0.1%
8.7.2 Carcinoma of other and ill-defined sites, NOS	1,522	0.4%
9 Miscellaneous specified neoplasms, NOS	8,643	2.4%
9.1 Other pediatric and embryonal tumors, NOS	834	0.2%
9.1.1 Wilms tumor	175	0.0%
9.1.2 Neuroblastoma	229	0.1%
9.1.3 Other pediatric and embryonal tumors, NOS	430	0.1%
9.2 Other specified and embryonal tumors, NOS	7,809	2.1%
9.2.1 Paraganglioma and glomus tumors	296	0.1%
9.2.2 Other specified gonadal tumors	611	0.2%
9.2.3 Myeloma, mast cell, misc lymphoreticular neo, NOS	1,510	0.4%

Values	Frequency	Percentage
9.2.4 Other specified neoplasms, NOS	5,392	1.5%
10 Unspecified Malignant Neoplasms	1,922	0.5%
Unclassified and Non-Malignant	42,375	11.6%

SEER*Stat Item Name: **Lymphoma subtype recode/WHO 2008**
Source of Standard: SEER
Source Item Name: Derived from NAACCR *Primary site*, *Histologic code ICD-O-3*, and *Behavior code ICD-O-3*
Source Item Number: NAACCR 400 (*Primary site*), 522 (*Histologic code ICD-O-3*), and 523 (*Behavior code ICD-O-3*)

Description

This variable was based on ICD-O-3, updated for Hematopoietic codes based on *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008)*. It was designed to facilitate epidemiologic studies of lymphoma subtypes.

Considerations for use

- This recode variable is defined by the SEER Program. It was adapted from a proposed nested classification of lymphoid neoplasms in:
  Morton LM, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood* 2007;110:695–708.
- This variable is recommended to be used for analyses where the years of diagnosis are 2008 or later as it provides the most up-to-date definitions of lymphoma.
- More information is available at https://seer.cancer.gov/lymphomarecode.

Values	Frequency	Percentage
Lymphoid Neoplasm	1,448,592	6.9%
1 Hodgkin Lymphoma	104,365	0.5%
1(a) Classical Hodgkin lymphoma	97,994	0.5%
1(a)1 Lymphocyte-rich/mixed cell/lymphocyte depleted	16,178	0.1%
1(a)1.1 Lymphocyte-rich	3,782	0.0%
1(a)1.2 Mixed cellularity	11,122	0.1%
1(a)1.3 Lymphocyte-depleted	1,274	0.0%
1(a)2 Nodular sclerosis	53,833	0.3%
1(a)3 Classical Hodgkin lymphoma, NOS	27,983	0.1%
1(b) Nodular lymphocyte predominant Hodgkin lymphoma	6,371	0.0%
2 Non-Hodgkin lymphoma	1,306,537	6.2%
2(a) Non-Hodgkin lymphoma, B-cell	1,209,787	5.8%
2(a)1 Precursor Non-Hodgkin lymphoma, B-cell	49,145	0.2%
2(a)2 Mature Non-Hodgkin lymphoma, B-cell	1,094,534	5.2%
2(a)2.1 Chronic/Sm/Prolymphocytic/Mantle B-cell NHL	272,457	1.3%
2(a)2.1.1 Chronic/Small lymphocytic leuk/lymph	237,343	1.1%
2(a)2.1.2 Prolymphocytic leukemia, B-cell	1,071	0.0%
2(a)2.1.3 Mantle-cell lymphoma	34,043	0.2%
2(a)2.2 Lymphoplasmacytic lymphoma/Waldenstrom	25,210	0.1%
2(a)2.2.1 Lymphoplasmacytic lymphoma	10,747	0.1%
2(a)2.2.2 Waldenstrom macroglubulinemia	14,463	0.1%
2(a)2.3 Diffuse large B-cell lymphoma (DLBCL)	278,910	1.3%
2(a)2.3.1 DLBCL, NOS	275,644	1.3%
2(a)2.3.2 Intravascular large B-cell lymphoma	548	0.0%
2(a)2.3.3 Primary effusion lymphoma	414	0.0%

Values	Frequency	Percentage
2(a)2.3.4 Mediastinal large B-cell lymphoma	2,304	0.0%
2(a)2.4 Burkitt lymphoma/leukemia	15,313	0.1%
2(a)2.5 Marginal-zone lymphoma (MZL)	77,444	0.4%
2(a)2.5.1 Splenic MZL	7,358	0.0%
2(a)2.5.2 Extranodal MZL, MALT type	45,915	0.2%
2(a)2.5.3 Nodal MZL	24,171	0.1%
2(a)2.6 Follicular lymphoma	150,153	0.7%
2(a)2.7 Hairy-cell leukemia	11,543	0.1%
2(a)2.8 Plasma cell neoplasms	263,301	1.3%
2(a)2.8.1 Plasmacytoma	16,986	0.1%
2(a)2.8.2 Multiple myeloma/plasma-cell leuk	246,315	1.2%
2(a)2.9 Heavy chain disease	203	0.0%
2(a)3 Non-Hodgkin lymphoma, B-cell, NOS	66,108	0.3%
2(b) Non-Hodgkin lymphoma, T-cell	85,314	0.4%
2(b)1 Precursor Non-Hodgkin lymphoma, T-cell	2,687	0.0%
2(b)2 Mature Non-Hodgkin lymphoma, T-cell	82,335	0.4%
2(b)2.1 Mycosis fungoides/Sezary syndrome	18,269	0.1%
2(b)2.1.1 Mycosis fungoides	17,547	0.1%
2(b)2.1.2 Sezary syndrome	722	0.0%
2(b)2.2 Peripheral T-cell lymphoma	44,864	0.2%
2(b)2.2.1 Peripheral T-cell lymphoma, NOS	16,299	0.1%
2(b)2.2.2 Angioimmunoblastic T-cell lymphoma	5,425	0.0%
2(b)2.2.3 Subcutan panniculitis-like T-cell lymph	436	0.0%
2(b)2.2.4 Anaplastic lar cell lymph, T-/Null-cell	9,291	0.0%
2(b)2.2.5 Hepatosplenic T-cell lymphoma	402	0.0%
2(b)2.2.6 Enteropathy-type T-cell lymphoma	556	0.0%
2(b)2.2.7 Cutaneous T-cell lymphoma, NOS	8,997	0.0%
2(b)2.2.8 Prim cutaneous anaplastic lar cell lymph	3,458	0.0%
2(b)2.3 Adult T-cell leukemia/lymphoma	9,779	0.0%
2(b)2.4 NK/T-cell lymph, nasal-type/aggres NK leuk	2,550	0.0%
2(b)2.5 T-cell large granular lymphocytic leukemia	5,275	0.0%
2(b)2.6 Prolymphocytic leukemia, T-cell	1,598	0.0%
2(b)3 Non-Hodgkin lymphoma, NOS, T-cell	292	0.0%
2(c) Non-Hodgkin lymphoma, unknown lineage	11,436	0.1%
2(c)1 Precursor lymphoblastic leuk/lymph, unk lineage	3,969	0.0%
2(c)2 Prolymphocytic leukemia, unknown lineage	417	0.0%
2(c)3 Non-Hodgkin lymphoma, NOS, unknown lineage	7,050	0.0%
3 Composite Hodgkin lymphoma and NHL	3,344	0.0%
4 Lymphoid neoplasm, NOS	34,346	0.2%
Unclassified	19,472,921	93.1%

SEER*Stat Item Name: **Behavior Recode for analysis derived/WHO2008**
Source of Standard: NAACCR
Source Item Name: Behavior code ICD-O-3
Source Item Number: 523

Description

This variable is the behavior code from *The International Classification of Diseases for Oncology*, Third Edition (ICD-O-3) for the primary tumor being reported. With the updates and implementation of ICD-O, some histologies changed from malignant behavior to benign or borderline malignancy and vice versa. This Behavior Recode variable takes those changes into consideration.

"Malignant" indicates a histology whose behavior did not change. "Only malignant in ICD-O-3" indicates histologies whose behavior changed to malignant in edition ICD-O-3. Additional changes to histologic behavior were made effective with cases diagnosed in 2010 and later; new histologies were added, while the behavior of others changed (primarily affecting hematopoetic malignancies). "Only malignant 2010+" limits the analysis to those histologies.

Considerations for use

- This database includes cases with invasive (malignant) and *in situ* behavior, benign/borderline behavior for brain/CNS cases diagnosed in 2004 and forward are also included. Caution should be used to select the correct behavior for analyses. Malignant behavior (specifically the "Malignant" category) is the default selection for cases in this database in SEER*Stat. If necessary for the analysis, "Only malignant in ICD-O-3" or "Only malignant 2010+" may be selected to further restrict case selection. If an analysis requires cases with *in situ* behavior, the "Malignant Only" selection should be unchecked on the "Selection" tab.
- Behavior code ICD-O-3 is required for cancer cases diagnosed on or after January 1, 2001. Statistics generated from the database using this behavior code variable are comparable to the methodology used to generate the U.S. Cancer Statistics official federal cancer statistics.
- For more information, please see SEER coding manual at http://seer.cancer.gov/icd-o-3.

Values	Frequency	Percentage
Benign	568,356	2.7%
Borderline malignancy	54,981	0.3%
In situ	1,479,658	7.1%
Malignant	18,485,512	88.4%
Only malignant in ICD-O-3	308,294	1.5%
Only malignant 2010+	24,712	0.1%

SEER*Stat Item Name: **Merged Summary Stage 2000**
Source of Standard: NPCR
Source Item Name: Combined from *Derived SS2000* and *SEER Summary Stage 2000*
Source Item Number: Derived from NAACCR 3020 (*Derived SS2000*) and 759 (*SEER Summary Stage 2000*)

Description

This is a merged stage variable created using two variables: *SEER Summary Stage 2000*, which records stage from diagnosis years 2001–2003, *Derived SS2000*, which records stage from diagnostic years 2004–2015, and *SEER Summary Stage 2000* for diagnostic year 2016. This stage variable can be used for diagnosis years 2001–2016.

Considerations for use

- The coding logic for this merged variable is:
  - If a case was diagnosed between 2001 and 2003, the Summary Stage 2000 variable value was used.
  - If a case was diagnosed between 2004 and 2015, then the Derived Summary Stage 2000 (Derived SS2000) variable was used.
  - If a case was diagnosed in 2016, then the *SEER Summary Stage 2000* variable was used.
  - If the Derived Summary Stage 2000 variable was blank and a valid value was available for the Summary Stage 2000 variable, that value was used to populate the merged variable. For example, if the case was diagnosed in 2013 and Derived Summary Stage was blank, but Summary Stage had a value of local, then the merged variable was coded as local stage. Otherwise, the merged variable would be left blank.
- For more information about SEER Summary Stage 2000 and Derived SS2000 variables, see https://cancerstaging.org/cstage/Pages/default.aspx.

Values	Frequency	Percentage
In situ	1,897,013	9.1%
Localized only	8,449,962	40.4%
Regional, direct extension only	1,366,846	6.5%
Regional, regional lymph nodes only	1,481,565	7.1%
Regional, direct extension and regional lymph nodes	856,870	4.1%
Regional, NOS	195,072	0.9%
Distant site(s)/node(s) involved	4,638,087	22.2%
Not applicable	603,890	2.9%
Unknown/unstaged/unspecified	1,431,779	6.8%
Blanks(s)	429	0.0%

SEER*Stat Item Name: **RX Summ – Surg Prim Site**
Source of Standard: SEER / CoC
Source Item Name: *RX Summ—Surg Prim Site*
Source Item Number: 1290

Description

Site-specific codes for the type of surgery to the primary site performed as part of the first course of treatment. This includes treatment given at all facilities as part of the first course of treatment.

Considerations for use

- Data for this variable are available for **female breast.**
- For breast surgery codes, refer to the *SEER Program Coding and Staging Manual 2016, Appendix C: Site Specific Coding Modules, Breast Surgery Codes* – https://seer.cancer.gov/archive/manuals/2016/AppendixC/Surgery_Codes_Breast_2016.pdf
- In **addition to the site-specific codes, refer to the most recent version of *FORDS* and SEER Program Code manual for additional instructions:**
  - *FORDS* manual - https://www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals
  - SEER Program Code manual - https://seer.cancer.gov/tools/codingmanuals

Note: This frequency table is restricted to female breast cases.

Values	Frequency	Percentage
00-99	3,352,135	100.0%

SEER*Stat Item Name: **CS Site-Specific Factor 1**
Source of Standard: AJCC
Source Item Name: *CS Site-Specific Factor 1*
Source Item Number: 2880

Description

The information recorded in *CS Site-Specific Factor 1* (SSF1) differs for each anatomic site and there is site-specific codes and coding structures for each anatomic site. In the U.S. Cancer Statistics Incidence Analytic database, SSF1 records information represents:

- **Female breast site: Estrogen Receptor (ER) Assay**
- **Brain site: World Health Organization (WHO) Grade Classification**.

Considerations for use

- Data for this variable are available for:
  - **Female breast**
  - **Brain** starting with **diagnosis year 2011**.
- For the site-specific codes, please refer to the Collaborative Stage Data Collection System
  - **Female breast:** Breast Estrogen Receptor Assay, available at https://web2.facs.org/cstage0205/breast/Breast_jag.html.
  - **Brain**: World Health Organization (WHO) Grade Classification, available at https://web2.facs.org/cstage0205/brain/Brain_jpo.html.

Note: This frequency table is restricted to female breast cases.

Values	Frequency	Percentage
00-99	3,336,557	99.5%
Blanks(s)	15,578	0.5%

Note: This frequency table is restricted to brain cases and diagnosis years ≥2011.

Values	Frequency	Percentage
00-99	148,712	99.9%
Blank(s)	214	0.1%

SEER*Stat Item Name: **CS Site-Specific Factor 2**
Source of Standard: AJCC
Source Item Name: *CS Site-Specific Factor 2*
Source Item Number: 2890

Description

The information recorded in CS Site-Specific Factor 2 (SSF2) differs for each anatomic site and there is site-specific codes and coding structures for each anatomic site.

In the U.S. Cancer Statistics Incidence Analytic database, SSF2 records information for **female breast**, specifically, **Progesterone Receptor (PR) Assay**.

Considerations for use

- Data for this variable are available for **female breast**
- Please refer to the Collaborative Stage Data Collection System for the specific codes for CS SSF2, Breast Progesterone Receptor Assay, available at http://web2.facs.org/cstage0205/breast/Breast_kac.html.

Note: This frequency table is restricted to female breast cases.

Values	Frequency	Percentage
00-99	3,336,553	99.5%
Blanks(s)	15,582	0.5%

SEER*Stat Item Name: **CS Site-Specific Factor 15**
Source of Standard: AJCC
Source Item Name: *CS Site-Specific Factor 15*
Source Item Number: 2869

Description

The information recorded in CS Site-Specific Factor 15 (SSF15) differs for each anatomic site and there is site-specific codes and coding structures for each anatomic site.

In the U.S. Cancer Statistics Incidence Analytic database, SSF2 records information for **female breast**, specifically, **Human Epidermal Growth Factor Receptor 2 (HER2): Summary Result of Testing**.

Considerations for use

- Data for this variable are available for **female breast** starting with **diagnosis year 2010**.
- Please refer to the Collaborative Stage Data Collection System for the specific codes for CS SSF3, Breast HER2, available at http://web2.facs.org/cstage0205/breast/Breast_sbg.html.

Note: This frequency table is restricted to female breast cases and diagnosis years ≥2010.

Values	Frequency	Percentage
00-99	2,035,198	99.8%
Blanks(s)	4,492	0.2%

SEER*Stat Item Name: **Laterality**

Source of Standard: NAACCR
Source Item Name: Laterality at Diagnosis (SEER)
Source Item Number: 410

Description

This variable identifies the side of a paired organ, or the side of the body on which the reportable tumor originated. This applies to the primary site only.

Considerations for use

For more information, please see:

- FORDS at www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals
- SEER coding manuals http://seer.cancer.gov/tools/codingmanuals/historical.html

Values	Frequency	Percentage
Not a paired site	11,349,682	54.2%
Right - origin of primary	4,645,882	22.2%
Left - origin of primary	4,312,619	20.6%
Only one side - side unspecified	36,698	0.2%
Bilateral, single primary	153,373	0.7%
Paired site: midline tumor	61,000	0.3%
Paired site, but no information concerning laterality	362,259	1.7%

SEER*Stat Item Name: **Sequence Number – Central**
Source of Standard: NAACCR
Source Item Name: Sequence Number – Central Revised
Source Item Number: 380

Description

This variable indicates the sequence of all reportable neoplasms over the patient's lifetime.

Considerations for use

- The sequence number may change over the patient's lifetime. If the patient was diagnosed with a single reportable neoplasm, and later diagnosed with a second reportable neoplasm, the sequence code for the first neoplasm changes from 00 to 01. A central registry may find that a patient with one or more known neoplasms had an earlier reportable neoplasm that had been unknown to the registry. Typically, a re-evaluation of all related sequence numbers is required whenever an additional neoplasm is identified.
- Standards define which neoplasms are reportable. Please see the SEER list at https://seer.cancer.gov/tools/casefinding/. It is assumed that these standards are the minimum definition of reportability. Individual central cancer registries may define additional neoplasms as reportable. Variability of assigning sequence numbers over time may exist for different registries, which may impact the coding of this variable.
- Because the time period of Sequence Number is a person's lifetime, reportable neoplasms not included in the central registry (those that occur outside the registry catchment area or before the reference date) also are allotted a sequence number. For example, a registry may contain a single record for a patient with a sequence number of 02 because the first reportable neoplasm preceded the central registry's reference date.
- If two or more reportable neoplasms are diagnosed at the same time, the lowest sequence number is assigned to the diagnosis with the worst prognosis. If no difference in prognosis is evident, the decision is arbitrary.
- Reportable non-malignant tumors diagnosed on or after January 1, 2004 are represented by sequence numbers labeled as "…state registry-defined neoplasm". Timing rules for sequencing these neoplasms are the same as timing rules for sequencing required *in situ* or invasive neoplasms.
- The *2007 Multiple Primary and Histology Coding Rules* may also affect the sequence number. For more information, please see https://seer.cancer.gov/tools/mphrules/mphrules_instructions.pdf.
- For more information, please see the SEER coding manual at https://seer.cancer.gov/tools/codingmanuals/historical.html.

Values	Frequency	Percentage
One primary only	14,971,085	^[1]
1st of 2 or more primaries	1,582,023	^[1]
2nd of 2 or more primaries	3,067,584	^[1]
3rd of 3 or more primaries	544,371	^[1]
4th of 4 or more primaries	100,361	^[1]
5th of 5 or more primaries	21,592	^[1]
6th or more primaries[2]	10,733	^[1]
Only one state registry-defined neoplasm	595,782	^[1]
1st of 2 or more state registry-defined neoplasms	11,989	^[1]
2nd of 2 or more state registry-defined neoplasms	13,992	^[1]
3rd of 3 or more state registry-defined neoplasms	992	^[1]
4th of 4 or more state registry-defined neoplasms	258	^[1]
5th of 5 or more state registry-defined neoplasms	104	^[1]
6th or more state registry-defined neoplasms[1]	66	^[1]

Values	Frequency	Percentage
Unknown sequence number - federally required *in situ* or malignant tumors	114	^[1]
Carcinoma *in situ* of the Cervix diagnosed 1/1/1996 or later	^[3]	^[3]
Unknown sequence number - state registry-defined neoplasms	311	^[1]

[1] Values are not reported due to the need for complementary cell suppression.

[2] Subsequent primaries (7 or higher) were collapsed into this category.

[3] Counts of fewer than 16 cases and the corresponding percentages are suppressed.

SEER*Stat Item Name: **Year of Birth**
Source of Standard: SEER / CoC
Source Item Name: Date of Birth
Source Item Number: 240

Description

Year of birth of the patient.

Considerations for use

- The month and day of birth are not provided for confidentiality reasons.
- If age at diagnosis and year of diagnosis are known, but year of birth is unknown, the year of birth should be calculated and so coded. Only the year is entered. Per the NAACCR Data Dictionary, registrars are instructed to estimate a date of birth rather than leave the birth date unknown.
- **This variable includes only count data**. Rates cannot be calculated using this variable, as no population data are associated with the variable.

Values	Frequency	Percentage
1890	^1	^1
1891	^1	^1
1892	^1	^1
1893	^1	^1
1894	^1	^1
1895	^1	^1
1896	^1	^1
1897	^1	^1
1898	^1	^1
1899	31	^2
1900	115	^2
1901	132	^2
1902	158	^2
1903	295	^2
1904	445	^2
1905	816	^2
1906	1,285	^2
1907	2,116	^2
1908	3,355	^2
1909	5,137	^2
1910	7,847	^2
1911	11,266	^2
1912	16,814	^2
1913	22,535	^2
1914	31,316	^2
1915	40,173	^2
1916	52,238	^2
1917	67,314	^2
1918	86,008	^2

Values	Frequency	Percentage
1919	102,844	^2
1920	135,152	^2
1921	165,687	^2
1922	187,759	^2
1923	216,206	^2
1924	248,019	^2
1925	271,056	^2
1926	298,574	^2
1927	330,690	^2
1928	349,741	^2
1929	365,840	^2
1930	395,402	^2
1931	400,974	^2
1932	416,560	^2
1933	412,895	^2
1934	442,462	^2
1935	458,715	^2
1936	468,722	^2
1937	485,590	^2
1938	505,657	^2
1939	504,538	^2
1940	518,012	^2
1941	536,234	^2
1942	581,612	^2
1943	584,891	^2
1944	541,894	^2
1945	515,232	^2
1946	588,961	^2
1947	627,814	^2
1948	570,323	^2
1949	540,385	^2
1950	507,484	^2
1951	497,652	^2
1952	483,324	^2
1953	459,814	^2
1954	448,482	^2
1955	424,570	^2
1956	409,339	^2
1957	390,911	^2
1958	360,515	^2
1959	338,188	^2
1960	314,684	^2
1961	290,095	^2

Values	Frequency	Percentage
1962	264,451	^2
1963	241,872	^2
1964	219,341	^2
1965	190,681	^2
1966	168,719	^2
1967	150,366	^2
1968	138,654	^2
1969	130,267	^2
1970	123,201	^2
1971	108,677	^2
1972	93,671	^2
1973	81,992	^2
1974	76,153	^2
1975	69,742	^2
1976	64,169	^2
1977	60,212	^2
1978	55,713	^2
1979	53,593	^2
1980	49,762	^2
1981	46,047	^2
1982	42,371	^2
1983	38,102	^2
1984	35,060	^2
1985	31,982	^2
1986	28,935	^2
1987	26,335	^2
1988	24,693	^2
1989	22,175	^2
1990	20,615	^2
1991	18,436	^2
1992	16,320	^2
1993	14,631	^2
1994	13,178	^2
1995	11,635	^2
1996	10,633	^2
1997	9,744	^2
1998	9,569	^2
1999	8,819	^2
2000	8,464	^2
2001	8,328	^2
2002	8,364	^2
2003	8,609	^2
2004	8,979	^2

Values	Frequency	Percentage
2005	8,800	^[2]
2006	8,581	^[2]
2007	8,092	^[2]
2008	7,622	^[2]
2009	6,764	^[2]
2010	6,171	^[2]
2011	5,368	^[2]
2012	4,667	^[2]
2013	3,998	^[2]
2014	2,798	^[2]
2015	1,747	^[2]
2016	700	^[2]
Blank(s)	^[1]	^[1]

[1] Values are not reported due to the need for complementary cell suppression.

[2] Counts of fewer than 16 cases and the corresponding percentages are suppressed.

SEER*Stat Item Name: **Month of Diagnosis**

Source of Standard: NAACCR
Source Item Name: Derived from *Date of initial diagnosis (CoC)*
Source Item Number: 390

Description

This variable is derived from *Date of initial diagnosis*, which is the date of initial diagnosis by a recognized medical practitioner for the cancer being reported, whether clinically or microscopically confirmed.

Considerations for use

- The day of diagnosis is not provided as an additional confidentiality measure.
- **This variable includes only count data**. Rates cannot be calculated using this variable, as no population data are associated with the variable.

Values	Frequency	Percentage
January	1,796,255	8.6%
February	1,617,436	7.7%
March	1,773,306	8.5%
April	1,735,527	8.3%
May	1,742,316	8.3%
June	1,792,183	8.6%
July	1,692,309	8.1%
August	1,756,910	8.4%
September	1,674,849	8.0%
October	1,764,892	8.4%
November	1,627,259	7.8%
December	1,613,775	7.7%
Blank(s)	334,496	1.6%

SEER*Stat Item Name: **Type of Reporting Source**
Source of Standard: NAACCR
Source Item Name: Type of reporting source
Source Item Number: 500

Description

This variable codes the source documents used to abstract the majority of information on the tumor being reported. This may not be the source of original casefinding (for example, if a case is identified through a pathology laboratory report review and all source documents used to abstract the case are from the physician's office, code 4 is assigned).

Considerations for use

- For cancers diagnosed prior to 2006, only the following categories were available for *Type of Reporting Source*:

Code	Definition
1	Hospital inpatient/outpatient or clinic
3	Laboratory only (hospital or private)
4	Physician's office/private medical practitioner (local medical doctor)
5	Nursing/convalescent home/hospice

  Since there was no code specific to radiation treatment facilities or ambulatory surgery centers, they were coded as either 1 or 4.

- For cancers diagnosed in 2006 and later, the following codes were added to allow identification of cases reported by radiation treatment centers, medical oncology clinics, and other hospital outpatient units/surgery centers.

Code	Definition
2	Radiation treatment centers, medical oncology clinics
8	Other hospital outpatient units/surgery centers

- Cases were coded in the following priority order: 1, 2, 8, 4, 3, 5. This reflects the addition of codes 2 and 8 and prioritizes laboratory reports over nursing home reports. The source facilities included in the previous code 1 (hospital inpatient and outpatient) are split between codes 1, 2, and 8.

Values	Frequency	Percentage
Hospital inpatient/outpatient or clinic	17,900,685	85.6%
Radiation treatment or medical oncology center (2006+)	526,078	2.5%
Laboratory only (hospital or private)	588,448	2.8%
Physician's office/private medical practitioner (LMD)	981,369	4.7%
Nursing/convalescent home/hospice	25,591	0.1%
Other hospital outpatient unit or surgery center (2006+)	899,342	4.3%

SEER*Stat Item Name: **Alcohol-Related Cancers**
Source of Standard: NPCR
Source Item Name: Derived from NAACCR's *Primary Site*, *Histologic Type ICD-O-3*, *Sex*
Source Item Number: Derived from NAACCR's 400 (*Primary Site*), 522 (*Histologic Type ICD-O-3*), 220 (*Sex*)

Description

Predefined variable created using ICD-O-3 site, histology and sex to define alcohol-related cancers[3,4].

Considerations for use

- Cancer registries do not routinely collect data on alcohol use, so the number of cancers associated with this risk factor cannot be determined definitively[5,6,7].
- However, other sources of information can be used to obtain the proportion of cancers probably caused by the risk factor, also known as the *attributable fraction*[8]. Then the number of *attributable* cancers can be estimated by multiplying the attributable fraction by the number of *associated* cancers.
- For more information on risk-factor related cancers, please see the publication listed below and online for the variable definitions:
  - Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive Cancer Incidence and Survival — United States, 2013. MMWR Morb Mortal Wkly Rep 2017;66:69–75. DOI: http://dx.doi.org/10.15585/mmwr.mm6603a1.
  - *Predefined SEER*Stat Variables for Calculating the Number of Associated Cancers for Selected Risk Factors* documentation available at https://www.cdc.gov/cancer/public-use.

  Please note that in official federal cancer statistics publications, CDC reports malignant (invasive) cancers with the exception of bladder cancers (includes *in situ* and invasive cancers).

Values	Frequency	Percentage
Lip, oral cavity, pharynx	490,723	7.8%
Esophagus	198,542	3.2%
Colon and rectum	1,813,989	28.8%
Liver	271,600	4.3%
Larynx	162,457	2.6%
Female breast cancer	3,352,135	53.3%

[3] International Agency for Research on Cancer. IARC monographs on the evaluation of carcinogenic risks to humans: Volume 96: Alcohol Consumption and Ethyl Carbamate. Lyon, France: International Agency for Research on Cancer; 2010. Available at http://monographs.iarc.fr/ENG/Monographs/vol96/.
[4] International Agency for Research on Cancer. IARC monographs on the evaluation of carcinogenic risks to humans: Volume 100E: Personal Habits and Indoor Combustions: Consumption of Alcoholic Beverages. Lyon, France: International Agency for Research on Cancer; 2012. Available at http://monographs.iarc.fr/ENG/Monographs/vol100E/.
[5] Henley S.J., Singh S.D., King J, et al. Invasive cancer incidence and survival—United States, 2013. *MMWR* 2017.
[6] Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *Journal of the National Cancer Institute* 2011;103(24):1827–1839.
[7] World Cancer Research Fund / American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective.* Washington, DC: AICR, 2007.
[8] Levine B. What does the population attributable fraction mean? *Preventing Chronic Disease* 2007;4(1):A14.

SEER*Stat Item Name: **HPV-Related Cancers**

Source of Standard: NPCR

Source Item Name: Derived from NAACCR's *Primary Site*, *Histologic Type ICD-O-3*, *Sex*, *Diagnostic Confirmation*

Source Item Number: Derived from NAACCR's 400 (*Primary Site*), 522 (*Histologic Type ICD-O-3*), 220 (*Sex*), 490 (*Diagnostic Confirmation*)

Description

Predefined variable created using ICD-O-3 site, histology and sex to define Human Papillomavirus (HPV)-related cancers[9,10,11,12,13].

Considerations for use

- Cancer registries do not routinely collect data on HPV-diagnoses, so the number of cancers associated with this risk factor cannot be determined definitively[14,15,16].
- However, other sources of information can be used to obtain the proportion of cancers probably caused by the risk factor, also known as the *attributable fraction*[17]. Then the number of *attributable* cancers can be estimated by multiplying the attributable fraction by the number of *associated* cancers.
- For more information on risk-factor related cancers, please see the following publication listed below and online for the variable definitions:
  - Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive Cancer Incidence and Survival — United States, 2013. MMWR Morb Mortal Wkly Rep 2017;66:69–75. DOI: http://dx.doi.org/10.15585/mmwr.mm6603a1.
  - *Predefined SEER*Stat Variables for Calculating the Number of Associated Cancers for Selected Risk Factors* documentation available at https://www.cdc.gov/cancer/public-use.

  Please note that in official federal cancer statistics publications, CDC reports malignant (invasive) cancers with the exception of bladder cancers (includes *in situ* and invasive cancers).

Values	Frequency	Percentage
Oropharyngeal squamous cell carcinoma	200,869	38.2%
Anal and rectal squamous cell carcinoma	82,179	15.6%
Vulvar squamous cell carcinoma	62,008	11.8%
Vaginal squamous cell carcinoma	11,889	2.3%

[9] Watson M, Saraiya M, Ahmed F, Cardinez CJ, Reichman ME, Weir HK, Richards TB. Using population-based cancer registry data to assess the burden of human papillomavirus-associated cancers in the United States: overview of methods. *Cancer* 2008;113(10 Suppl):2841–2854. Available at www.ncbi.nlm.nih.gov/pubmed/18980203.

[10] Saraiya M, Unger ER, Thompson TD, Lynch CF, Hernandez BY, Lyu CW, Steinau M, Watson M, Wilkinson EJ, Hopenhayn C, Copeland G, Cozen W, Peters ES, Huang Y, Saber MS, Altekruse S, Goodman MT; HPV Typing of Cancers Workgroup. US assessment of HPV types in cancers: implications for current and 9-valent HPV vaccines. *Journal of the National Cancer Institute* 2015;107(6):djv086. Available at www.ncbi.nlm.nih.gov/pubmed/25925419.

[11] International Agency for Research on Cancer. IARC monographs on the evaluation of carcinogenic risks to humans. Volume 90: Human Papillomaviruses. Lyon, France: International Agency for Research on Cancer; 2007. Available at http://monographs.iarc.fr/ENG/Monographs/vol90/.

[12] Viens LJ, Henley SJ, Watson M, Markowitz LE, Thomas CC, Thompson TD, Razzaghi H, Saraiya M, Centers for Disease Control and Prevention (CDC). Human papillomavirus–associated cancers—United States, 2008–2012. *MMWR* 2016;65(26):661–666. Available at www.cdc.gov/mmwr/volumes/65/wr/mm6526a1.htm.

[13] Centers for Disease Control and Prevention. How Many Cancers Are Linked with HPV Each Year? Atlanta, GA: U.S. Department of Health and Human Services. Available at www.cdc.gov/cancer/hpv/statistics/cases.htm.

[14] Henley S.J., Singh S.D., King J, et al. Invasive cancer incidence and survival—United States, 2013. *MMWR* 2017.

[15] Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *Journal of the National Cancer Institute* 2011;103(24):1827–1839.

[16] World Cancer Research Fund / American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective.* Washington, DC: AICR, 2007.

[17] Levine B. What does the population attributable fraction mean? *Preventing Chronic Disease* 2007;4(1):A14.

Penile squamous cell carcinoma	22,257	4.2%
Cervical carcinoma	147,220	28.0%

SEER*Stat Item Name: **Obesity-Related Cancers**
Source of Standard: NPCR
Source Item Name: Derived from NAACCR's *Primary Site*, *Histologic Type ICD-O-3*, *Sex*, *Diagnostic Confirmation*, Age at diagnosis
Source Item Number: Derived from NAACCR's 400 (*Primary Site*), 522 (*Histologic Type ICD-O-3*), 220 (*Sex*), 490 (*Diagnostic Confirmation*), 230 (*Age at diagnosis*)

Description

Predefined variable created using ICD-O-3 site, histology and sex to define obesity-related cancers[18,19,20].

Considerations for use

- Cancer registries do not routinely collect data on obesity, so the number of cancers associated with this risk factor cannot be determined definitively[21,22,23].
- However, other sources of information can be used to obtain the proportion of cancers probably caused by the risk factor, also known as the *attributable fraction*[24]. Then the number of *attributable* cancers can be estimated by multiplying the attributable fraction by the number of *associated* cancers.
- For more information on risk-factor related cancers, please see the publication listed below and online for the variable definitions:
  - Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive Cancer Incidence and Survival — United States, 2013. MMWR Morb Mortal Wkly Rep 2017;66:69–75. DOI: http://dx.doi.org/10.15585/mmwr.mm6603a1.
  - *Predefined SEER*Stat Variables for Calculating the Number of Associated Cancers for Selected Risk Factors* documentation available at https://www.cdc.gov/cancer/public-use.

  Please note that in official federal cancer statistics publications, CDC reports malignant (invasive) cancers with the exception of bladder cancers (includes *in situ* and invasive cancers).

Values	Frequency	Percentage
Esophageal adenocarcinoma	123,904	1.5%
Gastric cardia	85,266	1.1%
Colon & rectum	1,813,989	22.5%
Liver	271,600	3.4%
Gallbladder	47,117	0.6%
Pancreas	492,688	6.1%
Kidney	624,460	7.7%
Meningioma	332,727	4.1%

[18] Eheman C, Henley SJ, Ballard-Barbash R, Jacobs EJ, Schymura MJ, Noone AM, Pan L, Anderson RN, Fulton JE, Kohler BA, Jemal A, Ward E, Plescia M, Ries LA, Edwards BK. Annual report to the nation on the status of cancer, 1975–2008, featuring cancers associated with excess weight and lack of sufficient physical activity. *Cancer* 2012;118:2338–2366. Available at www.ncbi.nlm.nih.gov/pmc/articles/PMC4586174/.
[19] World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington, DC: AICR, 2007. Available at http://preventcancer.aicr.org/site/PageServer?pagename=research_science_expert_report.
[20] Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body fatness and cancer—viewpoint of the IARC Working Group. N Engl J Med 2016;375:794-798.
[21] Henley S.J., Singh S.D., King J, et al. Invasive cancer incidence and survival—United States, 2013. *MMWR* 2017.
[22] Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *Journal of the National Cancer Institute* 2011;103(24):1827–1839.
[23] World Cancer Research Fund / American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective.* Washington, DC: AICR, 2007.
[24] Levine B. What does the population attributable fraction mean? *Preventing Chronic Disease* 2007;4(1):A14.

Thyroid	523,962	6.5%
Multiple myeloma	245,089	3.0%
Post-menopausal female breast	2,667,487	33.1%
Corpus and uterus, NOS (not otherwise specified)	582,926	7.2%
Ovary	256,754	3.2%

SEER*Stat Item Name: **Physical Inactivity-Related Cancers**
Source of Standard: NPCR
Source Item Name: Derived from NAACCR's *Primary Site*, *Histologic Type ICD-O-3*, *Sex*
Source Item Number: Derived from NAACCR's 400 (*Primary Site*), 522 (*Histologic Type ICD-O-3*), 220 (*Sex*)

Description

Predefined variable created using ICD-O-3 site, histology and sex to define physical inactivity-related cancers[25,26].

Considerations for use

- Cancer registries do not routinely collect data on physical inactivity, so the number of cancers associated with this risk factor cannot be determined definitively[27,28,29].
- However, other sources of information can be used to obtain the proportion of cancers probably caused by the risk factor, also known as the *attributable fraction*[30]. Then the number of *attributable* cancers can be estimated by multiplying the attributable fraction by the number of *associated* cancers.
- For more information on risk-factor related cancers, please see the publication listed below and online for the variable definitions:
  - Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive Cancer Incidence and Survival — United States, 2013. MMWR Morb Mortal Wkly Rep 2017;66:69–75. DOI: http://dx.doi.org/10.15585/mmwr.mm6603a1.
  - *Predefined SEER*Stat Variables for Calculating the Number of Associated Cancers for Selected Risk Factors* documentation available at https://www.cdc.gov/cancer/public-usee.

  Please note that in official federal cancer statistics publications, CDC reports malignant (invasive) cancers with the exception of bladder cancers (includes *in situ* and invasive cancers).

Values	Frequency	Percentage
Colon	1,297,827	28.5%
Postmenopausal female breast	2,667,487	58.6%
Corpus and uterus, NOS (not otherwise specified)	582,926	12.8%

[25] Eheman C, Henley SJ, Ballard-Barbash R, Jacobs EJ, Schymura MJ, Noone AM, Pan L, Anderson RN, Fulton JE, Kohler BA, Jemal A, Ward E, Plescia M, Ries LA, Edwards BK. Annual report to the nation on the status of cancer, 1975–2008, featuring cancers associated with excess weight and lack of sufficient physical activity. *Cancer* 2012;118:2338–2366. Available at www.ncbi.nlm.nih.gov/pmc/articles/PMC4586174/.
[26] World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington, DC: AICR, 2007. Available at http://preventcancer.aicr.org/site/PageServer?pagename=research_science_expert_report.
[27] Henley S.J., Singh S.D., King J, et al. Invasive cancer incidence and survival—United States, 2013. *MMWR* 2017.
[28] Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *Journal of the National Cancer Institute* 2011;103(24):1827–1839.
[29] World Cancer Research Fund / American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective.* Washington, DC: AICR, 2007.
[30] Levine B. What does the population attributable fraction mean? *Preventing Chronic Disease* 2007;4(1):A14.

SEER*Stat Item Name: **Tobacco-Related Cancers**
Source of Standard: NPCR
Source Item Name: Derived from NAACCR's *Primary Site*, *Histologic Type ICD-O-3*, *Sex*
Source Item Number: Derived from NAACCR's 400 (*Primary Site*), 522 (*Histologic Type ICD-O-3*), 220 (*Sex*)

Description

Predefined variable created using ICD-O-3 site, histology and sex to define tobacco-related cancers[31].

Considerations for use

- Cancer registries do not routinely collect data on tobacco use, so the number of cancers associated with this risk factor cannot be determined definitively[32,33,34].
- However, other sources of information can be used to obtain the proportion of cancers probably caused by the risk factor, also known as the *attributable fraction*[35]. Then the number of *attributable* cancers can be estimated by multiplying the attributable fraction by the number of *associated* cancers.
- For more information on risk-factor related cancers, please see the publication listed below and online for the variable definitions:
  - Henley SJ, Singh SD, King J, Wilson RJ, O'Neil ME, Ryerson AB. Invasive Cancer Incidence and Survival — United States, 2013. MMWR Morb Mortal Wkly Rep 2017;66:69–75. DOI: http://dx.doi.org/10.15585/mmwr.mm6603a1.
  - *Predefined SEER*Stat Variables for Calculating the Number of Associated Cancers for Selected Risk Factors* documentation available at https://www.cdc.gov/cancer/public-use.

  Please note that in official federal cancer statistics publications, CDC reports malignant (invasive) cancers with the exception of bladder cancers (includes *in situ* and invasive cancers).

Values	Frequency	Percentage
Lip, oral cavity, pharynx	490,723	6.1%
Esophagus	198,542	2.5%
Stomach	273,714	3.4%
Colon and rectum	1,813,989	22.5%
Liver	271,600	3.4%
Pancreas	492,688	6.1%
Larynx	162,457	2.0%
Trachea, lung, bronchus	2,528,891	31.4%
Cervix uteri	153,523	1.9%
Kidney and renal pelvis	672,799	8.3%
Urinary bladder	845,841	10.5%
Acute myeloid leukemia	157,379	2.0%

[31] U.S. Department of Health and Human Services. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014. Available at www.cdc.gov/tobacco/data_statistics/sgr/50th-anniversary/.
[32] Henley S.J., Singh S.D., King J, et al. Invasive cancer incidence and survival—United States, 2013. *MMWR* 2017.
[33] Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *Journal of the National Cancer Institute* 2011;103(24):1827–1839.
[34] World Cancer Research Fund / American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective.* Washington, DC: AICR, 2007.
[35] Levine B. What does the population attributable fraction mean? *Preventing Chronic Disease* 2007;4(1):A14.

Additional Resources

- NPCR www.cdc.gov/cancer/npcr
- SEER https://seer.cancer.gov
- U.S. Cancer Statistics Publication Standard www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm
- NAACCR www.naaccr.org/
- NAACCR data dictionary https://www.naaccr.org/data-standards-data-dictionary
- American College of Surgeons Commission on Cancer (CoC) Registry Manuals: *Facility Oncology Registry Data Standards* (FORDS) or *Registry Operations and Data Standards* (ROADS) www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals
- SEER site recode ICD-O-3/WHO 2008 https://seer.cancer.gov/siterecode/icdo3_dwhoheme/index.html
- ICCC site recode ICD-O-3/WHO 2008 https://seer.cancer.gov/iccc/iccc-who2008.html
- AYA site recode ICD-O-3/WHO 2008 https://seer.cancer.gov/ayarecode/
- Lymphoma subtype recode ICD-O-3/WHO 2008 https://seer.cancer.gov/lymphomarecode/
- ICD-O-3 http://apps.who.int/iris/bitstream/10665/96612/1/9789241548496_eng.pdf
- Collaborative Staging Manual http://cancerstaging.org/cstage/manuals.html
- Census www.census.gov

Abbreviations

AI/AN	American Indian or Alaska Native
A/PI	Asian or Pacific Islander
AYA	Adolescent and young adult
CCR	Central cancer registry
CNS	Central nervous system
CoC	Commission on Cancer
CS	Collaborative Stage
Dx	Diagnosis
ICCC	International Classification of Childhood Cancer
ICD-O-3	*International Classification of Diseases for Oncology*, Third Edition
NAACCR	North American Association of Central Cancer Registries
NAPIIA NAACCR	Asian/Pacific Islander identification algorithm
NHIA NAACCR	Hispanic identification algorithm
NOS	Not otherwise specified
NPCR	National Program of Cancer Registries
SEER	Surveillance, Epidemiology, and End Results
USCS	U.S. Cancer Statistics
WHO	World Health Organization