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 2001–2015

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

NIH NATIONAL CANCER INSTITUTE

CS293113-D

Table of Contents

Message to Data Users 3

Overview of CDC's National Program of Cancer Registries and NCI's Surveillance, Epidemiology, and End Results Program 4

Variable List 7

Data Citations 8

Cautionary Notes 9

Checklist for a NPCR and SEER Incidence – U.S. Cancer Statistics 2001– 2015 Public Use Data Analysis 12

NPCR and SEER Incidence – U.S. Cancer Statistics 2001– 2015 Public Use Database: Variable Definition and Frequency 13

Additional Resources 65

Abbreviations 67

Appendix A. NPCR – Indian Health Services (IHS) Linkage Schedule 68

Message to Data Users

June 8, 2018

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. This database provides population-based data on newly diagnosed cancer cases across the *entire* United States population.

The NPCR and SEER Program are comprehensive surveillance systems that work collaboratively to collect, compile, and disseminate 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 efforts by reporting facilities, cancer registrars, central cancer registries, and 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 as well as the care and quality of lives for those diagnosed with cancer.

Sincerely,

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

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

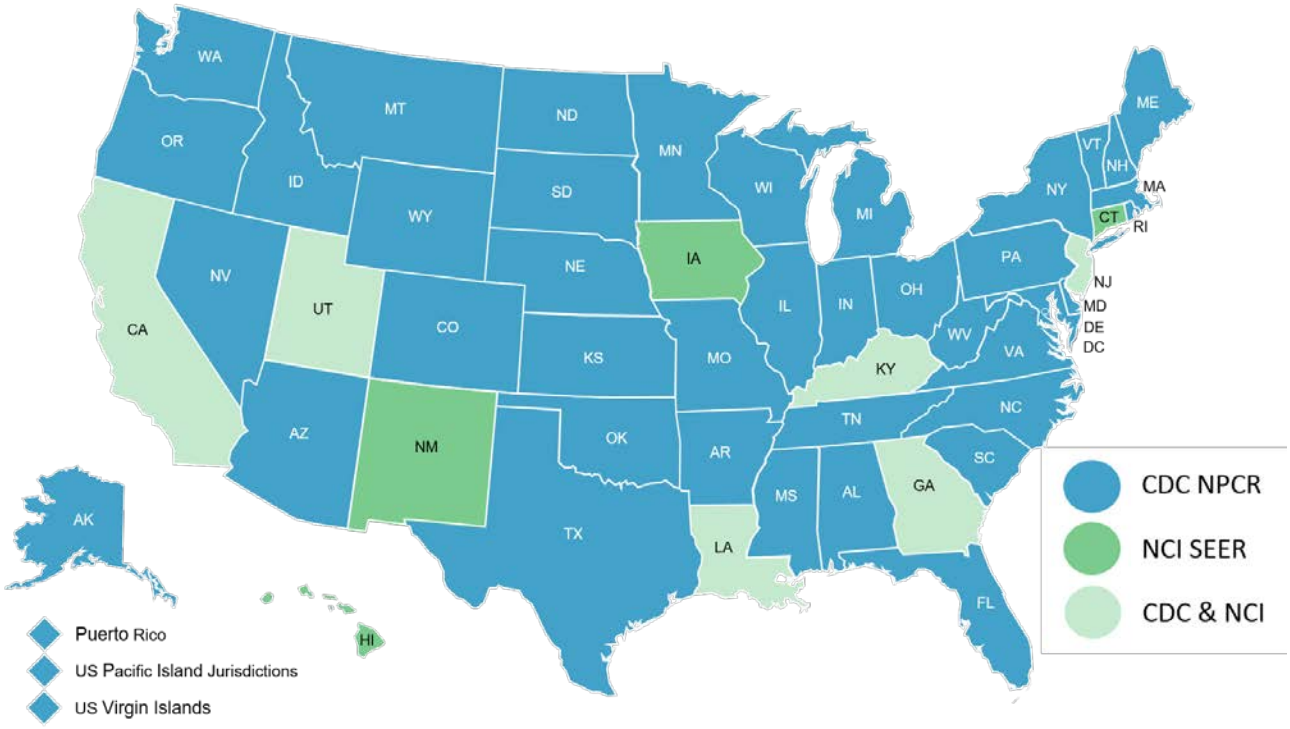
The National Program of Cancer Registries (NPCR), administered by the Centers for Disease Control and Prevention (CDC), was established by Congress in 1992. Through cooperative agreements, NPCR supports central cancer registries in 46 states, the District of Columbia, Puerto Rico, U.S. Pacific Island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

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 diagnosed with cancer to a central cancer registry, where the information is consolidated and goes through rigorous checks for quality and completeness. All hospitals are required by state law to report cancer cases to the central cancer registry in their respective states. On an annual basis, the central cancer registries submit demographic and clinical information about each cancer case to CDC and/or NCI.

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

Figure 1. Central cancer registry programs funded by NPCR and SEER in 2018



Data Collection, Submission, Quality, and Coverage

The most current data come from the November 2017 NPCR and SEER submissions, which include cancer cases diagnosed from January 1, 2001 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 to ensure case completeness and high quality. NPCR and SEER allow an interval of 23 months after the close of the diagnosis year for submission (for the 2015 data, NPCR required submission by November 30, 2017 and SEER required submission by November 1, 2017).

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 2001–2015 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 2001–2015 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 2001–2015 public research data. In the 2018-release of the public use database there is 100% population coverage for all 50 states and the District of Columbia for cases diagnosed from 2001 through 2015. Mississippi's cases diagnosed in 2001 and 2002 are not available and the U.S. population coverage those 2 years is 99%.

[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 2001–2015 Public Use Research Database.

Diagnosis year(s)	Percentage of U.S. population covered in database
2001 [b]	99%
2002 [b]	99%
2003	100%
2004	100%
2005	100%
2006	100%
2007	100%
2008	100%
2009	100%
2010	100%
2011	100%
2012	100%
2013	100%
2014	100%
2015	100%
2001–2015	99%
2006–2015 [c]	100%
2011–2015 [d]	100%

[a] The NPCR and SEER Incidence – U.S. Cancer Statistics 2001–2015 Public Use Research Database includes data submitted by all 50 states and District of Columbia. Puerto Rico, U.S. Pacific Island Jurisdiction, and U.S. Virgin Island data are not included in the database.

[b] Mississippi's cases diagnosed in 2001 and 2002 are not available in the 2001-2015 public use database.

[c] The most recently submitted 10 years of data.

[d] The most recently submitted 5 years of data.

Variable List

Table 3 shows all of the variables available in the 2001-2015 public use research database.

Table 3. Variables in the NPCR and SEER Incidence – U.S. Cancer Statistics 2001–2015 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
	USCS standard
	Race recode for USCS
	Program
	Region
	Stateraceethincl
	USCS0115
	USCS0615
	USCS1115
	Origin recode NHIA (Hispanic, Non-Hisp)
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
Race and Age (case data only)	NHIA derived Hisp origin
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 2001–2015 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 (2001-2015), 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 this 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]

Complementary Cell Suppression

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

Suppressing less than 16 cases

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

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

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

Race and Ethnicity Suppression

States have the option to suppress race-specific and Hispanic ethnicity-specific data every submission year. While these states can be included in an aggregated analysis, the affected state's race and ethnicity information cannot be reported at the state level.

The Merged System-Supplied variable, *state race eth suppress*, can be used to restrict your analysis to the states that are eligible to be included in a state-level analysis of race and ethnicity.

The following states have state-level race or ethnicity data presentation restrictions:

- Data for American Indians and Alaska Natives cannot be displayed for Delaware, Illinois, Kansas, Kentucky, New Jersey, and New York.
- Data for Asians and Pacific Islanders cannot be displayed for Delaware, Kansas and Kentucky.
- Hispanic ethnicity data cannot be displayed for Delaware, Kentucky, and Massachusetts.
- Race and ethnicity combinations—e.g., white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for Delaware, Kansas, Kentucky, Massachusetts, and Pennsylvania.

For more information, please refer to the *Race Recode for USCS*, *Origin recode NHIA (Hispanic, Non-Hisp)*, and *NHIA derived Hisp origin* variable descriptions in this document.

Case Level Data

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

Benign Central Nervous System (CNS) Tumors

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

Primary Site Variables[4]

Beginning in diagnosis year 2010, some of the lymphoma and leukemia ICD-O-3 codes were updated based on changes from the World Health Organization. The appropriate site recode variables to 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 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.

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 2001–2015 Public Use Data Analysis

Multi-year analyses

☐ The database includes variables that can be used to restrict analyses to the states meeting U.S. Cancer Statistics publication criteria during the most commonly analyzed multi-year time periods, specifically: all years of data in the database (variable *USCS0115* for diagnosis years 2001–2015), the most recent 10 years of data (*USCS0615* for diagnosis years 2006–2015) and the most recent 5 years of data (*USCS1115* for diagnosis years 2011–2015).

If you are conducting a multi-year analysis and want to restrict it to the states that met reporting standards during each of the years, did you use variable *USCS0115*, *USCS0615*, or *USCS1115* and also use the *Year of Diagnosis* variable to restrict to the corresponding year range on the SEER*Stat Selection tab?

- This is important to do during a trend analysis, as the same states need to be included for each year being analyzed for comparable comparisons.
- The *Year of Diagnosis* variable is used in combination with the predefined USCS variable to exclude the non-relevant years. For example, if *USCS1115* is used, then *Year of Diagnosis* should also be restricted to diagnosis years 2011–2015 in the SEER*Stat Selection tab.
- If you would like to analyze a range of years other than those predefined variables, please contact CDC at uscsdata@cdc.gov and we will create a new variable for you.[1]

Single year analyses

☐ If you are analyzing just 1 year of data, did you use the variable *USCS Standard* and restricted the analysis to the specific *Year of Diagnosis* in the SEER*Stat Selection tab? [2]

Common selection and reporting considerations

☐ If you are reporting **state-specific race, ethnicity or race/ethnicity combinations**, have you suppressed data from the registries that opted out of reporting these data items using the *state race eth suppress* variable or manually in the SEER*Stat Selection tab?[3]

☐ 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)?[4]

☐ 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?[4]

☐ 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?[6]

☐ 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 2001–2015 Public Use Research Database?[7]

---

[1] See *USCS0115*, *USCS0615*, and *USCS1115* variable descriptions.

[2] See *USCS Standard* variable description.

[3] See *state race eth suppress*, *Race Recode*, *Origin recode NHIA*, and *NHIA derived Hisp origin* variable descriptions.

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

[5] See *Diagnostic Confirmation* variable descriptions.

[6] See *Sex* variable description.

[7] See *Data Citation* section.

NPCR and SEER Incidence – U.S. Cancer Statistics 2001– 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	17,002	0.1%
01–04 years	55,826	0.2%
05–09 years	41,432	0.2%
10–14 years	48,815	0.2%
15–19 years	82,410	0.3%
20–24 years	132,007	0.5%
25–29 years	207,766	0.8%
30–34 years	320,825	1.3%
35–39 years	494,168	2.0%
40–44 years	862,187	3.5%
45–49 years	1,384,920	5.6%
50–54 years	2,071,664	8.4%
55–59 years	2,668,342	10.8%
60–64 years	3,097,366	12.5%
65–69 years	3,372,859	13.7%
70–74 years	3,150,120	12.8%
75–79 years	2,810,069	11.4%
80–84 years	2,130,793	8.6%
≥85 years	1,748,766	7.1%

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	12,252,411	49.6%
Female	12,444,926	50.4%

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
2001	1,440,654	5.8%
2002	1,460,628	5.9%
2003	1,473,738	6.0%
2004	1,535,531	6.2%
2005	1,570,851	6.4%
2006	1,610,646	6.5%
2007	1,664,840	6.7%
2008	1,688,587	6.8%
2009	1,714,256	6.9%
2010	1,704,950	6.9%
2011	1,743,581	7.1%
2012	1,739,646	7.0%
2013	1,765,795	7.1%
2014	1,784,627	7.2%
2015	1,799,007	7.3%

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	41,745	0.2%
Alabama	388,186	1.6%
Arkansas	236,418	1.0%
Arizona	436,764	1.8%
California	2,563,307	10.4%
Colorado	335,915	1.4%
Connecticut	339,314	1.4%
District of Columbia	46,579	0.2%
Delaware	84,499	0.3%
Florida	1,779,441	7.2%
Georgia	687,314	2.8%
Hawaii	106,333	0.4%
Idaho	113,181	0.5%
Illinois	1,044,232	4.2%
Indiana	516,188	2.1%
Iowa	275,086	1.1%
Kansas	229,774	0.9%
Kentucky	398,870	1.6%
Louisiana	369,651	1.5%
Massachusetts	593,077	2.4%
Maryland	455,799	1.8%
Maine	135,526	0.5%
Michigan	873,357	3.5%
Minnesota	426,935	1.7%
Missouri	492,806	2.0%
Mississippi	205,902	0.8%
Montana	86,956	0.4%
North Carolina	764,427	3.1%
North Dakota	56,053	0.2%
Nebraska	147,520	0.6%

Values	Frequency	Percentage
New Hampshire	124,256	0.5%
New Jersey	809,792	3.3%
New Mexico	138,033	0.6%
Nevada	177,831	0.7%
New York	1,742,587	7.1%
Ohio	977,919	4.0%
Oklahoma	292,745	1.2%
Oregon	318,654	1.3%
Pennsylvania	1,242,407	5.0%
Rhode Island	100,192	0.4%
South Carolina	383,107	1.6%
South Dakota	67,700	0.3%
Tennessee	502,113	2.0%
Texas	1,556,814	6.3%
Utah	150,006	0.6%
Virginia	576,243	2.3%
Vermont	58,878	0.2%
Washington	548,394	2.2%
Wisconsin	478,272	1.9%
West Virginia	179,569	0.7%
Wyoming	40,670	0.2%

SEER*Stat Item Name: **USCS standard**

Source of Standard: NPCR

Source Item Name: Derived from NAACCR's *State at diagnosis (CoC)* and U.S. Cancer Statistics publication standard

Source Item Number: Derived from NAACCR's 80

Description

This variable indicates the central cancer registries with cancer incidence data that are of high quality and meet the U.S. Cancer Statistics standard for a single year of analysis at the national level for all cancer sites combined.

Considerations for use

- This variable allows the selection of only those central cancer registries whose data meet the U.S. Cancer Statistics standard for an individual diagnosis year. The year of diagnosis should also be specified in the SEER*Stat Selection tab.
- If you are conducting a multiyear analysis and want to restrict the analysis to the states that met publication criteria for each of those years (for example, a trend analysis), we recommend using the predefined variables, *USCS1115* (includes diagnosis years 2011–2015), *USCS0615* (includes diagnosis years 2006–2015), or *USCS0115* (includes diagnosis years 2001–2015).
- If you would like to analyze a range of years other than those predefined variables, please contact CDC at uscsdata@cdc.gov and we will create a new variable for you that can be imported into SEER*Stat.
- The U.S. Cancer Statistics publication standard is available at www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm.

Number of central cancer registries[a]	Year of diagnosis	Frequency	Percentage
50	2001	1,440,654	6.3%
50	2002	1,460,628	6.4%
51	2003	1,473,738	6.4%
51	2004	1,535,531	6.7%
51	2005	1,570,851	6.9%
51	2006	1,610,646	7.0%
51	2007	1,664,840	7.3%
51	2008	1,688,587	7.4%
51	2009	1,714,256	7.5%
51	2010	1,704,950	7.4%
51	2011	1,743,581	7.6%
51	2012	1,739,646	7.6%
51	2013	1,765,795	7.7%
51	2014	1,784,627	7.8%
51	2015	1,440,654	6.3%

[a] Refer to Table 1 for the central cancer registries available by each diagnosis year.

SEER*Stat Item Name: **Race Recode for USCS**

Source of Standard: NAACCR
Source Item Name: Derived from *Race 1*, *Race 2*, and *race- NAPIIA (derived API)*
Source Item Number: 160 (*Race 1*), 161 (*Race 2*), and 192 (*race- NAPIIA (derived API)*)

Description

This variable indicates the derived code for the patient's race. Race is coded separately from Hispanic ethnicity. This variable is created using NAACCR variables *Race1*, *Race2*, the Indian Health Service (IHS) Link variable, and *race-NAPIIA (derived API)*. Race recode starts as *Race1*. If *Race1* is white and *Race 2* is a specified non-white race, then the value from *Race2* is used. After this check, if Race is still white and there is a positive IHS link, then *Race/Ethnicity* is set to American Indian/Alaskan Native (AI/AN).

Considerations for use

- This variable is available only in the 2001–2015 public use database; it is not available in the 2005–2015 database.
- States have the option to suppress race-specific and Hispanic-specific data every submission year. While these states can be included in an aggregated analysis, their race and ethnicity information cannot be reported at the state level.
- The following states have state-level race data presentation restrictions:
  - Data for American Indians and Alaska Natives cannot be displayed for Delaware, Illinois, Kansas, Kentucky, New Jersey, and New York.
  - Data for Asians and Pacific Islanders cannot be displayed for Delaware, Kansas and Kentucky.
  - Data for race and ethnicity combinations—white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for Delaware, Kansas, Kentucky, Massachusetts, and Pennsylvania.
- Race is defined by specific physical, hereditary, and cultural traditions or origins, not necessarily by birthplace, place of residence, or citizenship. "Origin" is defined by the U.S. Census Bureau as the heritage, nationality group, lineage, or in some cases, the country of birth of the person or the person's parents or ancestors before their arrival in the United States.
- IHS provides medical services to American Indians and Alaska Natives (AI/ANs) who are members of federally recognized tribes, estimated to be about 65% of the AI/AN population. To improve identification of AI/ANs, 31 NPCR registries with Purchase/Referred Care Service Delivery Area (PRCSDA) counties in their state annually link with the IHS patient registration database to identify AI/ANs who had not been coded as AI/AN (shown in Appendix A). All NPCR registries link every five years; linkages were performed by all NPCR states most recently in 2016. In 2016, SEER registries linked cancer cases diagnosed from 1994–2014.
  - When interpreting data results, be aware that AI/AN populations may be undercounted during years when linkages did not occur in all NPCR registries.
  - If a project is looking specifically at AI/AN populations, analysts may consider restricting the NPCR states included in the analysis to NPCR registries that conduct annual IHS linkages. See Appendix A for the list of these states.
- In all separate records of tumors for the same patient, the patient should have the same race code.
- The *Race Recode for USCS* variable contains "other unspecified" and "unknown" categories. These groups are coded as "unknown race" for the purpose of analyses as specified in the SEER documentation https://seer.cancer.gov/seerstat/variables/seer/race_ethnicity. Population data are not available for the "other race" and "unknown race" categories.
- For more information, please see
  - NAACCR data dictionary www.naaccr.org/StandardsandRegistryOperations/VolumeII.aspx
  - 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
White	21,036,166	85.2%
Black	2,532,714	10.3%
American Indian/Alaska Native	121,943	0.5%
Asian or Pacific Islander	664,714	2.7%
Other unspecified (1991+)	79,546	0.3%
Unknown	262,254	1.1%

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	23,838,571	96.5%
SEER	858,766	3.5%

SEER*Stat Item Name: **Region**

Source of Standard: NPCR
Source Item Name: Derived from NAACCR's *State at diagnosis (CoC)* and US Census Region
Source Item Number: Derived from NAACCR's 80

Description

This variable indicates the U.S. Census region in which the patient lived at the time of diagnosis. The NAACCR data item *Address at Diagnosis-state* is recoded into one of the four U.S. Census regions: Northeast, Midwest, South, and West.

Considerations for use

- There is a potential for bias in the incidence rates for Census regions as only data from state registries that met U.S. Cancer Statistics publication criteria are included in the database. It is recommended that age-adjusted incidence rates for U.S. Census regions be presented only if:
  - At least 80% of the population for the Census region was covered by cancer registries that met U.S. Cancer Statistics publication criteria.
  - The 95% confidence intervals around the observed age-adjusted regional incidence rates based on data from eligible registries for each of six major cancer sites (prostate, female breast, male colorectal, female colorectal, male lung and bronchus, and female lung and bronchus) included the estimate of the regional rate calculated using the specified: www.cdc.gov/cancer/npcr/uscs/data/00_bias_correction.htm.
- If any state in a region has a case count of fewer than 16, then the case counts for U.S. Census regions cannot be presented.
- See www.census.gov/geo/reference/gtc/gtc_census_divreg.html for a list of states in each region.

Values	Frequency	Percentage
Northeast	5,146,029	20.8%
Midwest	5,585,842	22.6%
South	8,907,677	36.1%
West	5,057,789	20.5%

SEER*Stat Item Name: **USCS0115**

Source of Standard: NPCR

Source Item Name: Derived from NAACCR's *State at diagnosis (CoC)* and U.S. Cancer Statistics publication standards

Source Item Number: Derived from NAACCR's 80

Description

This variable indicates whether a central cancer registry met the U.S. Cancer Statistics publication standard for all cancer sites combined each year in 2001–2015.

Considerations for use

- This variable is used for analysis of combined 2001–2015 data in the NPCR and SEER Incidence – U.S. Cancer Statistics 2001–2015 database. Data from states that did not meet the U.S. Cancer Statistics publication standard in any given year were excluded from the database for that year. Please refer to Table 1 for states available by year.
- If you are conducting a multiyear analysis and want to restrict the analysis to the states that met publication criteria for each of those years (for example, a trend analysis), we recommend using the predefined variables, *USCS1115* (includes diagnosis years 2011–2015), *USCS0615* (includes diagnosis years 2006–2015), or *USCS0115* (includes diagnosis years 2001–2015).
- If you would like to analyze a range of years other than those predefined variables, please contact CDC at uscsdata@cdc.gov and we will create a new variable for you that can be imported into SEER*Stat.
- The U.S. Cancer Statistics publication standard is available at www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm.

Values	Frequency	Percentage
Does not meet USCS standard from 2001–2015	205,902	0.8%
Meets USCS standard from 2001–2015	24,491,435	99.2%

SEER*Stat Item Name: **USCS0615**

Source of Standard: NPCR

Source Item Name: Derived from NAACCR's *State at diagnosis (CoC)* and U.S. Cancer Statistics publication standard

Source Item Number: Derived from NAACCR's 80

Description

This variable indicates whether a central cancer registry met the U.S. Cancer Statistics publication standard for all cancer sites combined each year in 2006–2015. When using this variable, restrict the diagnosis years to 2006–2015. This is done in SEER*Stat on the Selection tab using the "Year of diagnosis" variable.

Considerations for use

- This variable is used for analysis of combined 2006–2015 data in the NPCR and SEER Incidence – U.S. Cancer Statistics 2001–2015 database. Data from states that did not meet the U.S. Cancer Statistics publication standard in any given year were excluded from the database for that year. Please refer to Table 1 for states available by year.
- If you are conducting a multiyear analysis and want to restrict the analysis to the states that met publication criteria for each of those years (for example, a trend analysis), we recommend using the predefined variables, *USCS1115* (includes diagnosis years 2011–2015), *USCS0615* (includes diagnosis years 2006–2015), or *USCS0115* (includes diagnosis years 2001–2015).
- If you would like to analyze a range of years other than those predefined variables, please contact CDC at uscsdata@cdc.gov and we will create a new variable for you that can be imported into SEER*Stat.
- The U.S. Cancer Statistics publication standard is available at www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm.

Note: This frequency table is restricted to cancers diagnosed during 2006–2015.

Values	Frequency	Percentage
Meets USCS standard from 2006–2015	18,786,786	100.0%

SEER*Stat Item Name: **USCS1115**

Source of Standard: NPCR

Source Item Name: Derived from NAACCR's *State at diagnosis (CoC)* and U.S. Cancer Statistics publication standard

Source Item Number: Derived from NAACCR's 80

Description

This variable indicates whether a central cancer registry met the U.S. Cancer Statistics publication standard for all cancer sites combined each year in 2011–2015. When using this variable, restrict the diagnosis years to 2011–2015. This is done in SEER*Stat on the Selection tab using the "Year of diagnosis" variable.

Considerations for use

- This variable is used for analysis of combined 2011–2015 data in the NPCR and SEER Incidence – U.S. Cancer Statistics 2001–2015 database. Data from states that did not meet the U.S. Cancer Statistics publication standard in any given year were excluded from the database for that year. Please refer to Table 1 for states available by year.
- If you are conducting a multiyear analysis and want to restrict the analysis to the states that met publication criteria for each of those years (for example, a trend analysis), we recommend using the predefined variables, *USCS1115* (includes diagnosis years 2011–2015), *USCS0615* (includes diagnosis years 2006–2015), or *USCS0115* (includes diagnosis years 2001–2015).
- If you would like to analyze a range of years other than those predefined variables, please contact CDC at uscsdata@cdc.gov and we will create a new variable for you that can be imported into SEER*Stat.
- The U.S. Cancer Statistics publication standard is available at www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm.

Note: This frequency table is restricted to cancers diagnosed during 2011–2015.

Values	Frequency	Percentage
Meets USCS standard from 2011–2015	8,832,656	100.0%

SEER*Stat Item Name: **Origin recode NHIA (Hispanic, Non-Hisp)**
Source of Standard: NAACCR
Source Item Name: *NHIA derived Hisp Origin*
Source Item Number: 191

Description

This variable was derived from the NAACCR standard variables: *Spanish/Hispanic Origin [190]*; *Name-Last [2230]*; *Name-Maiden [2390]*; *Birthplace [250]*; *Race 1 [160]*; *IHS Link [192]*; and *Sex [220]*.

NAACCR Hispanic Identification Algorithm uses the combination of these variables to directly or indirectly classify cases as Hispanic for analytic purposes.

Considerations for use

- This variable is available only in the 2001–2015 database.
- States have the option to suppress race-specific and Hispanic-specific data every submission year. While these states can be included in an aggregated analysis, the state's race and ethnicity information cannot be reported at the state level.
- The following states have state-level race or ethnicity data presentation restrictions:
  - Data for Hispanic ethnicity cannot be displayed for Delaware, Kentucky, and Massachusetts.
  - Data for race and ethnicity combinations—white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for Delaware, Kansas, Kentucky, Massachusetts, and Pennsylvania.
- Blank values are allowed for states that chose not to include data for NHIA in this file.
- For more information, please see:
  NAACCR Race and Ethnicity Work Group. *NAACCR Guideline for Enhancing Hispanic/Latino Identification: Revised NAACCR Hispanic/Latino Identification Algorithm [NHIA v2.2.1].* Springfield (IL): North American Association of Central Cancer Registries. September 2011. Available at www.naaccr.org/wp-content/uploads/2016/11/NHIA_v2_2_1_09122011.pdf.

Values	Frequency	Percentage
Non-Spanish-Hispanic-Latino	23,048,725	93.3%
Spanish-Hispanic-Latino	1,645,247	6.7%
Invalid Value(s)	3,365	0.0%

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)	24,290,151	98.4%
C80.9 (Unknown primary site)	407,186	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	24,697,337	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	2,193,382	8.9%
Moderately differentiated; Grade II	6,116,833	24.8%
Poorly differentiated; Grade III	4,944,068	20.0%
Undifferentiated; anaplastic; Grade IV	796,425	3.2%
T-cell	84,596	0.3%
B-cell; pre-B; B-precursor	1,178,006	4.8%
Null cell; non T-non B	2,601	0.0%
NK cell; natural killer cell (1995+)	3,781	0.0%
Unknown	9,377,645	38.0%

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)	23,238,060	94.1%
Positive histology	22,275,504	90.2%
Positive exfoliative cytology, no positive histology	771,384	3.1%
Positive histology AND immunophenotyping AND/OR positive genetic studies	163,517	0.7%
Positive microscopic confirm, method not specified	27,655	0.1%
Positive laboratory test/marker study	105,736	0.4%
Direct visualization without microscopic confirmation	33,403	0.1%
Radiography without microscopic confirm	870,781	3.5%
Clinical diagnosis only	144,964	0.6%
Unknown	304,392	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

- This variable is a 5-digit ICD-O-3 morphology code. The first 4 digits indicate the histology (cell type), and the fifth digit is the behavior code. Please note that the ICD-O-3 morphology codes have been grouped by major morphology headings as found in the International Classification of Diseases for Oncology, Third Edition in the frequency table shown below. However, the morphology codes are not grouped in the database.
- 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 at 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	1,885,280	7.6%
809–811	Basal Cell Neoplasms	7,024	0.0%
812–813	Transitional Cell Papillomas and Carcinomas	1,075,034	4.4%
814–838	Adneomas and Adenocarcinomas	9,723,988	39.4%
839–842	Adnexal and Skin Appendage Neoplasms	28,099	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	6,930	0.0%
868–871	Paragangliomas and Glomus Tumors	4,015	0.0%
872–879	Nevi and Melanomas	1,624,817	6.6%
880	Soft Tissue Tumors and Sarcomas, NOS	50,955	0.2%
881–883	Fibromatous Neoplasms	54,321	0.2%
884	Myxomatous Neoplasms	1,219	0.0%
885–888	Lipomatous Neoplasms	36,227	0.1%
889–892	Myomatous Neoplasms	57,810	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	5,702	0.0%

ICD–O–3 Code	Label	Frequency	Percentage
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%
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	297,796	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	1,026,813	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	5,861	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)	**24,697,337**	100.0%
Oral Cavity and Pharynx	566,085	2.3%
Lip	34,721	0.1%
Tongue	164,661	0.7%
Salivary Gland	59,986	0.2%
Floor of Mouth	33,404	0.1%
Gum and Other Mouth	78,726	0.3%
Nasopharynx	26,253	0.1%
Tonsil	95,465	0.4%
Oropharynx	24,635	0.1%
Hypopharynx	34,946	0.1%
Other Oral Cavity and Pharynx	13,288	0.1%
Digestive System	4,156,225	16.8%
Esophagus	237,011	1.0%
Stomach	327,562	1.3%
Small Intestine	105,656	0.4%
Colon and Rectum	2,289,363	9.3%
Colon excluding Rectum	1,647,222	6.7%
Cecum	354,900	1.4%
Appendix	41,011	0.2%
Ascending Colon	316,265	1.3%
Hepatic Flexure	80,068	0.3%
Transverse Colon	149,551	0.6%
Splenic Flexure	50,664	0.2%
Descending Colon	97,478	0.4%
Sigmoid Colon	449,099	1.8%
Large Intestine, NOS	108,186	0.4%
Rectum and Rectosigmoid Junction	642,141	2.6%
Rectosigmoid Junction	167,444	0.7%
Rectum	474,697	1.9%
Anus, Anal Canal and Anorectum	96,158	0.4%
Liver and Intrahepatic Bile Duct	329,697	1.3%
Liver	293,569	1.2%

Values	Frequency	Percentage
Intrahepatic Bile Duct	36,128	0.1%
Gallbladder	55,878	0.2%
Other Biliary	82,500	0.3%
Pancreas	565,056	2.3%
Retroperitoneum	18,185	0.1%
Peritoneum, Omentum and Mesentery	28,273	0.1%
Other Digestive Organs	20,886	0.1%
Respiratory System	3,323,799	13.5%
Nose, Nasal Cavity and Middle Ear	33,625	0.1%
Larynx	199,260	0.8%
Lung and Bronchus	3,080,667	12.5%
Pleurae	1,426	0.0%
Trachea, Mediastinum and Other Respiratory Organs	8,821	0.0%
Bones and Joints	43,472	0.2%
Soft Tissue including Heart	151,780	0.6%
Skin excluding Basal and Squamous	1,657,724	6.7%
Melanoma of the Skin	1,579,993	6.4%
Other Non-Epithelial Skin	77,731	0.3%
Breast (female and male combined)	4,014,362	16.3%
Female Genital System	1,321,637	5.4%
Cervix Uteri	188,349	0.8%
Corpus and Uterus, NOS	670,015	2.7%
Corpus Uteri	650,673	2.6%
Uterus, NOS	19,342	0.1%
Ovary	318,511	1.3%
Vagina	21,920	0.1%
Vulva	96,978	0.4%
Other Female Genital Organs	25,864	0.1%
Male Genital System	3,241,615	13.1%
Prostate	3,086,534	12.5%
Testis	121,894	0.5%
Penis	27,543	0.1%
Other Male Genital Organs	5,644	0.0%
Urinary System	1,838,028	7.4%
Urinary Bladder	1,015,078	4.1%
Kidney and Renal Pelvis	764,418	3.1%
Ureter	40,439	0.2%
Other Urinary Organs	18,093	0.1%
Eye and Orbit	45,837	0.2%
Brain and Other Nervous System	751,495	3.0%
Brain	335,120	1.4%
Cranial Nerves Other Nervous System	416,375	1.7%
Endocrine System	745,141	3.0%
Thyroid	564,660	2.3%
Other Endocrine including Thymus	180,481	0.7%
Lymphoma	1,050,216	4.3%
Hodgkin Lymphoma	126,441	0.5%
Hodgkin – Nodal	122,849	0.5%
Hodgkin – Extranodal	3,592	0.0%

Values	Frequency	Percentage
Non-Hodgkin Lymphoma	923,775	3.7%
NHL – Nodal	628,805	2.5%
NHL – Extranodal	294,970	1.2%
Myeloma	296,894	1.2%
Leukemia	628,694	2.5%
Lymphocytic Leukemia	314,359	1.3%
Acute Lymphocytic Leukemia	70,076	0.3%
Chronic Lymphocytic Leukemia	224,004	0.9%
Other Lymphocytic Leukemia	20,279	0.1%
Myeloid and Monocytic Leukemia	282,521	1.1%
Acute Myeloid Leukemia	182,895	0.7%
Acute Monocytic Leukemia	11,293	0.0%
Chronic Myeloid Leukemia	79,962	0.3%
Other Myeloid/Monocytic Leukemia	8,371	0.0%
Other Leukemia	31,814	0.1%
Other Acute Leukemia	11,981	0.0%
Aleukemic, Subleukemic and NOS	19,833	0.1%
Mesothelioma	47,166	0.2%
Kaposi Sarcomae	19,157	0.1%
Miscellaneous	798,010	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	57,579	23.5%
I(a) Lymphoid leukemias	41,105	16.7%
I(b) Acute myeloid leukemias	9,980	4.1%
I(c) Chronic myeloproliferative diseases	2,913	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,866	0.8%
I(e) Unspecified and other specified leukemias	1,715	0.7%
II Lymphomas and reticuloendothelial neoplasms	33,018	13.5%
II(a) Hodgkin lymphomas	15,381	6.3%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	11,040	4.5%
II(c) Burkitt lymphoma	3,015	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	3,162	1.3%
II(e) Unspecified lymphomas	420	0.2%
III CNS and misc intracranial and intraspinal neoplasms	55,162	22.5%
III(a) Ependymomas and choroid plexus tumor	4,493	1.8%
III(b) Astrocytomas	20,604	8.4%
III(c) Intracranial and intraspinal embryonal tumors	7,578	3.1%
III(d) Other gliomas	6,880	2.8%
III(e) Other specified intracranial/intraspinal neoplasms	13,679	5.6%
III(f) Unspecified intracranial and intraspinal neoplasms	1,928	0.8%
IV Neuroblastoma and other peripheral nervous cell tumors	10,687	4.4%
IV(a) Neuroblastoma and ganglioneuroblastoma	10,383	4.2%
IV(b) Other peripheral nervous cell tumors	304	0.1%
V Retinoblastoma	3,969	1.6%
VI Renal tumors	8,584	3.5%
VI(a) Nephroblastoma and other nonepithelial renal tumors	7,743	3.2%
VI(b) Renal carcinomas	811	0.3%
VI(c) Unspecified malignant renal tumors	30	0.0%
VII Hepatic tumors	2,852	1.2%
VII(a) Hepatoblastoma	2,087	0.9%
VII(b) Hepatic carcinomas	739	0.3%
VII(c) Unspecified malignant hepatic tumors	26	0.0%

Values	Frequency	Percentage
VIII Malignant bone tumors	11,131	4.5%
VIII(a) Osteosarcomas	6,263	2.6%
VIII(b) Chondrosarcomas	430	0.2%
VIII(c) Ewing tumor and related sarcomas of bone	3,658	1.5%
VIII(d) Other specified malignant bone tumors	559	0.2%
VIII(e) Unspecified malignant bone tumors	221	0.1%
IX Soft tissue and other extraosseous sarcomas	15,027	6.1%
IX(a) Rhabdomyosarcomas	5,855	2.4%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	1,643	0.7%
IX(c) Kaposi sarcomae	59	0.0%
IX(d) Other specified soft tissue sarcomas	5,855	2.4%
IX(e) Unspecified soft tissue sarcomas	1,615	0.7%
X Germ cell & trophoblastic tumors & neoplasms of gonads	14,386	5.9%
X(a) Intracranial & intraspinal germ cell tumors	2,542	1.0%
X(b) Extracranial & extragonadal germ cell tumors	1,817	0.7%
X(c) Malignant gonadal germ cell tumors	9,081	3.7%
X(d) Gonadal carcinomas	536	0.2%
X(e) Other and unspecified malignant gonadal tumors	410	0.2%
XI Other malignant epithelial neoplasms and melanomas	23,343	9.5%
XI(a) Adrenocortical carcinomas	265	0.1%
XI(b) Thyroid carcinomas	10,011	4.1%
XI(c) Nasopharyngeal carcinomas	710	0.3%
XI(d) Malignant melanomas	6,363	2.6%
XI(e) Skin carcinomas	105	0.0%
XI(f) Other and unspecified carcinomas	5,889	2.4%
XII Other and unspecified malignant neoplasms	946	0.4%
XII(a) Other specified malignant tumors	500	0.2%
XII(b) Other unspecified malignant tumors	446	0.2%
Not classified by ICCC or *in situ*	8,801	3.6%

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	57,579	23.5%
I(a) Lymphoid leukemias	41,105	16.7%
I(a.1) Precursor cell leukemias	39,795	16.2%
I(a.2) Mature B-cell leukemias	992	0.4%
I(a.3) Mature T-cell and NK cell leukemias	136	0.1%
I(a.4) Lymphoid leukemia, NOS	182	0.1%
I(b) Acute myeloid leukemias	9,980	4.1%
I(c) Chronic myeloproliferative diseases	2,913	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,866	0.8%
I(e) Unspecified and other specified leukemias	1,715	0.7%
II Lymphomas and reticuloendothelial neoplasms	33,018	13.5%
II(a) Hodgkin lymphomas	15,381	6.3%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	11,040	4.5%
II(b.1) Precursor cell lymphomas	3,201	1.3%
II(b.2) Mature B-cell lymphomas except Burkitt lymphoma	4,446	1.8%
II(b.3) Mature T-cell and NK-cell lymphomas	2,797	1.1%
II(b.4) Non-Hodgkin lymphomas, NOS	596	0.2%
II(c) Burkitt lymphoma	3,015	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	3,162	1.3%
II(e) Unspecified lymphomas	420	0.2%
III CNS and misc intracranial and intraspinal neoplasms	55,162	22.5%
III(a) Ependymomas and choroid plexus tumor	4,493	1.8%
III(a.1) Ependymomas	3,451	1.4%
III(a.2) Choroid plexus tumor	1,042	0.4%
III(b) Astrocytomas	20,604	8.4%
III(c) Intracranial and intraspinal embryonal tumors	7,578	3.1%
III(c.1) Medulloblastomas	4,853	2.0%
III(c.2) PNET	1,662	0.7%
III(c.3) Medulloepithelioma	66	0.0%
III(c.4) Atypical teratoid/rhabdoid tumor	997	0.4%

Values	Frequency	Percentage
III(d) Other gliomas	6,880	2.8%
III(d.1) Oligodendrogliomas	868	0.4%
III(d.2) Mixed and unspecified gliomas	5,865	2.4%
III(d.3) Neuroepithelial glial tumors of uncertain orig	147	0.1%
III(e) Other specified intracranial/intraspinal neoplasms	13,679	5.6%
III(e.1) Pituitary adenomas and carcinomas	5,635	2.3%
III(e.2) Tumors of sellar region (craniopharyngiomas)	2,027	0.8%
III(e.3) Pineal parenchymal tumors	564	0.2%
III(e.4) Neuronal and mixed neuronal-glial tumors	3,937	1.6%
III(e.5) Meningiomas	1,516	0.6%
III(f) Unspecified intracranial and intraspinal neoplasms	1,928	0.8%
IV Neuroblastoma and other peripheral nervous cell tumors	10,687	4.4%
IV(a) Neuroblastoma and ganglioneuroblastoma	10,383	4.2%
IV(b) Other peripheral nervous cell tumors	304	0.1%
V Retinoblastoma	3,969	1.6%
VI Renal tumors	^[1]	^[1]
VI(a) Nephroblastoma and other nonepithelial renal tumors	^[1]	^[1]
VI(a.1) Nephroblastoma	7,279	3.0%
VI(a.2) Rhabdoid renal tumor	199	0.1%
VI(a.3) Kidney sarcomas	253	0.1%
VI(a.4) pPNET of kidney	^[2]	^[2]
VI(b) Renal carcinomas	811	0.3%
VI(c) Unspecified malignant renal tumors	30	0.0%
VII Hepatic tumors	2,852	1.2%
VII(a) Hepatoblastoma	2,087	0.9%
VII(b) Hepatic carcinomas	739	0.3%
VII(c) Unspecified malignant hepatic tumors	26	0.0%
VIII Malignant bone tumors	11,131	4.5%
VIII(a) Osteosarcomas	6,263	2.6%
VIII(b) Chondrosarcomas	430	0.2%
VIII(c) Ewing tumor and related sarcomas of bone	3,658	1.5%
VIII(c.1) Ewing tumor and Askin tumor of bone	3,513	1.4%
VIII(c.2) pPNET of bone	145	0.1%
VIII(d) Other specified malignant bone tumors	559	0.2%
VIII(d.1) Malignant fibrous neoplasms of bone	55	0.0%
VIII(d.2) Malignant chordomas	263	0.1%
VIII(d.3) Odontogenic malignant tumors	68	0.0%
VIII(d.4) Miscellaneous malignant bone tumors	173	0.1%
VIII(e) Unspecified malignant bone tumors	221	0.1%
IX Soft tissue and other extraosseous sarcomas	15,027	6.1%
IX(a) Rhabdomyosarcomas	5,855	2.4%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	1,643	0.7%
IX(b.1) Fibroblastic and myofibroblastic tumors	855	0.3%
IX(b.2) Nerve sheath tumors	767	0.3%
IX(b.3) Other fibromatous neoplasms	21	0.0%
IX(c) Kaposi sarcomae	59	0.0%
IX(d) Other specified soft tissue sarcomas	5,855	2.4%
IX(d.1) Ewing tumor and Askin tumor of soft tissue	759	0.3%
IX(d.2) pPNET of soft tissue	359	0.1%
IX(d.3) Extrarenal rhabdoid tumor	306	0.1%
IX(d.4) Liposarcomas	341	0.1%

Values	Frequency	Percentage
IX(d.5) Fibrohistiocytic tumors	1,456	0.6%
IX(d.6) Leiomyosarcomas	240	0.1%
IX(d.7) Synovial sarcomas	1,310	0.5%
IX(d.8) Blood vessel tumors	222	0.1%
IX(d.9) Osseous & chondromatous neoplasms of soft tissue	124	0.1%
IX(d.10) Alveolar soft parts sarcoma	200	0.1%
IX(d.11) Miscellaneous soft tissue sarcomas	538	0.2%
IX(e) Unspecified soft tissue sarcomas	1,615	0.7%
X Germ cell & trophoblastic tumors & neoplasms of gonads	14,386	5.9%
X(a) Intracranial & intraspinal germ cell tumors	2,542	1.0%
X(a.1) Intracranial & intraspinal germinomas	1,579	0.6%
X(a.2) Intracranial & intraspinal teratomas	643	0.3%
X(a.3) Intracranial & intraspinal embryonal carcinomas	30	0.0%
X(a.4) Intracranial & intraspinal yolk sac tumor	38	0.0%
X(a.5) Intracranial & intraspinal choriocarcinoma	22	0.0%
X(a.6) Intracranial & intraspinal tumors of mixed forms	230	0.1%
X(b) Extracranial & extragonadal germ cell tumors	1,817	0.7%
X(b.1) Germinomas: extracranial/extragonadal	191	0.1%
X(b.2) Malignant teratomas: extracranial/extragonadal	714	0.3%
X(b.3) Embryonal carcinomas: extracranial/extragonadal	19	0.0%
X(b.4) Yolk sac tumor: extracranial/extragonadal	416	0.2%
X(b.5) Choriocarcinomas: extracranial/extragonadal	212	0.1%
X(b.6) Other mixed germ cell: extracranial/extragonadal	265	0.1%
X(c) Malignant gonadal germ cell tumors	9,081	3.7%
X(c.1) Malignant gonadal germinomas	1,916	0.8%
X(c.2) Malignant gonadal teratomas	1,600	0.7%
X(c.3) Gonadal embryonal carcinomas	899	0.4%
X(c.4) Gonadal yolk sac tumor	862	0.4%
X(c.5) Gonadal choriocarcinoma	88	0.0%
X(c.6) Malignant gonadal tumors of mixed forms	3,716	1.5%
X(d) Gonadal carcinomas	536	0.2%
X(e) Other and unspecified malignant gonadal tumors	410	0.2%
XI Other malignant epithelial neoplasms and melanomas	23,343	9.5%
XI(a) Adrenocortical carcinomas	265	0.1%
XI(b) Thyroid carcinomas	10,011	4.1%
XI(c) Nasopharyngeal carcinomas	710	0.3%
XI(d) Malignant melanomas	6,363	2.6%
XI(e) Skin carcinomas	105	0.0%
XI(f) Other and unspecified carcinomas	5,889	2.4%
XI(f.1) Carcinomas of salivary glands	1,100	0.4%
XI(f.2) Carcinomas of colon and rectum	733	0.3%
XI(f.3) Carcinomas of appendix	1,055	0.4%
XI(f.4) Carcinomas of lung	565	0.2%
XI(f.5) Carcinomas of thymus	87	0.0%
XI(f.6) Carcinomas of breast	243	0.1%
XI(f.7) Carcinomas of cervix uteri	180	0.1%
XI(f.8) Carcinomas of bladder	370	0.2%
XI(f.9) Carcinomas of eye	37	0.0%
XI(f.10) Carcinomas of other specified sites	1,324	0.5%
XI(f.11) Carcinomas of unspecified site	195	0.1%

XII Other and unspecified malignant neoplasms	^[1]	^[1]
XII(a) Other specified malignant tumors	^[1]	^[1]
XII(a.1) Gastrointestinal stromal tumor	117	0.0%
XII(a.2) Pancreatoblastoma	41	0.0%
XII(a.3) Pulmonary blastoma and pleuropulmonary blastoma	237	0.1%
XII(a.4) Other complex mixed and stromal neoplasms	56	0.0%
XII(a.5) Mesothelioma	48	0.0%
XII(a.6) Other specified malignant tumors	^[2]	^[2]
XII(b) Other unspecified malignant tumors	446	0.2%
Not classified by ICCC or *in situ*	8,801	3.6%

[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	26,160	6.2%
1.1 Acute lymphoid leukemia	10,176	2.4%
1.2 Acute myeloid leukemia	9,930	2.4%
1.3 Chronic myeloid leukemia	4,066	1.0%
1.4 Other and unspecified leukemia	1,988	0.5%
2 Lymphomas	61,163	14.5%
2.1 Non-Hodgkin lymphoma	23,730	5.6%
2.2 Hodgkin lymphoma	37,433	8.9%
3 CNS and Oth Intracranial and Intraspinal Neo (all behav)	37,757	8.9%
3.1. Astrocytoma	12,659	3.0%
3.1.1 Specified low-grade astrocytic tumors	5,011	1.2%
3.1.2 Glioblastoma and anaplastic astrocytoma	4,738	1.1%
3.1.3 Astrocytoma, NOS	2,910	0.7%
3.2 Other glioma	6,777	1.6%
3.3 Ependymoma	2,462	0.6%
3.4. Medulloblastoma and other PNET	2,149	0.5%
3.4.1 Medulloblastoma	1,217	0.3%
3.4.2 Supratentorial PNET	932	0.2%
3.5 Other specified intracranial and intraspinal neoplasms	11,634	2.8%
3.6 Unspecified intracranial and intraspinal neoplasms	2,076	0.5%
3.6.1 Unspec malignant intracranial and intraspinal neo	380	0.1%
3.6.2 Unspec ben/border intracran. and intraspin neo	1,696	0.4%
4 Osseous & Chondromatous Neoplasms	10,872	2.6%
4.1 Osteosarcoma	4,728	1.1%
4.2 Chondrosarcoma	1,479	0.4%
4.3 Ewing tumor	3,648	0.9%
4.4 Other specified and unspecified bone tumors	1,017	0.2%
5 Soft Tissue Sarcomas	17,914	4.2%
5.1 Fibromatous neoplasms	4,447	1.1%
5.2 Rhabdomyosarcoma	2,017	0.5%
5.3 Other soft tissue sarcoma	11,450	2.7%

Values	Frequency	Percentage
5.3.1 Specified soft tissue sarcoma	8,878	2.1%
5.3.1.1 Specified (excluding Kaposi sarcoma)	7,123	1.7%
5.3.1.2 Kaposi sarcomae	1,755	0.4%
5.3.2 Unspecified soft tissue sarcoma	2,572	0.6%
6 Germ Cell and Trophoblastic Neoplasms	49,652	11.8%
6.1 Germ cell and trophoblastic neoplasms of gonads	45,353	10.7%
6.2 Germ cell and trophoblastic neo of nongonadal sites	4,299	1.0%
6.2.1 Intracranial (all behaviors)	1,626	0.4%
6.2.2 Other nongonadal	2,673	0.6%
7 Melanoma and Skin Carcinomas	38,808	9.2%
7.1 Melanoma	38,474	9.1%
7.2 Skin carcinomas	334	0.1%
8 Carcinomas	123,869	29.3%
8.1 Thyroid carcinoma	53,525	12.7%
8.2 Other carcinoma of head and neck	6,993	1.7%
8.2.1 Nasopharyngeal carcinoma	1,278	0.3%
8.2.2 Other sites in lip, oral cavity and pharynx	5,045	1.2%
8.2.3 Nasal cav,mid ear,sinus,larynx,ill-def head/neck	670	0.2%
8.3 Carcinoma of trachea,bronchus, and lung	2,941	0.7%
8.4 Carcinoma of breast	15,544	3.7%
8.5 Carcinoma of genitourinary tract	24,622	5.8%
8.5.1 Carcinoma of kidney	5,076	1.2%
8.5.2 Carcinoma of bladder	2,146	0.5%
8.5.3 Carcinoma of gonads	3,817	0.9%
8.5.4 Carcinoma of cervix and uterus	12,912	3.1%
8.5.5 Carc of oth and ill-defined sites	671	0.2%
8.6 Carcinoma of gastrointestinal tract	17,968	4.3%
8.6.1 Carcinoma of colon and rectum	12,035	2.9%
8.6.2 Carcinoma of stomach	1,752	0.4%
8.6.3 Carcinoma of liver and intrahepatic bile ducts	1,687	0.4%
8.6.4 Carcinoma of pancreas	1,263	0.3%
8.6.5 Carc oth and ill-def sites, gastrointestinal tract	1,231	0.3%
8.7 Carcinoma of other and ill-defined sites	2,276	0.5%
8.7.1 Adrenocortical carcinoma	356	0.1%
8.7.2 Carcinoma of other and ill-defined sites, NOS	1,920	0.5%
9 Miscellaneous specified neoplasms, NOS	9,596	2.3%
9.1 Other pediatric and embryonal tumors, NOS	973	0.2%
9.1.1 Wilms tumor	189	0.0%
9.1.2 Neuroblastoma	283	0.1%
9.1.3 Other pediatric and embryonal tumors, NOS	501	0.1%
9.2 Other specified and embryonal tumors, NOS	8,623	2.0%
9.2.1 Paraganglioma and glomus tumors	360	0.1%
9.2.2 Other specified gonadal tumors	726	0.2%
9.2.3 Myeloma, mast cell, misc lymphoreticular neo, NOS	1,545	0.4%
9.2.4 Other specified neoplasms, NOS	5,992	1.4%
10 Unspecified Malignant Neoplasms	2,243	0.5%
Unclassified and Non-Malignant	44,149	10.5%

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	1,685,399	6.8%
1 Hodgkin Lymphoma	126,441	0.5%
1(a) Classical Hodgkin lymphoma	119,713	0.5%
1(a)1 Lymphocyte-rich/mixed cell/lymphocyte depleted	20,415	0.1%
1(a)1.1 Lymphocyte-rich	4,412	0.0%
1(a)1.2 Mixed cellularity	14,447	0.1%
1(a)1.3 Lymphocyte-depleted	1,556	0.0%
1(a)2 Nodular sclerosis	68,641	0.3%
1(a)3 Classical Hodgkin lymphoma, NOS	30,657	0.1%
1(b) Nodular lymphocyte predominant Hodgkin lymphoma	6,728	0.0%
2 Non-Hodgkin lymphoma	1,509,483	6.1%
2(a) Non-Hodgkin lymphoma, B-cell	1,390,675	5.6%
2(a)1 Precursor Non-Hodgkin lymphoma, B-cell	54,691	0.2%
2(a)2 Mature Non-Hodgkin lymphoma, B-cell	1,264,840	5.1%
2(a)2.1 Chronic/Sm/Prolymphocytic/Mantle B-cell NHL	316,946	1.3%
2(a)2.1.1 Chronic/Small lymphocytic leuk/lymph	277,187	1.1%
2(a)2.1.2 Prolymphocytic leukemia, B-cell	1,287	0.0%
2(a)2.1.3 Mantle-cell lymphoma	38,472	0.2%
2(a)2.2 Lymphoplasmacytic lymphoma/Waldenstrom	28,877	0.1%
2(a)2.2.1 Lymphoplasmacytic lymphoma	12,344	0.0%
2(a)2.2.2 Waldenstrom macroglubulinemia	16,533	0.1%
2(a)2.3 Diffuse large B-cell lymphoma (DLBCL)	324,991	1.3%
2(a)2.3.1 DLBCL, NOS	321,690	1.3%
2(a)2.3.2 Intravascular large B-cell lymphoma	584	0.0%
2(a)2.3.3 Primary effusion lymphoma	422	0.0%
2(a)2.3.4 Mediastinal large B-cell lymphoma	2,295	0.0%
2(a)2.4 Burkitt lymphoma/leukemia	18,378	0.1%

Values	Frequency	Percentage
2(a)2.5 Marginal-zone lymphoma (MZL)	86,911	0.4%
2(a)2.5.1 Splenic MZL	7,988	0.0%
2(a)2.5.2 Extranodal MZL, MALT type	52,275	0.2%
2(a)2.5.3 Nodal MZL	26,648	0.1%
2(a)2.6 Follicular lymphoma	176,694	0.7%
2(a)2.7 Hairy-cell leukemia	13,732	0.1%
2(a)2.8 Plasma cell neoplasms	298,072	1.2%
2(a)2.8.1 Plasmacytoma	19,532	0.1%
2(a)2.8.2 Multiple myeloma/plasma-cell leuk	278,540	1.1%
2(a)2.9 Heavy chain disease	239	0.0%
2(a)3 Non-Hodgkin lymphoma, B-cell, NOS	71,144	0.3%
2(b) Non-Hodgkin lymphoma, T-cell	97,443	0.4%
2(b)1 Precursor Non-Hodgkin lymphoma, T-cell	4,775	0.0%
2(b)2 Mature Non-Hodgkin lymphoma, T-cell	92,124	0.4%
2(b)2.1 Mycosis fungoides/Sezary syndrome	20,041	0.1%
2(b)2.1.1 Mycosis fungoides	19,296	0.1%
2(b)2.1.2 Sezary syndrome	745	0.0%
2(b)2.2 Peripheral T-cell lymphoma	53,249	0.2%
2(b)2.2.1 Peripheral T-cell lymphoma, NOS	18,740	0.1%
2(b)2.2.2 Angioimmunoblastic T-cell lymphoma	5,856	0.0%
2(b)2.2.3 Subcutan panniculitis-like T-cell lymph	488	0.0%
2(b)2.2.4 Anaplastic lar cell lymph, T-/Null-cell	11,993	0.0%
2(b)2.2.5 Hepatosplenic T-cell lymphoma	414	0.0%
2(b)2.2.6 Enteropathy-type T-cell lymphoma	643	0.0%
2(b)2.2.7 Cutaneous T-cell lymphoma, NOS	10,979	0.0%
2(b)2.2.8 Prim cutaneous anaplastic lar cell lymph	4,136	0.0%
2(b)2.3 Adult T-cell leukemia/lymphoma	9,818	0.0%
2(b)2.4 NK/T-cell lymph, nasal-type/aggres NK leuk	2,842	0.0%
2(b)2.5 T-cell large granular lymphocytic leukemia	4,468	0.0%
2(b)2.6 Prolymphocytic leukemia, T-cell	1,706	0.0%
2(b)3 Non-Hodgkin lymphoma, NOS, T-cell	544	0.0%
2(c) Non-Hodgkin lymphoma, unknown lineage	21,365	0.1%
2(c)1 Precursor lymphoblastic leuk/lymph, unk lineage	8,281	0.0%
2(c)2 Prolymphocytic leukemia, unknown lineage	634	0.0%
2(c)3 Non-Hodgkin lymphoma, NOS, unknown lineage	12,450	0.1%
3 Composite Hodgkin lymphoma and NHL	2,663	0.0%
4 Lymphoid neoplasm, NOS	46,812	0.2%
Unclassified	23,011,938	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	535,691	2.2%
Borderline malignancy	52,818	0.2%
In situ	1,697,196	6.9%
Malignant	22,052,486	89.3%
Only malignant in ICD-O-3	339,408	1.4%
Only malignant 2010+	19,738	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 2000* was blank, but *Summary Stage 2000* 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	2,108,909	8.5%
Localized only	10,124,448	41.0%
Regional, direct extension only	1,642,307	6.6%
Regional, regional lymph nodes only	1,741,130	7.0%
Regional, direct extension and regional lymph nodes	1,020,484	4.1%
Regional, NOS	252,125	1.0%
Distant site(s)/node(s) involved	5,390,351	21.8%
Not applicable	588,327	2.4%
Unknown/unstaged/unspecified	1,828,404	7.4%
Blanks	852	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	13,579,791	55.0%
Right - origin of primary	5,384,217	21.8%
Left - origin of primary	4,985,464	20.2%
Only one side - side unspecified	50,969	0.2%
Bilateral, single primary	191,390	0.8%
Paired site: midline tumor	51,554	0.2%
Paired site, but no information concerning laterality	453,952	1.8%

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	17,803,669	^[1]
1st of 2 or more primaries	2,080,729	^[1]
2nd of 2 or more primaries	3,501,508	^[1]
3rd of 3 or more primaries	588,456	^[1]
4th of 4 or more primaries	102,872	^[1]
5th of 5 or more primaries	21,133	^[1]
6th or more primaries[2]	9,999	^[1]
Only one state registry-defined neoplasm	563,211	^[1]
1st of 2 or more state registry-defined neoplasms	11,212	^[1]
2nd of 2 or more state registry-defined neoplasms	12,678	^[1]
3rd of 3 or more state registry-defined neoplasms	877	^[1]
4th of 4 or more state registry-defined neoplasms	220	^[1]
5th of 5 or more state registry-defined neoplasms	101	^[1]
6th or more state registry-defined neoplasms[1]	61	^[1]

Values	Frequency	Percentage
Unknown sequence number - federally required *in situ* or malignant tumors	115	^[1]
Carcinoma *in situ* of the Cervix diagnosed 1/1/1996 or later	^[3]	^[3]
Unknown sequence number - state registry-defined neoplasms	331	^[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: **NHIA Derived Hisp Origin**
Source of Standard: NAACCR
Source Item Name: *NHIA Derived Hispanic Origin*
Source Item Number: 191

Description

The NAACCR Hispanic Identification Algorithm uses a combination of data items to directly or indirectly classify cases as Hispanic for analytic purposes. Cases are classified based on individual's birth place (Non-Hispanic, Mexican, Puerto Rican, Cuban).

Considerations for use

- **This variable includes only count data**; rates cannot be calculated using this variable, as no population data are associated with the variable. Use the variable *Origin recode NHIA (Hispanic/Non-Hispanic)* to calculate age-adjusted rates by ethnicity.
- Blank values are allowed for states that chose not to include data for NHIA in this file.
- The following states have state-level race or ethnicity data presentation restrictions:
  - Data for Hispanic ethnicity cannot be displayed for Delaware, Kentucky, and Massachusetts.
  - Data for race and ethnicity combinations—white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for Delaware, Kansas, Kentucky, Massachusetts, and Pennsylvania.

Values	Frequency	Percentage
Non-Spanish-Hispanic-Latino	23,048,725	93.3%
Mexican	382,103	1.5%
Puerto Rican	89,712	0.4%
Cuban	69,517	0.3%
South or Central American excluding Brazil	127,888	0.5%
Other specified Spanish/Hispanic origin including Europe	39,532	0.2%
Spanish/Hispanic/Latino, NOS	687,221	2.8%
NHIA surname match only	218,911	0.9%
Dominican Republic	30,363	0.1%
Invalid Value(s)	3,365	0.0%

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 it.

Values	Frequency	Percentage
1890	^1	^1
1891	^1	^1
1892	^1	^1
1893	^1	^1
1894	17	^2
1895	28	^2
1896	44	^2
1897	70	^2
1898	145	^2
1899	222	^2
1900	559	^2
1901	855	^2
1902	1,271	^2
1903	1,891	^2
1904	2,943	^2
1905	4,553	^2
1906	6,528	^2
1907	9,991	^2
1908	14,351	^2
1909	19,413	^2
1910	27,023	^2
1911	35,068	^2
1912	48,346	^2
1913	61,150	^2
1914	79,105	^2
1915	95,225	^2
1916	115,586	^2
1917	140,578	^2
1918	171,067	^2

Values	Frequency	Percentage
1919	194,407	^2
1920	242,160	^2
1921	284,655	^2
1922	309,962	^2
1923	345,294	^2
1924	386,090	^2
1925	411,969	^2
1926	439,748	^2
1927	475,984	^2
1928	491,301	^2
1929	503,298	^2
1930	532,993	^2
1931	528,884	^2
1932	541,150	^2
1933	528,435	^2
1934	558,047	^2
1935	570,976	^2
1936	574,669	^2
1937	585,606	^2
1938	601,628	^2
1939	591,105	^2
1940	598,860	^2
1941	613,016	^2
1942	658,948	^2
1943	656,177	^2
1944	601,768	^2
1945	567,968	^2
1946	642,688	^2
1947	680,654	^2
1948	612,022	^2
1949	575,609	^2
1950	537,766	^2
1951	526,213	^2
1952	509,015	^2
1953	482,464	^2
1954	467,178	^2
1955	439,876	^2
1956	422,073	^2
1957	401,478	^2
1958	368,509	^2
1959	344,141	^2
1960	319,779	^2
1961	294,074	^2

Values	Frequency	Percentage
1962	266,335	^2
1963	242,574	^2
1964	217,919	^2
1965	187,795	^2
1966	166,629	^2
1967	149,024	^2
1968	137,414	^2
1969	128,717	^2
1970	121,710	^2
1971	107,169	^2
1972	92,330	^2
1973	80,405	^2
1974	74,869	^2
1975	68,111	^2
1976	62,553	^2
1977	59,156	^2
1978	54,387	^2
1979	52,133	^2
1980	48,334	^2
1981	44,638	^2
1982	41,253	^2
1983	36,736	^2
1984	33,840	^2
1985	30,722	^2
1986	27,775	^2
1987	25,382	^2
1988	23,568	^2
1989	21,499	^2
1990	19,645	^2
1991	17,725	^2
1992	15,817	^2
1993	14,111	^2
1994	13,022	^2
1995	11,665	^2
1996	11,001	^2
1997	10,621	^2
1998	11,074	^2
1999	10,617	^2
2000	10,730	^2
2001	10,487	^2
2002	9,773	^2
2003	9,240	^2
2004	8,719	^2

Values	Frequency	Percentage
2005	8,118	^[2]
2006	7,854	^[2]
2007	7,364	^[2]
2008	6,962	^[2]
2009	6,083	^[2]
2010	5,449	^[2]
2011	4,515	^[2]
2012	3,711	^[2]
2013	2,874	^[2]
2014	1,769	^[2]
2015	724	^[2]
Blank(s)	^[1]	^[1]

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

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

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

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

Description

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

Considerations for use

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

Values	Frequency	Percentage
January	2,173,217	8.8%
February	1,916,744	7.8%
March	2,100,813	8.5%
April	2,072,940	8.4%
May	2,080,235	8.4%
June	2,127,250	8.6%
July	2,035,891	8.2%
August	2,078,561	8.4%
September	1,985,790	8.0%
October	2,117,729	8.6%
November	1,932,001	7.8%
December	1,910,900	7.7%
Blank(s)	165,266	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	21,576,223	87.4%
Radiation treatment or medical oncology center (2006+)	486,696	2.0%
Laboratory only (hospital or private)	655,254	2.7%
Physician's office/private medical practitioner (LMD)	1,046,905	4.2%
Nursing/convalescent home/hospice	32,472	0.1%
Other hospital outpatient unit or surgery center (2006+)	899,508	3.6%
Blanks	279	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	557,838	7.6%
Esophagus	190,967	2.6%
Colon and rectum	2,239,388	30.5%
Liver	198,029	2.7%
Larynx	196,067	2.7%
Female breast cancer	3,949,589	53.9%

[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	222,889	36.5%
Anal and rectal squamous cell carcinoma	91,835	15.0%
Vulvar squamous cell carcinoma	74,850	12.3%
Vaginal squamous cell carcinoma	14,436	2.4%
Penile squamous cell carcinoma	25,725	4.2%
Cervical carcinoma	180,996	29.6%

[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	144,145	1.6%
Gastric cardia	99,026	1.1%
Colon & rectum	2,239,388	25.2%
Liver	198,029	2.2%
Gallbladder	52,456	0.6%
Pancreas	473,071	5.3%
Kidney	644,350	7.3%
Meningioma	138,920	1.6%
Thyroid	561,932	6.3%
Multiple myeloma	248,709	2.8%
Post-menopausal female breast	3,110,221	35.1%
Corpus and uterus, NOS (not otherwise specified)	663,852	7.5%

[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.

Ovary	299,079	3.4%

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	1,606,225	29.9%
Postmenopausal female breast	3,110,221	57.8%
Corpus and uterus, NOS (not otherwise specified)	663,852	12.3%

[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	557,838	6.1%
Esophagus	230,161	2.5%
Stomach	319,962	3.5%
Colon and rectum	2,239,388	24.7%
Liver	198,029	2.2%
Pancreas	473,071	5.2%
Larynx	196,067	2.2%
Trachea, lung, bronchus	2,802,226	30.9%
Cervix uteri	185,004	2.0%
Kidney and renal pelvis	698,657	7.7%
Urinary bladder	1,001,899	11.0%
Acute myeloid leukemia	174,347	1.9%

[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.

SEER*Stat Item Name: **State race eth suppress**
Source of Standard: NPCR
Source Item Name: Derived from *Addr at DX - state* and state-level race or ethnicity reporting restrictions
Source Item Number: Derived from NAACCR 80 (Addr at Dx - state)

Description

This variable was created specifically for this database. It provides the selection of states that are eligible to be included in a state-level analysis of race, ethnicity, or race and ethnicity combination.

Considerations for use

- States have the option to suppress race-specific and Hispanic ethnicity-specific data every submission year. While these states can be included in an aggregated analysis, the affected state's race and ethnicity information cannot be reported at the state level.
- The following states have state-level race or ethnicity data presentation restrictions:
  - Data for American Indians and Alaska Natives cannot be displayed for Delaware, Illinois, Kansas, Kentucky, New Jersey, and New York.
  - Data for Asians and Pacific Islanders cannot be displayed for Delaware, Kansas and Kentucky.
  - Hispanic ethnicity data cannot be displayed for Delaware, Kentucky, and Massachusetts.
  - Race and ethnicity combinations—white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for Delaware, Kansas, Kentucky, Massachusetts, and Pennsylvania.
- For more information, please refer to the *Race Recode for USCS*, *Origin recode NHIA (Hispanic, Non-Hisp)*, and *NHIA derived Hisp origin* variable descriptions in this document.

Values	Frequency	Percentage
All races	24,697,337	26.1%
White	21,036,166	22.3%
White Non-Hispanic	17,255,267	18.3%
White Hispanic	1,485,051	1.6%
Black	2,532,714	2.7%
Black Non-Hispanic	2,312,131	2.4%
Black Hispanic	43,024	0.0%
American Indian/Alaska Native	116,279	0.1%
AIAN Non-Hispanic	113,711	0.1%
AIAN Hispanic	5,276	0.0%
Asian/Pacific Islander	660,222	0.7%
API Non-Hispanic	626,386	0.7%
API Hispanic	9,928	0.0%
Hispanic	1,618,689	1.7%
Non-Hispanic	21,998,837	23.3%

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
PRCSDA	Purchased/Referred Care Service Delivery Area
SEER	Surveillance, Epidemiology, and End Results
SS	Summary Stage
USCS	U.S. Cancer Statistics
WHO	World Health Organization

Appendix A. NPCR – Indian Health Services (IHS) Linkage Schedule

All NPCR-funded registries link with the Indian Health Service every five years. The most recent linkage year was 2016.

All state central cancer registries with Purchase/Referred Care Service Delivery Area[36] (PRCSDA) counties link with the Indian Health Service every year. These include:

- Alabama
- Alaska
- Arizona
- California
- Colorado
- Florida
- Idaho
- Indiana
- Kansas
- Louisiana
- Maine
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Montana
- Nebraska
- Nevada
- New York
- North Carolina
- North Dakota
- Oklahoma
- Oregon
- Pennsylvania
- Rhode Island
- South Carolina
- South Dakota
- Texas
- Washington
- Wisconsin
- Wyoming

[36] Purchased/Referred Care Service Delivery Area (PRCSDA) was previously referred to as Contract Health Service Delivery Area (CHSDA).