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

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

November, 2017 Submission

Diagnosis Years 2005–2015

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

NIH NATIONAL CANCER INSTITUTE

CS293113-C

Table of Contents

Message to Data Users

June 8, 2017

We are pleased to share the 2018-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 with the public. This database provides population-based cancer statistics on the *entire* United States population.

The NPCR and SEER Program are comprehensive and complex surveillance systems that combined, involve compiling and disseminating information on more than 1.7 million cancer cases annually. Cancer registry data provide a foundation of cancer surveillance activities that include identifying disparities in cancer burden, investigating potential causes of cancer, and evaluating and monitoring cancer prevention and screening activities.

This publicly available data source is the result of tremendous efforts by reporting facility staff, cancer registrars, central cancer registry staff, CDC NPCR and NCI SEER staff, and contractors. I thank everyone for their continued diligence in contributing to these important data, which 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 data source, researchers can have a positive impact on comprehensive cancer prevention and control and 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

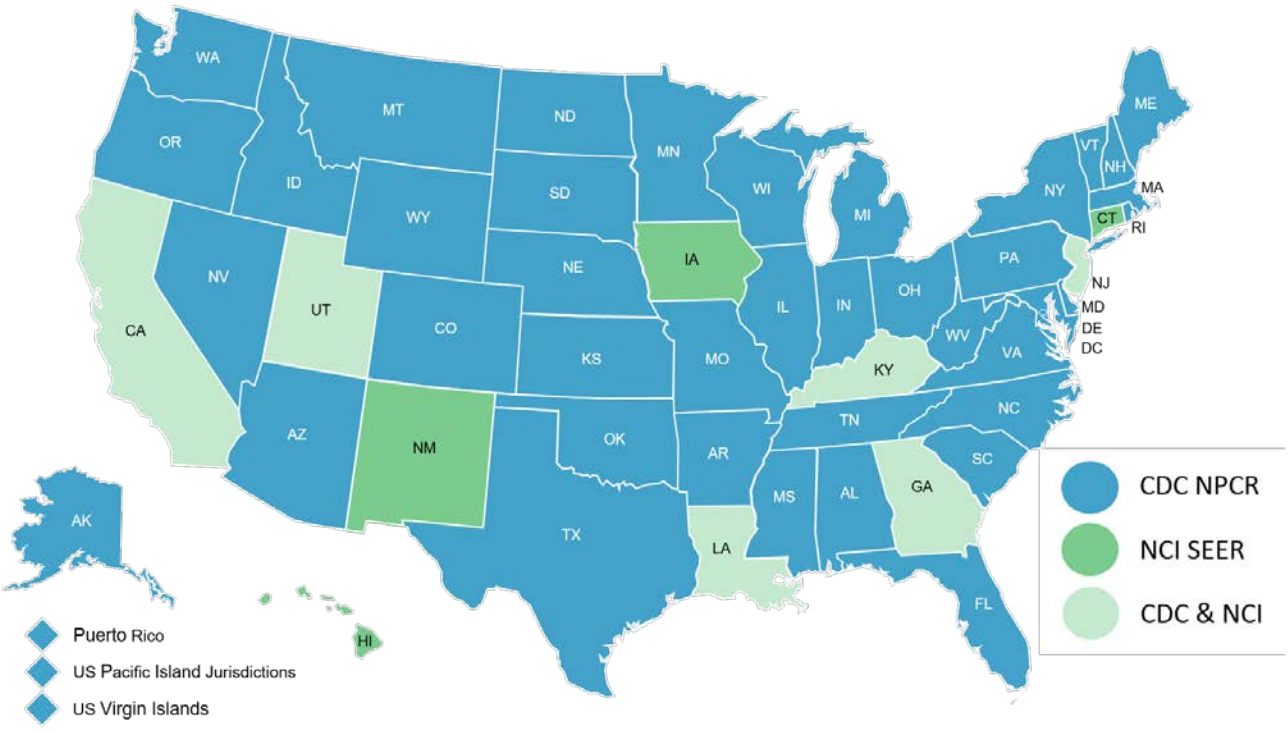
The National Program of Cancer Registries (NPCR), administered by the Centers for Disease Control and Prevention (CDC), was established by Congress in 1992. Through cooperative agreements, NPCR supports central cancer registries in 46 states, the District of Columbia, Puerto Rico, U.S. Pacific Island Jurisdictions, and the U.S. Virgin Islands (see map 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 20 U.S. geographic areas, including 5 states (see map 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. On an annual basis, the central cancer registries submit demographic and clinical information about each cancer case to CDC and/or NCI. **None of the information submitted to CDC or NCI contains personally identifiable information about individual patients.**

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 funded by NPCR and SEER in 2017



Data Collection, Submission, Quality, and Coverage

The most current data come from the November 2017 NPCR and SEER submissions, which cover cancer cases diagnosed from January 1, 2005 through December 31, 2015. 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. NPCR and SEER allow an interval of 23 months after the close of the diagnosis year for submission (for the 2015 data, NPCR required submission by November 30, 2017 and SEER required submission by November 1, 2017) to ensure case completeness and high quality.

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 2005–2015 Public Use Database

Two NPCR and SEER Incidence – U.S. Cancer Statistics public use databases are available for researchers: the 2001–2015 database and the 2005–2015 database. **This data standards document is specific to the 2005–2015 database**.

The 2001–2015 database includes race and ethnicity variables, while the 2005–2015 database does not. The 2005–2015 database includes Puerto Rico data, while the 2001–2015 database does not.

- The 2001–2015 database's population denominators are race-specific, ethnicity-specific, and sex-specific county population estimates from the U.S. Census (July 1, 2010–2016 bridged–race vintage 2016 population estimates), modified by SEER and aggregated to the state and national levels.

- The 2005–2015 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 2005–2015 public research data. In the 2018-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 2015.

[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 2005–2015 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%

[a] The NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2015 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 2005-2015 public use research database.

Table 3. Variables in the NPCR and SEER Incidence – U.S. Cancer Statistics 2005–2015 Public Use Research Database

SEER*Stat Category	SEER*Stat Variable Name
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
Extent of Disease – CS	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

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
SS	Summary stage
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 2005–2015 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 Research Database, Nov 2017 submission (2005-2015), United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. Released June 2018, based on November 2017 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 and SEER Summary Stage *in situ* because the information needed to distinguish between *in situ* and invasive bladder cancers is not always available or reliable.[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.

Suppression Rules[2-3]

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.

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.

Note: 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[9]

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

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 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 2005–2015 and 2001–2015 public use databases.

Reporting Delay

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 2005–2015 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 2005–2015 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 – 2015 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.

- *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 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.

- All cases with an unknown age or with sex other than male or female have been excluded from this database and are unavailable. 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.

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	13,079	0.1%
01–04 years	42,702	0.2%
05–09 years	32,192	0.2%
10–14 years	37,716	0.2%
15–19 years	64,209	0.3%
20–24 years	103,110	0.5%
25–29 years	164,169	0.9%
30–34 years	245,957	1.3%
35–39 years	371,829	2.0%
40–44 years	640,697	3.4%
45–49 years	1,050,854	5.5%
50–54 years	1,603,094	8.5%
55–59 years	2,088,011	11.0%
60–64 years	2,451,182	12.9%
65–69 years	2,646,776	14.0%
70–74 years	2,371,522	12.5%
75–79 years	2,066,407	10.9%
80–84 years	1,598,810	8.4%
85+ years	1,356,712	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	9,359,389	49.4%
Female	9,589,639	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 www.naaccr.org/StandardsandRegistryOperations/VolumeII.aspx
  - 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	1,583,482	8.4%
2006	1,623,678	8.6%
2007	1,677,981	8.9%
2008	1,702,118	9.0%
2009	1,728,293	9.1%
2010	1,719,751	9.1%
2011	1,759,187	9.3%
2012	1,755,601	9.3%
2013	1,782,216	9.4%
2014	1,800,858	9.5%
2015	1,815,863	9.6%

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 www.naaccr.org/StandardsandRegistryOperations/VolumeII.aspx
  - FORDS variable "state at diagnosis" at www.facs.org/quality-programs/cancer/ncdb/registrymanuals/cocmanuals

Values	Frequency	Percentage
Alaska	31,927	0.2%
Alabama	298,287	1.6%
Arkansas	179,346	0.9%
Arizona	336,056	1.8%
California	1,948,819	10.3%
Colorado	259,148	1.4%
Connecticut	254,116	1.3%
District of Columbia	34,199	0.2%
Delaware	65,853	0.3%
Florida	1,345,468	7.1%
Georgia	535,923	2.8%
Hawaii	81,342	0.4%
Idaho	88,035	0.5%
Illinois	787,511	4.2%
Indiana	389,290	2.1%
Iowa	206,811	1.1%
Kansas	174,418	0.9%
Kentucky	305,418	1.6%
Louisiana	278,869	1.5%
Massachusetts	440,260	2.3%
Maryland	344,452	1.8%
Maine	101,503	0.5%
Michigan	649,020	3.4%
Minnesota	325,673	1.7%
Missouri	374,238	2.0%
Mississippi	178,402	0.9%
Montana	66,449	0.4%
North Carolina	599,559	3.2%
North Dakota	42,346	0.2%
Nebraska	110,943	0.6%

Values	Frequency	Percentage
New Hampshire	94,513	0.5%
New Jersey	605,656	3.2%
New Mexico	105,156	0.6%
Nevada	134,869	0.7%
New York	1,321,322	7.0%
Ohio	735,395	3.9%
Oklahoma	220,921	1.2%
Oregon	241,040	1.3%
Pennsylvania	932,197	4.9%
Puerto Rico	162,242	0.9%
Rhode Island	73,972	0.4%
South Carolina	295,566	1.6%
South Dakota	51,108	0.3%
Tennessee	396,916	2.1%
Texas	1,190,358	6.3%
Utah	117,953	0.6%
Virginia	442,866	2.3%
Vermont	44,697	0.2%
Washington	419,470	2.2%
Wisconsin	363,373	1.9%
West Virginia	134,835	0.7%
Wyoming	30,922	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 2001–2015 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 2001–2015 diagnosis years (i.e., Connecticut, Hawaii, Iowa, and New Mexico).

Values	Frequency	Percentage
NPCR	18,301,603	96.6%
SEER	647,425	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)	18,653,755	98.4%
C80.9 (Unknown primary site)	295,273	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 3–5 of the NAACCR 2010 Implementation Guidelines. www.naaccr.org/StandardsandRegistryOperations/ImplementationGuidelines.aspx.
- 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

Values	Frequency	Percentage
8000–9992	18,949,028	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.”

Values	Frequency	Percentage
Well differentiated; Grade I	1,717,471	9.1%
Moderately differentiated; Grade II	4,504,748	23.8%
Poorly differentiated; Grade III	3,792,005	20.0%
Undifferentiated; anaplastic; Grade IV	598,185	3.2%
T-cell	67,876	0.4%
B-cell; pre-B; B-precursor	977,455	5.2%
Null cell; non T-non B	1,709	0.0%
NK cell; natural killer cell (1995+)	3,075	0.0%
Unknown	7,286,504	38.5%

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 that "diagnostic confirmation=microscopically confirmed" is selected.
- 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)	17,784,711	93.9%
Positive histology	17,034,640	89.9%
Positive exfoliative cytology, no positive histology	566,557	3.0%
Positive histology AND immunophenotyping AND/OR positive genetic studies	163,318	0.9%
Positive microscopic confirm, method not specified	20,196	0.1%
Positive laboratory test/marker study	80,236	0.4%
Direct visualization without microscopic confirmation	22,197	0.1%
Radiography without microscopic confirm	725,432	3.8%
Clinical diagnosis only	115,736	0.6%
Unknown	220,715	1.2%

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	574,552	2.3%
801–804	Epithelial Neoplasms, NOS	1,521,194	6.2%
805–808	Squamous Cell Neoplasms	1,885,283	7.6%
809–811	Basal Cell Neoplasms	7,025	0.0%
812–813	Transitional Cell Papillomas and Caricnomas	1,075,034	4.4%
814–838	Adneomas and Adenocarcinomas	9,723,989	39.4%
839–842	Adnexal and Skin Appendage Neoplasms	28,100	0.1%
843	Mucoepidermoid Neoplasms	21,432	0.1%
844–849	Cystic, Mucinous and Serous Neoplasms	670,283	2.7%
850–854	Ductal and Lobular Neoplasms	3,697,378	15.0%
855	Acinar Cell Neoplasms	53,688	0.2%
856–857	Complex Epithelial Neoplasms	83,524	0.3%
858	Thymic Epithelial Neoplasms	11,210	0.0%
859–867	specialized Gonadal Neoplasms	6,931	0.0%
868–871	Paragangliomas and Glomus Tumors	4,016	0.0%
872–879	Nevi and Melanomas	1,624,818	6.6%
880	Soft Tissue Tumors and Sarcomas, NOS	50,955	0.2%
881–883	Fibromatous Neoplasms	54,328	0.2%
884	Myxomatous Neoplasms	1,219	0.0%
885–888	Lipomatous Neoplasms	36,229	0.1%
889–892	Myomatous Neoplasms	57,811	0.2%
893–899	Complex Mixed and Stromal Neoplasms	119,215	0.5%
900–903	Fibroepithelial Neoplasms	7,299	0.0%
904	Synovial-Like Neoplasms	8,801	0.0%
905	Mesothelial Neoplasms	47,166	0.2%
906–909	Germ Cell Neoplasms	131,759	0.5%
910	Trophoblastic Neoplasms	5,704	0.0%
911	Mesonephromas	248	0.0%
912–916	Blood Vessel Tumors	54,419	0.2%
917	Lymphatic Vessel Tumors	210	0.0%
918–924	Osseous and Chondromatous Neoplasms	30,080	0.1%
925	Giant Cell Tumors	1,265	0.0%

ICD–O–3 Code	Label	Frequency	Percentage
926	Miscellaneous Bone Tumors	7,739	0.0%
927–934	Odotogenic Tumors	833	0.0%
935–937	Miscellaneous Tumors	14,900	0.1%
938–948	Gliomas	297,797	1.2%
949–952	Neuroepitheliomatous Neoplasms	25,013	0.1%
953	Meningiomas	319,393	1.3%
954–957	Nerve Sheath Tumors	82,534	0.3%
958	Granular Cell Tumors and Alveolar Soft Part Sarcomas	1,148	0.0%
959–972	Hodgkin and Non-Hodgkin Lymphomas	1,026,815	4.2%
973	Plasma Cell Tumors	299,431	1.2%
974	Mast Cell Tumors	2,175	0.0%
975	Neoplasms of Histiocytes and Accessory Lymphoid Cells	5,862	0.0%
976	Immunoproliferative Disease	17,113	0.1%
980–994	Leukemias	647,374	2.6%
995–996	Chronic Myeloproliferative Disorders	139,562	0.6%
997	Other Hematologic Disorders	12,378	0.1%
998–999	Myelodysplastic Syndromes	202,105	0.8%

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 is available at https://seer.cancer.gov/siterecode.

Values	Frequency	Percentage
All Sites (total)	**18,949,028**	100.0%
Oral Cavity and Pharynx	442,256	2.3%
Lip	24,995	0.1%
Tongue	132,148	0.7%
Salivary Gland	46,542	0.2%
Floor of Mouth	24,438	0.1%
Gum and Other Mouth	60,477	0.3%
Nasopharynx	19,945	0.1%
Tonsil	77,681	0.4%
Oropharynx	19,668	0.1%
Hypopharynx	26,029	0.1%
Other Oral Cavity and Pharynx	10,333	0.1%
Digestive System	3,128,249	16.5%
Esophagus	180,486	1.0%
Stomach	248,575	1.3%
Small Intestine	84,734	0.4%
Colon and Rectum	1,661,537	8.8%
Colon excluding Rectum	1,189,821	6.3%
Cecum	255,103	1.3%
Appendix	34,216	0.2%
Ascending Colon	233,666	1.2%
Hepatic Flexure	56,279	0.3%
Transverse Colon	109,702	0.6%
Splenic Flexure	35,662	0.2%
Descending Colon	70,901	0.4%
Sigmoid Colon	317,507	1.7%
Large Intestine, NOS	76,785	0.4%
Rectum and Rectosigmoid Junction	471,716	2.5%
Rectosigmoid Junction	119,095	0.6%
Rectum	352,621	1.9%
Anus, Anal Canal and Anorectum	76,794	0.4%
Liver and Intrahepatic Bile Duct	273,237	1.4%
Liver	242,875	1.3%

Values	Frequency	Percentage
Intrahepatic Bile Duct	30,362	0.2%
Gallbladder	42,868	0.2%
Other Biliary	64,312	0.3%
Pancreas	443,101	2.3%
Retroperitoneum	13,850	0.1%
Peritoneum, Omentum and Mesentery	21,617	0.1%
Other Digestive Organs	17,138	0.1%
Respiratory System	2,487,787	13.1%
Nose, Nasal Cavity and Middle Ear	25,799	0.1%
Larynx	148,809	0.8%
Lung and Bronchus	2,305,447	12.2%
Pleura	1,072	0.0%
Trachea, Mediastinum and Other Respiratory Organs	6,660	0.0%
Bones and Joints	32,889	0.2%
Soft Tissue including Heart	117,549	0.6%
Skin excluding Basal and Squamous	1,333,595	7.0%
Melanoma of the Skin	1,272,676	6.7%
Other Non-Epithelial Skin	60,919	0.3%
Breast (female and male combined)	3,067,381	16.2%
Female Genital System	1,011,244	5.3%
Cervix Uteri	140,061	0.7%
Corpus and Uterus, NOS	524,435	2.8%
Corpus Uteri	509,419	2.7%
Uterus, NOS	15,016	0.1%
Ovary	235,524	1.2%
Vagina	16,474	0.1%
Vulva	73,060	0.4%
Other Female Genital Organs	21,690	0.1%
Male Genital System	2,412,233	12.7%
Prostate	2,294,444	12.1%
Testis	91,911	0.5%
Penis	21,500	0.1%
Other Male Genital Organs	4,378	0.0%
Urinary System	1,420,193	7.5%
Urinary Bladder	769,332	4.1%
Kidney and Renal Pelvis	605,570	3.2%
Ureter	31,024	0.2%
Other Urinary Organs	14,267	0.1%
Eye and Orbit	35,133	0.2%
Brain and Other Nervous System	650,935	3.4%
Brain	262,739	1.4%
Cranial Nerves Other Nervous System	388,196	2.0%
Endocrine System	638,767	3.4%
Thyroid	473,989	2.5%
Other Endocrine including Thymus	164,778	0.9%
Lymphoma	801,190	4.2%
Hodgkin Lymphoma	95,042	0.5%
Hodgkin – Nodal	92,555	0.5%
Hodgkin – Extranodal	2,487	0.0%

Values	Frequency	Percentage
Non-Hodgkin Lymphoma	706,148	3.7%
NHL – Nodal	478,694	2.5%
NHL – Extranodal	227,454	1.2%
Myeloma	232,963	1.2%
Leukemia	483,833	2.6%
Lymphocytic Leukemia	240,810	1.3%
Acute Lymphocytic Leukemia	53,652	0.3%
Chronic Lymphocytic Leukemia	172,023	0.9%
Other Lymphocytic Leukemia	15,135	0.1%
Myeloid and Monocytic Leukemia	218,766	1.2%
Acute Myeloid Leukemia	141,555	0.7%
Acute Monocytic Leukemia	8,374	0.0%
Chronic Myeloid Leukemia	62,490	0.3%
Other Myeloid/Monocytic Leukemia	6,347	0.0%
Other Leukemia	24,257	0.1%
Other Acute Leukemia	8,626	0.0%
Aleukemic, Subleukemic and NOS	15,631	0.1%
Mesothelioma	35,059	0.2%
Kaposi Sarcoma	13,816	0.1%
Miscellaneous	603,956	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.
- Additional information is available at https://seer.cancer.gov/iccc/iccc3.html.

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

Values	Frequency	Percentage
I Leukemias, myeloproliferative & myelodysplastic diseases	43,561	22.9%
I(a) Lymphoid leukemias	31,085	16.4%
I(b) Acute myeloid leukemias	7,483	3.9%
I(c) Chronic myeloproliferative diseases	2,241	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,397	0.7%
I(e) Unspecified and other specified leukemias	1,355	0.7%
II Lymphomas and reticuloendothelial neoplasms	25,331	13.3%
II(a) Hodgkin lymphomas	11,497	6.1%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	8,392	4.4%
II(c) Burkitt lymphoma	2,236	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	2,883	1.5%
II(e) Unspecified lymphomas	323	0.2%
III CNS and misc intracranial and intraspinal neoplasms	44,561	23.5%
III(a) Ependymomas and choroid plexus tumor	3,600	1.9%
III(b) Astrocytomas	15,689	8.3%
III(c) Intracranial and intraspinal embryonal tumors	5,597	2.9%
III(d) Other gliomas	5,219	2.7%
III(e) Other specified intracranial/intraspinal neoplasms	12,682	6.7%
III(f) Unspecified intracranial and intraspinal neoplasms	1,774	0.9%
IV Neuroblastoma and other peripheral nervous cell tumors	8,014	4.2%
IV(a) Neuroblastoma and ganglioneuroblastoma	7,782	4.1%
IV(b) Other peripheral nervous cell tumors	232	0.1%
V Retinoblastoma	2,992	1.6%
VI Renal tumors	6,512	3.4%
VI(a) Nephroblastoma and other nonepithelial renal tumors	5,841	3.1%
VI(b) Renal carcinomas	650	0.3%
VI(c) Unspecified malignant renal tumors	21	0.0%
VII Hepatic tumors	2,216	1.2%
VII(a) Hepatoblastoma	1,637	0.9%
VII(b) Hepatic carcinomas	559	0.3%
VII(c) Unspecified malignant hepatic tumors	20	0.0%

Values	Frequency	Percentage
VIII Malignant bone tumors	8,266	4.4%
VIII(a) Osteosarcomas	4,642	2.4%
VIII(b) Chondrosarcomas	307	0.2%
VII(c) Ewing tumor and related sarcomas of bone	2,726	1.4%
VIII(d) Other specified malignant bone tumors	416	0.2%
VIII(e) Unspecified malignant bone tumors	175	0.1%
IX Soft tissue and other extraosseous sarcomas	11,242	5.9%
IX(a) Rhabdomyosarcomas	4,360	2.3%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	1,209	0.6%
IX(c) Kaposi sarcoma	50	0.0%
IX(d) Other specified soft tissue sarcomas	4,386	2.3%
IX(e) Unspecified soft tissue sarcomas	1,237	0.7%
X Germ cell & trophoblastic tumors & neoplasms of gonads	10,956	5.8%
X(a) Intracranial & intraspinal germ cell tumors	2,007	1.1%
X(b) Extracranial & extragonadal germ cell tumors	1,352	0.7%
X(c) Malignant gonadal germ cell tumors	6,873	3.6%
X(d) Gonadal carcinomas	420	0.2%
X(e) Other and unspecified malignant gonadal tumors	304	0.2%
XI Other malignant epithelial neoplasms and melanomas	17,930	9.4%
XI(a) Adrenocortical carcinomas	194	0.1%
XI(b) Thyroid carcinomas	8,144	4.3%
XI(c) Nasopharyngeal carcinomas	521	0.3%
XI(d) Malignant melanomas	4,361	2.3%
XI(e) Skin carcinomas	75	0.0%
XI(f) Other and unspecified carcinomas	4,635	2.4%
XII Other and unspecified malignant neoplasms	680	0.4%
XII(a) Other specified malignant tumors	370	0.2%
XII(b) Other unspecified malignant tumors	310	0.2%
Not classified by ICCC or *in situ*	7,637	4.0%

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, please visit https://seer.cancer.gov/iccc/iccc-who2008.html.
- Additional information is available at http://seer.cancer.gov/iccc/iccc3_ext.html.

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

Values	Frequency	Percentage
I Leukemias, myeloproliferative & myelodysplastic diseases	43,561	22.9%
I(a) Lymphoid leukemias	31,085	16.4%
I(a.1) Precursor cell leukemias	30,075	15.8%
I(a.2) Mature B-cell leukemias	793	0.4%
I(a.3) Mature T-cell and NK cell leukemias	106	0.1%
I(a.4) Lymphoid leukemia, NOS	111	0.1%
I(b) Acute myeloid leukemias	7,483	3.9%
I(c) Chronic myeloproliferative diseases	2,241	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,397	0.7%
I(e) Unspecified and other specified leukemias	1,355	0.7%
II Lymphomas and reticuloendothelial neoplasms	25,331	13.3%
II(a) Hodgkin lymphomas	11,497	6.1%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	8,392	4.4%
II(b.1) Precursor cell lymphomas	2,457	1.3%
II(b.2) Mature B-cell lymphomas except Burkitt lymphoma	3,391	1.8%
II(b.3) Mature T-cell and NK-cell lymphomas	2,117	1.1%
II(b.4) Non-Hodgkin lymphomas, NOS	427	0.2%
II(c) Burkitt lymphoma	2,236	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	2,883	1.5%
II(e) Unspecified lymphomas	323	0.2%
III CNS and misc intracranial and intraspinal neoplasms	44,561	23.5%
III(a) Ependymomas and choroid plexus tumor	3,600	1.9%
III(a.1) Ependymomas	2,671	1.4%
III(a.2) Choroid plexus tumor	929	0.5%
III(b) Astrocytomas	15,689	8.3%
III(c) Intracranial and intraspinal embryonal tumors	5,597	2.9%
III(c.1) Medulloblastomas	3,666	1.9%
III(c.2) PNET	1,070	0.6%
III(c.3) Medulloepithelioma	49	0.0%
III(c.4) Atypical teratoid/rhabdoid tumor	812	0.4%

Values	Frequency	Percentage
III(d) Other gliomas	5,219	2.7%
III(d.1) Oligodendrogliomas	577	0.3%
III(d.2) Mixed and unspecified gliomas	4,534	2.4%
III(d.3) Neuroepithelial glial tumors of uncertain orig	108	0.1%
III(e) Other specified intracranial/intraspinal neoplasms	12,682	6.7%
III(e.1) Pituitary adenomas and carcinomas	5,329	2.8%
III(e.2) Tumors of sellar region (craniopharyngiomas)	1,889	1.0%
III(e.3) Pineal parenchymal tumors	427	0.2%
III(e.4) Neuronal and mixed neuronal-glial tumors	3,631	1.9%
III(e.5) Meningiomas	1,406	0.7%
III(f) Unspecified intracranial and intraspinal neoplasms	1,774	0.9%
IV Neuroblastoma and other peripheral nervous cell tumors	8,014	4.2%
IV(a) Neuroblastoma and ganglioneuroblastoma	7,782	4.1%
IV(b) Other peripheral nervous cell tumors	232	0.1%
V Retinoblastoma	2,992	1.6%
VI Renal tumors	^1	^1
VI(a) Nephroblastoma and other nonepithelial renal tumors	^1	^1
VI(a.1) Nephroblastoma	5,492	2.9%
VI(a.2) Rhabdoid renal tumor	153	0.1%
VI(a.3) Kidney sarcomas	188	0.1%
VI(a.4) pPNET of kidney	^2	^2
VI(b) Renal carcinomas	650	0.3%
VI(c) Unspecified malignant renal tumors	21	0.0%
VII Hepatic tumors	2,216	1.2%
VII(a) Hepatoblastoma	1,637	0.9%
VII(b) Hepatic carcinomas	559	0.3%
VII(c) Unspecified malignant hepatic tumors	20	0.0%
VIII Malignant bone tumors	8,266	4.4%
VIII(a) Osteosarcomas	4,642	2.4%
VIII(b) Chondrosarcomas	307	0.2%
VIII(c) Ewing tumor and related sarcomas of bone	2,726	1.4%
VIII(c.1) Ewing tumor and Askin tumor of bone	2,617	1.4%
VIII(c.2) pPNET of bone	109	0.1%
VIII(d) Other specified malignant bone tumors	416	0.2%
VIII(d.1) Malignant fibrous neoplasms of bone	40	0.0%
VIII(d.2) Malignant chordomas	195	0.1%
VIII(d.3) Odontogenic malignant tumors	55	0.0%
VIII(d.4) Miscellaneous malignant bone tumors	126	0.1%
VIII(e) Unspecified malignant bone tumors	175	0.1%
IX Soft tissue and other extraosseous sarcomas	^1	^1
IX(a) Rhabdomyosarcomas	4,360	2.3%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	^1	^1
IX(b.1) Fibroblastic and myofibroblastic tumors	641	0.3%
IX(b.2) Nerve sheath tumors	553	0.3%
IX(b.3) Other fibromatous neoplasms	^2	^2
IX(c) Kaposi sarcoma	50	0.0%
IX(d) Other specified soft tissue sarcomas	4,386	2.3%
IX(d.1) Ewing tumor and Askin tumor of soft tissue	586	0.3%
IX(d.2) pPNET of soft tissue	254	0.1%
IX(d.3) Extrarenal rhabdoid tumor	251	0.1%
IX(d.4) Liposarcomas	260	0.1%

Values	Frequency	Percentage
IX(d.5) Fibrohistiocytic tumors	1,037	0.5%
IX(d.6) Leiomyosarcomas	167	0.1%
IX(d.7) Synovial sarcomas	984	0.5%
IX(d.8) Blood vessel tumors	182	0.1%
IX(d.9) Osseous & chondromatous neoplasms of soft tissue	93	0.0%
IX(d.10) Alveolar soft parts sarcoma	153	0.1%
IX(d.11) Miscellaneous soft tissue sarcomas	419	0.2%
IX(e) Unspecified soft tissue sarcomas	1,237	0.7%
X Germ cell & trophoblastic tumors & neoplasms of gonads	^1	^1
X(a) Intracranial & intraspinal germ cell tumors	2,007	1.1%
X(a.1) Intracranial & intraspinal germinomas	1,187	0.6%
X(a.2) Intracranial & intraspinal teratomas	566	0.3%
X(a.3) Intracranial & intraspinal embryonal carcinomas	25	0.0%
X(a.4) Intracranial & intraspinal yolk sac tumor	28	0.0%
X(a.5) Intracranial & intraspinal choriocarcinoma	21	0.0%
X(a.6) Intracranial & intraspinal tumors of mixed forms	180	0.1%
X(b) Extracranial & extragonadal germ cell tumors	^1	^1
X(b.1) Germinomas: extracranial/extragonadal	144	0.1%
X(b.2) Malignant teratomas: extracranial/extragonadal	521	0.3%
X(b.3) Embryonal carcinomas: extracranial/extragonadal	^2	^2
X(b.4) Yolk sac tumor: extracranial/extragonadal	308	0.2%
X(b.5) Choriocarcinomas: extracranial/extragonadal	151	0.1%
X(b.6) Other mixed germ cell: extracranial/extragonadal	215	0.1%
X(c) Malignant gonadal germ cell tumors	6,873	3.6%
X(c.1) Malignant gonadal germinomas	1,457	0.8%
X(c.2) Malignant gonadal teratomas	1,169	0.6%
X(c.3) Gonadal embryonal carcinomas	648	0.3%
X(c.4) Gonadal yolk sac tumor	639	0.3%
X(c.5) Gonadal choriocarcinoma	62	0.0%
X(c.6) Malignant gonadal tumors of mixed forms	2,898	1.5%
X(d) Gonadal carcinomas	420	0.2%
X(e) Other and unspecified malignant gonadal tumors	304	0.2%
XI Other malignant epithelial neoplasms and melanomas	17,930	9.4%
XI(a) Adrenocortical carcinomas	194	0.1%
XI(b) Thyroid carcinomas	8,144	4.3%
XI(c) Nasopharyngeal carcinomas	521	0.3%
XI(d) Malignant melanomas	4,361	2.3%
XI(e) Skin carcinomas	75	0.0%
XI(f) Other and unspecified carcinomas	4,635	2.4%
XI(f.1) Carcinomas of salivary glands	838	0.4%
XI(f.2) Carcinomas of colon and rectum	560	0.3%
XI(f.3) Carcinomas of appendix	966	0.5%
XI(f.4) Carcinomas of lung	424	0.2%
XI(f.5) Carcinomas of thymus	64	0.0%
XI(f.6) Carcinomas of breast	164	0.1%
XI(f.7) Carcinomas of cervix uteri	125	0.1%
XI(f.8) Carcinomas of bladder	271	0.1%
XI(f.9) Carcinomas of eye	26	0.0%
XI(f.10) Carcinomas of other specified sites	1,065	0.6%
XI(f.11) Carcinomas of unspecified site	132	0.1%

XII Other and unspecified malignant neoplasms	^[1]	^[1]
XII(a) Other specified malignant tumors	^[1]	^[1]
XII(a.1) Gastrointestinal stromal tumor	78	0.0%
XII(a.2) Pancreatoblastoma	31	0.0%
XII(a.3) Pulmonary blastoma and pleuropulmonary blastoma	189	0.1%
XII(a.4) Other complex mixed and stromal neoplasms	38	0.0%
XII(a.5) Mesothelioma	34	0.0%
XII(a.6) Other specified malignant tumors	^[2]	^[2]
XII(b) Other unspecified malignant tumors	310	0.2%
Not classified by ICCC or *in situ*	7,637	4.0%

[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, please visit https://seer.cancer.gov/ayarecode.

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

Values	Frequency	Percentage
1 Leukemias	20,240	6.1%
1.1 Acute lymphoid leukemia	7,833	2.4%
1.2 Acute myeloid leukemia	7,592	2.3%
1.3 Chronic myeloid leukemia	3,246	1.0%
1.4 Other and unspecified leukemia	1,569	0.5%
2 Lymphomas	46,377	14.0%
2.1 Non-Hodgkin lymphoma	18,207	5.5%
2.2 Hodgkin lymphoma	28,170	8.5%
3 CNS and Oth Intracranial and Intraspinal Neo (all behav)	31,181	9.4%
3.1. Astrocytoma	9,622	2.9%
3.1.1 Specified low-grade astrocytic tumors	3,801	1.1%
3.1.2 Glioblastoma and anaplastic astrocytoma	3,646	1.1%
3.1.3 Astrocytoma, NOS	2,175	0.7%
3.2 Other glioma	5,229	1.6%
3.3 Ependymoma	2,036	0.6%
3.4. Medulloblastoma and other PNET	1,558	0.5%
3.4.1 Medulloblastoma	925	0.3%
3.4.2 Supratentorial PNET	633	0.2%
3.5 Other specified intracranial and intraspinal neoplasms	10,859	3.3%
3.6 Unspecified intracranial and intraspinal neoplasms	1,877	0.6%
3.6.1 Unspec malignant intracranial and intraspinal neo	276	0.1%
3.6.2 Unspec ben/border intracran. and intraspin neo	1,601	0.5%
4 Osseous & Chondromatous Neoplasms	8,079	2.4%
4.1 Osteosarcoma	3,476	1.0%
4.2 Chondrosarcoma	1,095	0.3%
4.3 Ewing tumor	2,745	0.8%
4.4 Other specified and unspecified bone tumors	763	0.2%
5 Soft Tissue Sarcomas	13,549	4.1%
5.1 Fibromatous neoplasms	3,228	1.0%
5.2 Rhabdomyosarcoma	1,522	0.5%

Values	Frequency	Percentage
5.3 Other soft tissue sarcoma	8,799	2.7%
5.3.1 Specified soft tissue sarcoma	6,765	2.0%
5.3.1.1 Specified (excluding Kaposi sarcoma)	5,368	1.6%
5.3.1.2 Kaposi sarcoma	1,397	0.4%
5.3.2 Unspecified soft tissue sarcoma	2,034	0.6%
6 Germ Cell and Trophoblastic Neoplasms	38,297	11.6%
6.1 Germ cell and trophoblastic neoplasms of gonads	34,980	10.6%
6.2 Germ cell and trophoblastic neo of nongonadal sites	3,317	1.0%
6.2.1 Intracranial (all behaviors)	1,268	0.4%
6.2.2 Other nongonadal	2,049	0.6%
7 Melanoma and Skin Carcinomas	28,254	8.5%
7.1 Melanoma	27,997	8.4%
7.2 Skin carcinomas	257	0.1%
8 Carcinomas	98,187	29.6%
8.1 Thyroid carcinoma	43,503	13.1%
8.2 Other carcinoma of head and neck	5,324	1.6%
8.2.1 Nasopharyngeal carcinoma	930	0.3%
8.2.2 Other sites in lip, oral cavity and pharynx	3,890	1.2%
8.2.3 Nasal cav,mid ear,sinus,larynx,ill-def head/neck	504	0.2%
8.3 Carcinoma of trachea,bronchus, and lung	2,260	0.7%
8.4 Carcinoma of breast	12,073	3.6%
8.5 Carcinoma of genitourinary tract	18,938	5.7%
8.5.1 Carcinoma of kidney	4,207	1.3%
8.5.2 Carcinoma of bladder	1,572	0.5%
8.5.3 Carcinoma of gonads	2,925	0.9%
8.5.4 Carcinoma of cervix and uterus	9,718	2.9%
8.5.5 Carc of oth and ill-defined sites	516	0.2%
8.6 Carcinoma of gastrointestinal tract	14,397	4.3%
8.6.1 Carcinoma of colon and rectum	9,692	2.9%
8.6.2 Carcinoma of stomach	1,378	0.4%
8.6.3 Carcinoma of liver and intrahepatic bile ducts	1,302	0.4%
8.6.4 Carcinoma of pancreas	1,055	0.3%
8.6.5 Carc oth and ill-def sites, gastrointestinal tract	970	0.3%
8.7 Carcinoma of other and ill-defined sites	1,692	0.5%
8.7.1 Adrenocortical carcinoma	283	0.1%
8.7.2 Carcinoma of other and ill-defined sites, NOS	1,409	0.4%
9 Miscellaneous specified neoplasms, NOS	7,668	2.3%
9.1 Other pediatric and embryonal tumors, NOS	751	0.2%
9.1.1 Wilms tumor	160	0.0%
9.1.2 Neuroblastoma	209	0.1%
9.1.3 Other pediatric and embryonal tumors, NOS	382	0.1%
9.2 Other specified and embryonal tumors, NOS	6,917	2.1%
9.2.1 Paraganglioma and glomus tumors	271	0.1%
9.2.2 Other specified gonadal tumors	560	0.2%
9.2.3 Myeloma, mast cell, misc lymphoreticular neo, NOS	1,334	0.4%
9.2.4 Other specified neoplasms, NOS	4,752	1.4%
10 Unspecified Malignant Neoplasms	1,708	0.5%
Unclassified and Non-Malignant	37,948	11.4%

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	1,293,384	6.8%
1 Hodgkin Lymphoma	95,042	0.5%
1(a) Classical Hodgkin lymphoma	89,411	0.5%
1(a)1 Lymphocyte-rich/mixed cell/lymphocyte depleted	14,953	0.1%
1(a)1.1 Lymphocyte-rich	3,424	0.0%
1(a)1.2 Mixed cellularity	10,332	0.1%
1(a)1.3 Lymphocyte-depleted	1,197	0.0%
1(a)2 Nodular sclerosis	49,619	0.3%
1(a)3 Classical Hodgkin lymphoma, NOS	24,839	0.1%
1(b) Nodular lymphocyte predominant Hodgkin lymphoma	5,631	0.0%
2 Non-Hodgkin lymphoma	1,164,114	6.1%
2(a) Non-Hodgkin lymphoma, B-cell	1,077,683	5.7%
2(a)1 Precursor Non-Hodgkin lymphoma, B-cell	44,313	0.2%
2(a)2 Mature Non-Hodgkin lymphoma, B-cell	976,939	5.2%
2(a)2.1 Chronic/Sm/Prolymphocytic/Mantle B-cell NHL	241,867	1.3%
2(a)2.1.1 Chronic/Small lymphocytic leuk/lymph	210,348	1.1%
2(a)2.1.2 Prolymphocytic leukemia, B-cell	980	0.0%
2(a)2.1.3 Mantle-cell lymphoma	30,539	0.2%
2(a)2.2 Lymphoplasmacytic lymphoma/Waldenstrom	21,975	0.1%
2(a)2.2.1 Lymphoplasmacytic lymphoma	9,671	0.1%
2(a)2.2.2 Waldenstrom macroglubulinemia	12,304	0.1%
2(a)2.3 Diffuse large B-cell lymphoma (DLBCL)	251,126	1.3%
2(a)2.3.1 DLBCL, NOS	248,296	1.3%
2(a)2.3.2 Intravascular large B-cell lymphoma	478	0.0%
2(a)2.3.3 Primary effusion lymphoma	374	0.0%
2(a)2.3.4 Mediastinal large B-cell lymphoma	1,978	0.0%
2(a)2.4 Burkitt lymphoma/leukemia	14,016	0.1%
2(a)2.5 Marginal-zone lymphoma (MZL)	69,247	0.4%

Values	Frequency	Percentage
2(a)2.5.1 Splenic MZL	6,538	0.0%
2(a)2.5.2 Extranodal MZL, MALT type	41,061	0.2%
2(a)2.5.3 Nodal MZL	21,648	0.1%
2(a)2.6 Follicular lymphoma	134,489	0.7%
2(a)2.7 Hairy-cell leukemia	10,319	0.1%
2(a)2.8 Plasma cell neoplasms	233,711	1.2%
2(a)2.8.1 Plasmacytoma	15,347	0.1%
2(a)2.8.2 Multiple myeloma/plasma-cell leuk	218,364	1.2%
2(a)2.9 Heavy chain disease	189	0.0%
2(a)3 Non-Hodgkin lymphoma, B-cell, NOS	56,431	0.3%
2(b) Non-Hodgkin lymphoma, T-cell	75,865	0.4%
2(b)1 Precursor Non-Hodgkin lymphoma, T-cell	2,673	0.0%
2(b)2 Mature Non-Hodgkin lymphoma, T-cell	72,921	0.4%
2(b)2.1 Mycosis fungoides/Sezary syndrome	15,974	0.1%
2(b)2.1.1 Mycosis fungoides	15,364	0.1%
2(b)2.1.2 Sezary syndrome	610	0.0%
2(b)2.2 Peripheral T-cell lymphoma	40,431	0.2%
2(b)2.2.1 Peripheral T-cell lymphoma, NOS	14,550	0.1%
2(b)2.2.2 Angioimmunoblastic T-cell lymphoma	4,886	0.0%
2(b)2.2.3 Subcutan panniculitis-like T-cell lymph	385	0.0%
2(b)2.2.4 Anaplastic lar cell lymph, T-/Null-cell	8,434	0.0%
2(b)2.2.5 Hepatosplenic T-cell lymphoma	344	0.0%
2(b)2.2.6 Enteropathy-type T-cell lymphoma	504	0.0%
2(b)2.2.7 Cutaneous T-cell lymphoma, NOS	8,163	0.0%
2(b)2.2.8 Prim cutaneous anaplastic lar cell lymph	3,165	0.0%
2(b)2.3 Adult T-cell leukemia/lymphoma	8,476	0.0%
2(b)2.4 NK/T-cell lymph, nasal-type/aggres NK leuk	2,248	0.0%
2(b)2.5 T-cell large granular lymphocytic leukemia	4,370	0.0%
2(b)2.6 Prolymphocytic leukemia, T-cell	1,422	0.0%
2(b)3 Non-Hodgkin lymphoma, NOS, T-cell	271	0.0%
2(c) Non-Hodgkin lymphoma, unknown lineage	10,566	0.1%
2(c)1 Precursor lymphoblastic leuk/lymph, unk lineage	3,896	0.0%
2(c)2 Prolymphocytic leukemia, unknown lineage	396	0.0%
2(c)3 Non-Hodgkin lymphoma, NOS, unknown lineage	6,274	0.0%
3 Composite Hodgkin lymphoma and NHL	2,471	0.0%
4 Lymphoid neoplasm, NOS	31,757	0.2%
Unclassified	17,655,644	93.2%

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 (including "Malignant", "Only malignant in ICD-O-3", and "Only malignant 2010+"categories) 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	505,634	2.7%
Borderline malignancy	49,637	0.3%
In situ	1,332,700	7.0%
Malignant	16,775,580	88.5%
Only malignant in ICD-O-3	265,343	1.4%
Only malignant 2010+	20,134	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 other variables: *SEER Summary Stage 2000*, which records stage from diagnosis years 2001–2003, and *Derived SS2000*, which records stage from diagnostic years 2004–2015. This stage variable can be used for diagnosis years 2001–2015.

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 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, please review https://cancerstaging.org/cstage/Pages/default.aspx.

Values	Frequency	Percentage
In situ	1,711,126	9.0%
Localized only	7,681,284	40.5%
Regional, direct extension only	1,251,642	6.6%
Regional, regional lymph nodes only	1,339,015	7.1%
Regional, direct extension and regional lymph nodes	775,406	4.1%
Regional, NOS	175,603	0.9%
Distant site(s)/node(s) involved	4,190,007	22.1%
Not applicable	555,094	2.9%
Unknown/unstaged/unspecified	1,269,224	6.7%
Blanks(s)	627	0.0%

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	10,300,863	54.4%
Right - origin of primary	4,197,101	22.1%
Left - origin of primary	3,893,800	20.5%
Only one side - side unspecified	34,265	0.2%
Bilateral, single primary	140,193	0.7%
Paired site: midline tumor	51,431	0.3%
Paired site, but no information concerning laterality	331,375	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	13,660,427	^[1]
1st of 2 or more primaries	1,372,336	^[1]
2nd of 2 or more primaries	2,761,979	^[1]
3rd of 3 or more primaries	483,603	^[1]
4th of 4 or more primaries	87,574	^[1]
5th of 5 or more primaries	18,481	^[1]
6th or more primaries[2]	8,961	^[1]
Only one state registry-defined neoplasm	531,392	^[1]
1st of 2 or more state registry-defined neoplasms	10,376	^[1]
2nd of 2 or more state registry-defined neoplasms	12,148	^[1]
3rd of 3 or more state registry-defined neoplasms	846	^[1]
4th of 4 or more state registry-defined neoplasms	214	^[1]
5th of 5 or more state registry-defined neoplasms	94	^[1]
6th or more state registry-defined neoplasms[1]	43	^[1]

Values	Frequency	Percentage
Unknown sequence number - federally required *in situ* or malignant tumors	111	^[1]
Carcinoma *in situ* of the Cervix diagnosed 1/1/1996 or later	^[3]	^[3]
Unknown sequence number - state registry-defined neoplasms	257	^[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	30	^[2]
1900	115	^[2]
1901	131	^[2]
1902	158	^[2]
1903	293	^[2]
1904	440	^[2]
1905	812	^[2]
1906	1,276	^[2]
1907	2,106	^[2]
1908	3,347	^[2]
1909	5,118	^[2]
1910	7,801	^[2]
1911	11,196	^[2]
1912	16,726	^[2]
1913	22,370	^[2]
1914	31,083	^[2]
1915	39,773	^[2]
1916	51,675	^[2]
1917	66,396	^[2]
1918	84,613	^[2]

Values	Frequency	Percentage
1919	100,936	^2
1920	132,260	^2
1921	161,506	^2
1922	182,240	^2
1923	209,169	^2
1924	238,879	^2
1925	260,322	^2
1926	285,215	^2
1927	314,404	^2
1928	331,113	^2
1929	345,339	^2
1930	371,616	^2
1931	375,576	^2
1932	388,802	^2
1933	384,212	^2
1934	410,683	^2
1935	424,993	^2
1936	432,896	^2
1937	446,847	^2
1938	463,809	^2
1939	461,657	^2
1940	473,010	^2
1941	488,045	^2
1942	527,747	^2
1943	529,276	^2
1944	489,003	^2
1945	464,315	^2
1946	528,601	^2
1947	561,650	^2
1948	508,300	^2
1949	479,442	^2
1950	448,914	^2
1951	440,125	^2
1952	427,845	^2
1953	406,753	^2
1954	395,973	^2
1955	374,558	^2
1956	360,273	^2
1957	343,960	^2
1958	316,630	^2
1959	296,712	^2
1960	275,784	^2
1961	254,692	^2

Values	Frequency	Percentage
1962	231,489	^2
1963	211,263	^2
1964	190,916	^2
1965	164,594	^2
1966	145,872	^2
1967	130,606	^2
1968	120,157	^2
1969	112,580	^2
1970	106,580	^2
1971	93,805	^2
1972	80,823	^2
1973	70,415	^2
1974	65,641	^2
1975	59,775	^2
1976	54,992	^2
1977	51,980	^2
1978	48,140	^2
1979	46,135	^2
1980	42,780	^2
1981	39,459	^2
1982	36,473	^2
1983	32,627	^2
1984	29,961	^2
1985	27,243	^2
1986	24,532	^2
1987	22,475	^2
1988	20,987	^2
1989	19,032	^2
1990	17,587	^2
1991	15,802	^2
1992	13,968	^2
1993	12,415	^2
1994	11,343	^2
1995	9,904	^2
1996	9,080	^2
1997	8,343	^2
1998	8,182	^2
1999	7,544	^2
2000	7,294	^2
2001	7,320	^2
2002	7,453	^2
2003	7,766	^2
2004	8,209	^2

Values	Frequency	Percentage
2005	8,181	^[2]
2006	7,922	^[2]
2007	7,423	^[2]
2008	7,009	^[2]
2009	6,137	^[2]
2010	5,504	^[2]
2011	4,559	^[2]
2012	3,749	^[2]
2013	2,905	^[2]
2014	1,779	^[2]
2015	730	
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	1,648,709	8.7%
February	1,473,047	7.8%
March	1,616,669	8.5%
April	1,589,634	8.4%
May	1,595,085	8.4%
June	1,635,614	8.6%
July	1,556,287	8.2%
August	1,601,007	8.4%
September	1,529,412	8.1%
October	1,620,196	8.6%
November	1,483,036	7.8%
December	1,470,829	7.8%
Blank(s)	129,503	0.7%

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	16,281,070	85.9%
Radiation treatment or medical oncology center (2006+)	465,223	2.5%
Laboratory only (hospital or private)	526,035	2.8%
Physician's office/private medical practitioner (LMD)	862,061	4.5%
Nursing/convalescent home/hospice	22,709	0.1%
Other hospital outpatient unit or surgery center (2006+)	791,684	4.2%
Blank(s)	246	0.0%

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 following publication:
  - 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.

  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	435,824	7.9%
Esophagus	146,108	2.6%
Colon and rectum	1,624,842	29.4%
Liver	158,872	2.9%
Larynx	146,412	2.6%
Female breast cancer	3,019,754	54.6%

[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 Papillomaviraus (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:
  - 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.

  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	179,582	37.8%
Anal and rectal squamous cell carcinoma	73,766	15.5%
Vulvar squamous cell carcinoma	56,687	11.9%
Vaginal squamous cell carcinoma	10,859	2.3%
Penile squamous cell carcinoma	20,125	4.2%
Cervical carcinoma	134,393	28.3%

[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.

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 following publication:
  - 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.

  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	112,174	1.6%
Gastric cardia	76,533	1.1%
Colon & rectum	1,624,842	23.8%
Liver	158,872	2.3%
Gallbladder	40,254	0.6%
Pancreas	376,099	5.5%
Kidney	510,445	7.5%
Meningioma	128,802	1.9%
Thyroid	471,872	6.9%
Multiple myeloma	195,955	2.9%
Post-menopausal female breast	2,393,791	35.0%

[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.

Corpus and uterus, NOS (not otherwise specified)	519,828	7.6%
Ovary	221,492	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 following publication:
  - 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.

  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	1,159,821	28.5%
Postmenopausal female breast	2,393,791	58.8%
Corpus and uterus, NOS (not otherwise specified)	519,828	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 following publication:
  - 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.

  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	435,824	6.4%
Esophagus	175,412	2.6%
Stomach	242,991	3.5%
Colon and rectum	1,624,842	23.7%
Liver	158,872	2.3%
Pancreas	376,099	5.5%
Larynx	146,412	2.1%
Trachea, lung, bronchus	2,099,799	30.7%
Cervix uteri	137,568	2.0%
Kidney and renal pelvis	552,928	8.1%
Urinary bladder	759,121	11.1%
Acute myeloid leukemia	135,191	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 www.naaccr.org/StandardsandRegistryOperations/VolumeII.aspx
- 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
SS	Summary Stage
USCS	U.S. Cancer Statistics
WHO	World Health Organization