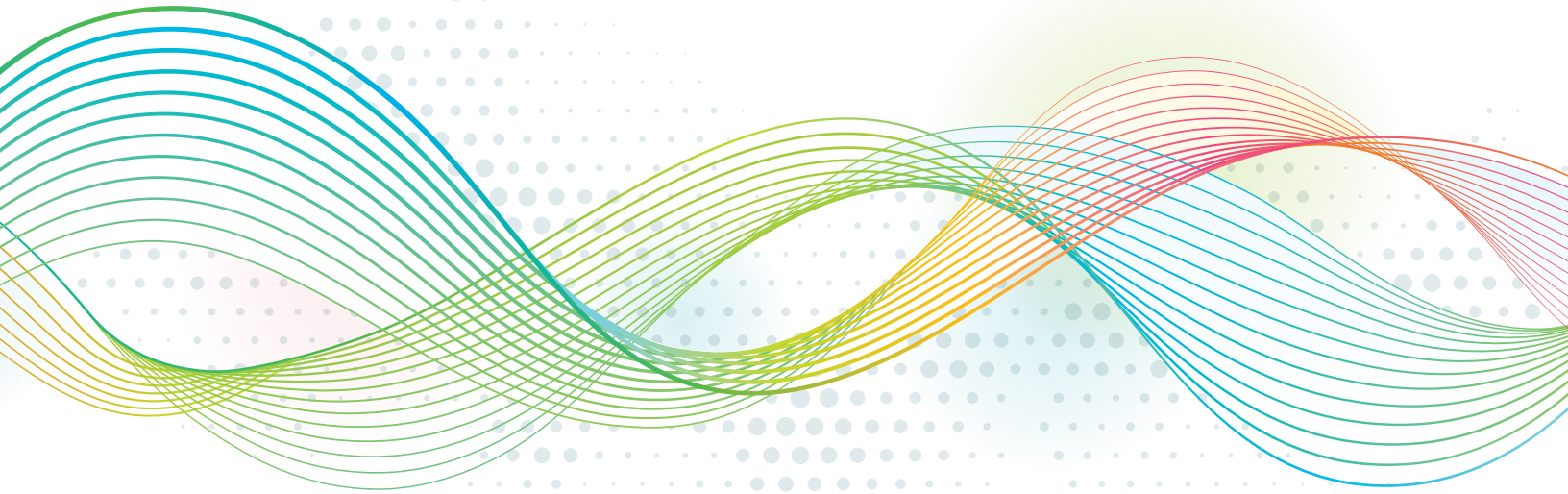
National Program of Cancer Registries (NPCR) and Surveillance, Epidemiology & End Results (SEER)

NPCR and SEER Incidence – USCS 2001–2014 Public Use Database Data Standards and Data Dictionary



November, 2016 Submission

Diagnosis Years 2001–2014

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

NIH NATIONAL CANCER INSTITUTE

CS273827-C

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 8

Data Citations 9

Cautionary Notes 9

Checklist for a NPCR and SEER Incidence – USCS 2001– 2014 Public Use Data Analysis 12

NPCR and SEER Incidence – USCS 2001– 2014 Public Use Database: Variable Definition and Frequency 13

Additional Resources 60

Appendix A. NPCR – Indian Health Services (IHS) Linkage Schedule 62

Message to Data Users

August 9, 2017

We are pleased to share for the first time a combined 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. We anticipate it will be a valuable resource for the research community.

This free, publicly available data source is the result of tremendous efforts undertaken every day 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 action for cancer prevention and control.

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. Although our database includes tens of millions of cases of cancer collected over 14 years, we know that each case represents an individual with cancer and those who care for that individual.

We encourage researchers to use our data to inform scientific inquiries, programs, and policies. Through use of this combined NPCR and SEER 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

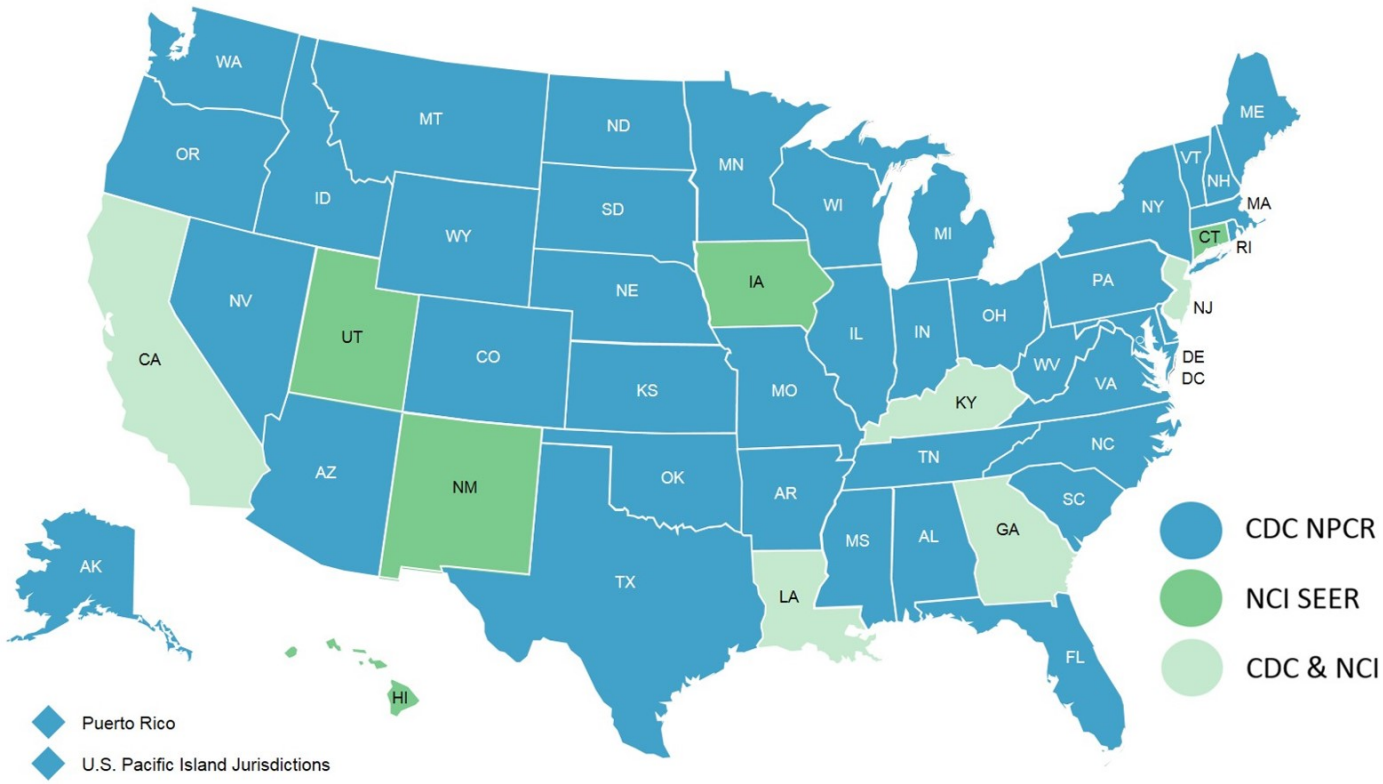
The National Program of Cancer Registries (NPCR), administered by the Centers for Disease Control and Prevention (CDC), was established by Congress in 1992. Through cooperative agreements, NPCR supports central cancer registries in 45 states, the District of Columbia, Puerto Rico, and the U.S. Pacific Island Jurisdictions (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 United States Cancer Statistics (USCS).

The cancer registries funded by CDC and NCI routinely collect data on patient demographics, primary tumor site, tumor morphology, stage at diagnosis, first course of treatment, and outcomes. Medical facilities such as hospitals, doctor's offices, pathology laboratories, and other treatment centers send demographic and clinical information related to people with cancer to a central cancer registry, where the information is consolidated and goes through rigorous checks for quality and completeness. All hospitals are required by state law to report cancer cases to the central cancer registry in their respective states. On an annual basis, the central cancer registries submit demographic and clinical information about each cancer case to CDC and/or NCI. **None of the information submitted to CDC or NCI contains personally identifiable information about individual patients.**

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

Figure 1. Central cancer registry programs funded by NPCR and SEER during 2016 data collection



Data Collection, Submission, Quality, and Coverage

The most current data come from the November 2016 NPCR and SEER submissions, which include cancer cases diagnosed in 2001 through 2014. 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 2014 data, NPCR required submission by November 30, 2016 and SEER required submission by November 1, 2016).

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 USCS public research data file, they must have met the following quality and completeness criteria for publication—

- 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 – USCS 2001–2014 Public Use Database

Two NPCR and SEER Incidence – USCS public use databases are available for researchers: the 2001–2014 database and the 2005–2014 database. This data standards document is specific to the 2001–2014 database.

The 2001–2014 database includes race and ethnicity variables, while the 2005–2014 database does not. The 2005–2014 database includes Puerto Rico data, while the 2001–2014 database does not.

- The 2001–2014 database's population denominators are race-specific, ethnicity-specific, and sex-specific county population estimates from the U.S. Census (July 1, 2010–2015 bridged–race vintage 2015 population estimates), modified by SEER and aggregated to the state and national levels.

- The 2005–2014 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 – USCS 2001–2014 public research data.

For more detail on data availability by central cancer registry from 2001–2014, see Table 2 below.

Table 1. U.S. population coverage, NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database.

Diagnosis year(s) [a]	Percentage of U.S. population covered in database
2001	99.0%
2002	98.8%
2003	100.0%
2004	100.0%
2005	100.0%
2006	100.0%
2007	100.0%
2008	100.0%
2009	100.0%
2010	100.0%
2011	99.1%
2012	100.0%
2013	100.0%
2014	100.0%
2001–2014	98.0%
2005–2014[b]	99.1%
2010–2014[c]	99.1%

[a]For the calculated percent population coverage for a range of years not shown in Table 1 (for example, 2008–2013), please send a request to: uscsdata@cdc.gov.

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

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

Table 2. Central cancer registry data included in the NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database[a]

Registry	Year of Diagnosis													
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Alabama														
Alaska														
Arizona														
Arkansas														
California														
Colorado														
Connecticut														
Delaware														
District of Columbia		NA												
Florida														
Georgia														
Hawaii														
Idaho														
Illinois														
Indiana														
Iowa														
Kansas														
Kentucky														
Louisiana														
Maine														
Maryland														
Massachusetts														
Michigan														
Minnesota														
Mississippi	NS	NA												
Missouri														
Montana														
Nebraska														
Nevada											NA			
New Hampshire														
New Jersey														
New Mexico														
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														
Wyoming														

Shaded box: Data meet the *United States Cancer Statistics (USCS)* publication criteria and are available in the NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database.

NS: Data were not submitted and are not available in the 2001–2014 public use research database.

NA: Data did not meet USCS quality and completeness criteria for publication and are not available in the 2001–2014 public use research database.

[a] Puerto Rico and U.S. Pacific Island Jurisdiction data are not included in the NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database.

Variable List

Table 3 shows all of the variables available in the 2001-2014 public use research database.

Table 3. Variables in the NPCR and SEER Incidence – USCS 2001–2014 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	Sex
	Year of diagnosis
	Addr at DX – state
	USCS standard
	USCS0114
	USCS0514
	Program
	Race recode for USCS
	USCS1014
	Region
	Stateraceethincl
	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

Abbreviations used in 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	United States 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:** 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 2001–2014 database:** National Program of Cancer Registries and Surveillance, Epidemiology, and End Results SEER*Stat Database: NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database, United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. Released August 2017, based on November 2016 submissions. Available at www.cdc.gov/cancer/public-use.

Cautionary Notes

Before using the database, analysts should read and understand the following nuances of the NPCR and SEER Incidence – USCS 2001–2014 Public Use Research data. If you have questions regarding these notes, please contact CDC at uscsdata@cdc.gov.

Exclusions

Cancer cases that were identified only through death certificate or autopsy reports have been excluded from this database.

Suppression Rules[1-2]

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

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

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

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

Note: As a further mechanism to protect data confidentiality, 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[3]

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[4-7]

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[8]

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 2004–2014.

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 2014, 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 2014 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. 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.
2. Doyle P, Lane JI, Theeuwes JM, Zayatz LM. *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies.* Amsterdam: Elsevier Science; 2001.
3. 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.
4. *International Classification of Diseases for Oncology.* Third Edition, First Revision. Geneva: World Health Organization, 2013.

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

6. 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/.

7. 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.

8. 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 – USCS 2001– 2014 Public Use Data Analysis

Multi-year analyses: USCS0114, USCS0514, or USCS1014

- ☐ The variable "USCS1014" includes states meeting USCS publication criteria for diagnosis years 2010–2014, "USCS0514" for diagnosis years 2005–2014, and "USCS0114" for diagnosis years 2001–2014. If you are reporting state-specific ethnicity or race/ethnicity combinations, have you suppressed data from the registries that opted out of reporting these data items using the "Stateraceethincl" variable or manually in the SEER*Stat Selection tab?[1]

- ☐ If you are conducting a multiyear analysis and want to restrict it to the states that met reporting standards during each of the years, did you use variable "USCS0114", "USCS0514", or "USCS1014" 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.

  - The "Year of Diagnosis" variable is used in combination with the predefined USCS variable to exclude the nonrelevant years. For example, if "USCS1014" is used, then "Year of Diagnosis" should also be restricted to diagnosis years 2010–2014 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.[2]

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? [3]

Common selection and reporting considerations

- ☐ 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 – USCS 2001–2014 Public Use Research Database?[7]

---

[1] See "Stateraceethincl," "Race Recode," and "Origin recode NHIA" variable descriptions.

[2] See "USCS0114", "USCS0514", and "USCS1014" variable descriptions.

[3] See "USCS standard" variable description.

[4] See Cautionary Notes/Primary Site Variables section.

[5] See "Diagnostic Confirmation" variable descriptions.

[6] See "Sex" variable description.

[7] See Data Citation section.

NPCR and SEER Incidence – USCS 2001– 2014 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	15,651	0.1%
01–04 years	51,600	0.2%
05–09 years	38,112	0.2%
10–14 years	44,883	0.2%
15–19 years	75,950	0.3%
20–24 years	121,830	0.5%
25–29 years	190,892	0.8%
30–34 years	294,702	1.3%
35–39 years	457,725	2.0%
40–44 years	803,609	3.5%
45–49 years	1,292,587	5.7%
50–54 years	1,918,386	8.4%
55–59 years	2,454,505	10.8%
60–64 years	2,836,300	12.5%
65–69 years	3,077,623	13.5%
70–74 years	2,895,838	12.7%
75–79 years	2,609,463	11.5%
80–84 years	1,984,725	8.7%
≥85 years	1,613,874	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	11,306,415	49.6%
Female	11,471,840	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	1,438,190	6.3%
2002	1,454,889	6.4%
2003	1,470,727	6.5%
2004	1,532,081	6.7%
2005	1,567,061	6.9%
2006	1,606,382	7.1%
2007	1,659,749	7.3%
2008	1,683,573	7.4%
2009	1,707,331	7.5%
2010	1,697,860	7.5%
2011	1,722,826	7.6%
2012	1,729,273	7.6%
2013	1,753,861	7.7%
2014	1,754,452	7.7%

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	38,693	0.2%
Alabama	359,242	1.6%
Arkansas	217,713	1.0%
Arizona	401,521	1.8%
California	2,370,311	10.4%
Colorado	309,981	1.4%
Connecticut	315,405	1.4%
District of Columbia	40,445	0.2%
Delaware	77,677	0.3%
Florida	1,615,245	7.1%
Georgia	629,340	2.8%
Hawaii	98,115	0.4%
Idaho	103,817	0.5%
Illinois	966,484	4.2%
Indiana	476,749	2.1%
Iowa	254,936	1.1%
Kansas	212,933	0.9%
Kentucky	368,568	1.6%
Louisiana	341,876	1.5%
Massachusetts	552,796	2.4%
Maryland	420,024	1.8%
Maine	125,800	0.6%
Michigan	811,949	3.6%
Minnesota	393,979	1.7%
Missouri	456,026	2.0%
Mississippi	188,510	0.8%
Montana	80,117	0.4%
North Carolina	698,947	3.1%
North Dakota	51,704	0.2%
Nebraska	136,493	0.6%
New Hampshire	114,585	0.5%

Values	Frequency	Percentage
New Jersey	750,463	3.3%
Nevada	127,902	0.6%
New Mexico	152,435	0.7%
New York	1,612,194	7.1%
Ohio	901,712	4.0%
Oklahoma	271,168	1.2%
Oregon	295,340	1.3%
Pennsylvania	1,152,457	5.1%
Rhode Island	93,430	0.4%
South Carolina	350,985	1.5%
South Dakota	62,654	0.3%
Tennessee	461,873	2.0%
Texas	1,437,405	6.3%
Utah	137,640	0.6%
Virginia	533,122	2.3%
Vermont	54,465	0.2%
Washington	507,651	2.2%
Wisconsin	441,207	1.9%
West Virginia	166,464	0.7%
Wyoming	37,707	0.2%

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

Source of Standard: NPCR
Source Item Name: Derived from NAACCR's "State at diagnosis (CoC)" and USCS 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 USCS 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 USCS 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, USCS1014 (includes diagnosis years 2010–2014), USCS0515 (includes diagnosis years 2005–2014), or USCS0114 (includes diagnosis years 2001–2014).
- 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 USCS 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	1,438,190	6.3%
49	2002	1,454,889	6.4%
51	2003	1,470,727	6.5%
51	2004	1,532,081	6.7%
51	2005	1,567,061	6.9%
51	2006	1,606,382	7.1%
51	2007	1,659,749	7.3%
51	2008	1,683,573	7.4%
51	2009	1,707,331	7.5%
51	2010	1,697,860	7.5%
50	2011	1,722,826	7.6%
51	2012	1,729,273	7.6%
51	2013	1,753,861	7.7%
51	2014	1,754,452	7.7%

[a] Refer to Table 2 for the list of central cancer registries included in each diagnosis year.

SEER*Stat Item Name: **USCS0114**

Source of Standard: NPCR
Source Item Name: Derived from NAACCR's "State at diagnosis (CoC)" and USCS publication standards
Source Item Number: Derived from NAACCR's 80

Description

This variable indicates whether a central cancer registry met the USCS publication standard for all cancer sites combined each year in 2001–2014.

Considerations for use

- This variable is used for analysis of combined 2001–2014 data in the NPCR and SEER Incidence – USCS 2001–2014 database. Data from states that did not meet the USCS publication standard in any given year were excluded from the database for that year. Please refer to Table 2 for the list of states by year that were included and excluded by this variable.
- 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, USCS1014 (includes diagnosis years 2010–2014), USCS0515 (includes diagnosis years 2005–2014), or USCS0114 (includes diagnosis years 2001–2014).
- 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 USCS publication standard is available at www.cdc.gov/cancer/npcr/uscs/technical_notes/criteria.htm.

Values	Frequency	Percentage
Does not meet USCS standard from 2001–2014	381,390	1.7%
Meets USCS standard from 2001–2014	22,396,865	98.3%

SEER*Stat Item Name: **USCS0514**

Source of Standard: NPCR
Source Item Name: Derived from NAACCR's "State at diagnosis (CoC)" and USCS publication standard
Source Item Number: Derived from NAACCR's 80

Description

This variable indicates whether a central cancer registry met the USCS publication standard for all cancer sites combined each year in 2005–2014. When using this variable, restrict the diagnosis years to 2005–2014. 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 2005–2014 data in the NPCR and SEER Incidence – USCS 2001–2014 database. Data from states that did not meet the USCS publication standard in any given year were excluded from the database for that year. Please refer to Table 2 for the list of states by year that were included and excluded by this variable.
- 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, USCS1014 (includes diagnosis years 2010–2014), USCS0515 (includes diagnosis years 2005–2014), or USCS0114 (includes diagnosis years 2001–2014).
- 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 USCS 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 2005–2014.

Values	Frequency	Percentage
Does not meet USCS standard from 2005–2014	109,505	0.6%
Meets USCS standard from 2005–2014	16,772,863	99.4%

SEER*Stat Item Name: **USCS1014**

Source of Standard: NPCR
Source Item Name: Derived from NAACCR's "State at diagnosis (CoC)" and USCS publication standard
Source Item Number: Derived from NAACCR's 80

Description

This variable indicates whether a central cancer registry met the USCS publication standard for all cancer sites combined each year in 2010–2014. When using this variable, restrict the diagnosis years to 2010–2014. 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 2010–2014 data in the NPCR and SEER Incidence – USCS 2001–2014 database. Data from states that did not meet the USCS publication standard in any given year were excluded from the database for that year. Please refer to Table 2 for the list of states by year that were included and excluded by this variable.
- 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, USCS1014 (includes diagnosis years 2010–2014), USCS0515 (includes diagnosis years 2005–2014), or USCS0114 (includes diagnosis years 2001–2014).
- 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 USCS 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 2010–2014.

Values	Frequency	Percentage
Does not meet USCS standard from 2010–2014	49,477	0.6%
Meets USCS standard from 2010–2014	8,608,795	99.4%

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

States that received funding from both programs (California, Georgia, Kentucky, Louisiana, and New Jersey) are categorized as "NPCR" states. "SEER" refers to states receiving funding only from SEER (Connecticut, Hawaii, Iowa, New Mexico, and Utah).

Values	Frequency	Percentage
NPCR	21,844,257	95.9%
SEER	933,998	4.1%

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 USCS 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 USCS 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	4,771,595	20.9%
Midwest	5,166,826	22.7%
South	8,178,604	35.9%
West	4,661,230	20.5%

SEER*Stat Item Name: **Stateraceethincl**

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 and ethnicity.

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 Arkansas, Delaware, and Kentucky.
  - Hispanic ethnicity data cannot be displayed for Arkansas, Delaware, Kentucky, Massachusetts, North Dakota, and Virginia.
  - Race and ethnicity combinations—white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for Pennsylvania.
- For more information, please refer to the “Race Recode for USCS” and “Origin recode NHIA (Hispanic, Non-Hisp)” variable descriptions in this document.

Values	Frequency	Percentage
Exclude state for race/ethnicity state-level analyses	6,083,099	26.7%
Include state for race/ethnicity state-level analyses	16,695,156	73.3%

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 2001–2014 public use database; it is not available in the 2005–2014 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 Arkansas, Delaware, and Kentucky.
- 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 Contract Health Service Delivery Area (CHSDA) 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	19,413,986	85.2%
Black	2,323,399	10.2%
American Indian/Alaska Native	110,418	0.5%
Asian or Pacific Islander	600,439	2.6%
Other unspecified (1991+)	74,252	0.3%
Unknown	255,761	1.1%

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 2001–2013 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, Massachusetts, North Dakota, and Virginia.
  - Data for race and ethnicity combinations—white Hispanic, white non-Hispanic, black Hispanic, and black non-Hispanic—cannot be displayed for 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	21,272,041	93.4%
Spanish-Hispanic-Latino	1,502,920	6.6%
Invalid Value(s)	3,294	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)	22,396,563	98.3%
C80.9 (Unknown primary site)	381,692	1.7%

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	22,778,255	100%

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

Description

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

Considerations for use

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

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

  The percentage of cases with a known grade varies by primary cancer site. Rules for coding the tumor grade differ for some primary sites. As a result, it may be appropriate to have a tumor grade coded as “9 – unknown.”

Values	Frequency	Percentage
Well differentiated; Grade I	1,979,040	8.7%
Moderately differentiated; Grade II	5,673,313	24.9%
Poorly differentiated; Grade III	4,640,325	20.4%
Undifferentiated; anaplastic; Grade IV	726,857	3.2%
T-cell	76,810	0.3%
B-cell; pre-B; B-precursor	1,051,916	4.6%
Null cell; non T-non B	2,479	0.0%
NK cell; natural killer cell (1995+)	3,421	0.0%
Unknown	8,624,094	37.9%

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)	21,446,240	94.2%
Positive histology	20,583,803	90.4%
Positive exfoliative cytology, no positive histology	717,953	3.2%
Positive histology AND immunophenotyping AND/OR positive genetic studies	119,475	0.5%
Positive microscopic confirm, method not specified	25,009	0.1%
Positive laboratory test/marker study	98,415	0.4%
Direct visualization without microscopic confirmation	31,792	0.1%
Radiography without microscopic confirm	784,201	3.4%
Clinical diagnosis only	132,806	0.6%
Unknown	284,800	1.3%

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

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

Description

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

Considerations for use

- 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	531,501	2.3%
801–804	Epithelial Neoplasms, NOS	1,440,003	6.3%
805–808	Squamous Cell Neoplasms	1,742,401	7.6%
809–811	Basal Cell Neoplasms	6,521	0.0%
812–813	Transitional Cell Papillomas and Carcinomas	992,103	4.4%
814–838	Adneomas and Adenocarcinomas	8,979,569	39.4%
839–842	Adnexal and Skin Appendage Neoplasms	25,691	0.1%
843	Mucoepidermoid Neoplasms	19,828	0.1%
844–849	Cystic, Mucinous and Serous Neoplasms	622,379	2.7%
850–854	Ductal and Lobular Neoplasms	3,408,369	15.0%
855	Acinar Cell Neoplasms	49,866	0.2%
856–857	Complex Epithelial Neoplasms	76,825	0.3%
858	Thymic Epithelial Neoplasms	10,230	0.0%
859–867	specialized Gonadal Neoplasms	6,270	0.0%
868–871	Paragangliomas and Glomus Tumors	3,645	0.0%
872–879	Nevi and Melanomas	1,464,296	6.4%
880	Soft Tissue Tumors and Sarcomas, NOS	46,300	0.2%
881–883	Fibromatous Neoplasms	50,963	0.2%
884	Myxomatous Neoplasms	1,098	0.0%
885–888	Lipomatous Neoplasms	33,166	0.1%
889–892	Myomatous Neoplasms	53,511	0.2%
893–899	Complex Mixed and Stromal Neoplasms	108,848	0.5%
900–903	Fibroepithelial Neoplasms	6,816	0.0%
904	Synovial-Like Neoplasms	8,165	0.0%
905	Mesothelial Neoplasms	43,870	0.2%
906–909	Germ Cell Neoplasms	122,128	0.5%
910	Trophoblastic Neoplasms	5,284	0.0%
911	Mesonephromas	209	0.0%

ICD–O–3 Code	Label	Frequency	Percentage
912–916	Blood Vessel Tumors	49,806	0.2%
917	Lymphatic Vessel Tumors	182	0.0%
918–924	Osseous and Chondromatous Neoplasms	27,932	0.1%
925	Giant Cell Tumors	1,172	0.0%
926	Miscellaneous Bone Tumors	7,172	0.0%
927–934	Odotogenic Tumors	757	0.0%
935–937	Miscellaneous Tumors	13,550	0.1%
938–948	Gliomas	275,318	1.2%
949–952	Neuroepitheliomatous Neoplasms	22,923	0.1%
953	Meningiomas	284,173	1.2%
954–957	Nerve Sheath Tumors	73,862	0.3%
958	Granular Cell Tumors and Alveolar Soft Part Sarcomas	1,063	0.0%
959–972	Hodgkin and Non-Hodgkin Lymphomas	951,648	4.2%
973	Plasma Cell Tumors	272,721	1.2%
974	Mast Cell Tumors	1,924	0.0%
975	Neoplasms of Histiocytes and Accessory Lymphoid Cells	5,007	0.0%
976	Immunoproliferative Disease	15,541	0.1%
980–994	Leukemias	590,120	2.6%
995–996	Chronic Myeloproliferative Disorders	126,853	0.6%
997	Other Hematologic Disorders	9,797	0.0%
998–999	Myelodysplastic Syndromes	186,876	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)	**22,778,255**	100.0%
Oral Cavity and Pharynx	517,977	2.3%
Lip	32,168	0.1%
Tongue	149,606	0.7%
Salivary Gland	55,137	0.2%
Floor of Mouth	31,098	0.1%
Gum and Other Mouth	72,426	0.3%
Nasopharynx	24,246	0.1%
Tonsil	86,501	0.4%
Oropharynx	22,333	0.1%
Hypopharynx	32,353	0.1%
Other Oral Cavity and Pharynx	12,109	0.1%
Digestive System	3,846,155	16.9%
Esophagus	218,998	1.0%
Stomach	302,995	1.3%
Small Intestine	96,137	0.4%
Colon and Rectum	2,136,980	9.4%
Colon excluding Rectum	1,539,170	6.8%
Cecum	332,524	1.5%
Appendix	35,418	0.2%
Ascending Colon	295,097	1.3%
Hepatic Flexure	75,261	0.3%
Transverse Colon	139,479	0.6%
Splenic Flexure	47,584	0.2%
Descending Colon	91,028	0.4%
Sigmoid Colon	421,363	1.8%
Large Intestine, NOS	101,416	0.4%
Rectum and Rectosigmoid Junction	597,810	2.6%
Rectosigmoid Junction	157,242	0.7%
Rectum	440,568	1.9%
Anus, Anal Canal and Anorectum	87,987	0.4%
Liver and Intrahepatic Bile Duct	296,673	1.3%
Liver	265,300	1.2%

Values	Frequency	Percentage
Intrahepatic Bile Duct	31,373	0.1%
Gallbladder	51,675	0.2%
Other Biliary	75,946	0.3%
Pancreas	517,159	2.3%
Retroperitoneum	16,814	0.1%
Peritoneum, Omentum and Mesentery	26,305	0.1%
Other Digestive Organs	18,486	0.1%
Respiratory System	3,083,557	13.5%
Nose, Nasal Cavity and Middle Ear	31,071	0.1%
Larynx	185,341	0.8%
Lung and Bronchus	2,857,617	12.5%
Pleurae	1,318	0.0%
Trachea, Mediastinum and Other Respiratory Organs	8,210	0.0%
Bones and Joints	40,238	0.2%
Soft Tissue including Heart	139,904	0.6%
Skin excluding Basal and Squamous	1,494,418	6.6%
Melanoma of the Skin	1,423,127	6.2%
Other Non-Epithelial Skin	71,291	0.3%
Breast (female and male combined)	3,704,811	16.3%
Female Genital System	1,219,141	5.4%
Cervix Uteri	175,231	0.8%
Corpus and Uterus, NOS	614,089	2.7%
Corpus Uteri	596,496	2.6%
Uterus, NOS	17,593	0.1%
Ovary	296,245	1.3%
Vagina	20,526	0.1%
Vulva	90,171	0.4%
Other Female Genital Organs	22,879	0.1%
Male Genital System	3,028,472	13.3%
Prostate	2,884,909	12.7%
Testis	112,965	0.5%
Penis	25,400	0.1%
Other Male Genital Organs	5,198	0.0%
Urinary System	1,689,325	7.4%
Urinary Bladder	937,412	4.1%
Kidney and Renal Pelvis	698,278	3.1%
Ureter	37,185	0.2%
Other Urinary Organs	16,450	0.1%
Eye and Orbit	42,119	0.2%
Brain and Other Nervous System	680,249	3.0%
Brain	309,114	1.4%
Cranial Nerves Other Nervous System	371,135	1.6%
Endocrine System	674,912	3.0%
Thyroid	514,268	2.3%
Other Endocrine including Thymus	160,644	0.7%
Lymphoma	970,703	4.3%
Hodgkin Lymphoma	117,836	0.5%
Hodgkin – Nodal	114,376	0.5%
Hodgkin – Extranodal	3,460	0.0%

Values	Frequency	Percentage
Non-Hodgkin Lymphoma	852,867	3.7%
NHL – Nodal	581,190	2.6%
NHL – Extranodal	271,677	1.2%
Myeloma	270,452	1.2%
Leukemia	575,369	2.5%
Lymphocytic Leukemia	287,112	1.3%
Acute Lymphocytic Leukemia	64,643	0.3%
Chronic Lymphocytic Leukemia	203,674	0.9%
Other Lymphocytic Leukemia	18,795	0.1%
Myeloid and Monocytic Leukemia	258,864	1.1%
Acute Myeloid Leukemia	167,720	0.7%
Acute Monocytic Leukemia	10,620	0.0%
Chronic Myeloid Leukemia	72,739	0.3%
Other Myeloid/Monocytic Leukemia	7,785	0.0%
Other Leukemia	29,393	0.1%
Other Acute Leukemia	11,244	0.0%
Aleukemic, Subleukemic and NOS	18,149	0.1%
Mesothelioma	43,870	0.2%
Kaposi Sarcomae	17,920	0.1%
Miscellaneous	738,663	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	53,294	23.6%
I(a) Lymphoid leukemias	38,060	16.8%
I(b) Acute myeloid leukemias	9,264	4.1%
I(c) Chronic myeloproliferative diseases	2,671	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,739	0.8%
I(e) Unspecified and other specified leukemias	1,560	0.7%
II Lymphomas and reticuloendothelial neoplasms	30,418	13.4%
II(a) Hodgkin lymphomas	14,332	6.3%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	10,221	4.5%
II(c) Burkitt lymphoma	2,812	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	2,669	1.2%
II(e) Unspecified lymphomas	384	0.2%
III CNS and misc intracranial and intraspinal neoplasms	50,519	22.3%
III(a) Ependymomas and choroid plexus tumor	4,145	1.8%
III(b) Astrocytomas	19,083	8.4%
III(c) Intracranial and intraspinal embryonal tumors	7,069	3.1%
III(d) Other gliomas	6,339	2.8%
III(e) Other specified intracranial/intraspinal neoplasms	12,128	5.4%
III(f) Unspecified intracranial and intraspinal neoplasms	1,755	0.8%
IV Neuroblastoma and other peripheral nervous cell tumors	9,870	4.4%
IV(a) Neuroblastoma and ganglioneuroblastoma	9,588	4.2%
IV(b) Other peripheral nervous cell tumors	282	0.1%
V Retinoblastoma	3,659	1.6%
VI Renal tumors	7,911	3.5%
VI(a) Nephroblastoma and other nonepithelial renal tumors	7,139	3.2%
VI(b) Renal carcinomas	744	0.3%
VI(c) Unspecified malignant renal tumors	28	0.0%
VII Hepatic tumors	2,613	1.2%
VII(a) Hepatoblastoma	1,908	0.8%
VII(b) Hepatic carcinomas	681	0.3%
VII(c) Unspecified malignant hepatic tumors	24	0.0%

Values	Frequency	Percentage
VIII Malignant bone tumors	10,362	4.6%
VIII(a) Osteosarcomas	5,845	2.6%
VIII(b) Chondrosarcomas	400	0.2%
VIII(c) Ewing tumor and related sarcomas of bone	3,399	1.5%
VIII(d) Other specified malignant bone tumors	513	0.2%
VIII(e) Unspecified malignant bone tumors	205	0.1%
IX Soft tissue and other extraosseous sarcomas	13,970	6.2%
IX(a) Rhabdomyosarcomas	5,454	2.4%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	1,519	0.7%
IX(c) Kaposi sarcomae	55	0.0%
IX(d) Other specified soft tissue sarcomas	5,448	2.4%
IX(e) Unspecified soft tissue sarcomas	1,494	0.7%
X Germ cell & trophoblastic tumors & neoplasms of gonads	13,349	5.9%
X(a) Intracranial & intraspinal germ cell tumors	2,342	1.0%
X(b) Extracranial & extragonadal germ cell tumors	1,689	0.7%
X(c) Malignant gonadal germ cell tumors	8,447	3.7%
X(d) Gonadal carcinomas	493	0.2%
X(e) Other and unspecified malignant gonadal tumors	378	0.2%
XI Other malignant epithelial neoplasms and melanomas	21,375	9.4%
XI(a) Adrenocortical carcinomas	253	0.1%
XI(b) Thyroid carcinomas	9,104	4.0%
XI(c) Nasopharyngeal carcinomas	660	0.3%
XI(d) Malignant melanomas	6,043	2.7%
XI(e) Skin carcinomas	101	0.0%
XI(f) Other and unspecified carcinomas	5,214	2.3%
XII Other and unspecified malignant neoplasms	881	0.4%
XII(a) Other specified malignant tumors	455	0.2%
XII(b) Other unspecified malignant tumors	426	0.2%
Not classified by ICCC or *in situ*	7,975	3.5%

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	53,294	23.6%
I(a) Lymphoid leukemias	38,060	16.8%
I(a.1) Precursor cell leukemias	36,914	16.3%
I(a.2) Mature B-cell leukemias	867	0.4%
I(a.3) Mature T-cell and NK cell leukemias	119	0.1%
I(a.4) Lymphoid leukemia, NOS	160	0.1%
I(b) Acute myeloid leukemias	9,264	4.1%
I(c) Chronic myeloproliferative diseases	2,671	1.2%
I(d) Myelodysplastic syndrome and other myeloproliferative	1,739	0.8%
I(e) Unspecified and other specified leukemias	1,560	0.7%
II Lymphomas and reticuloendothelial neoplasms	30,418	13.4%
II(a) Hodgkin lymphomas	14,332	6.3%
II(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)	10,221	4.5%
II(b.1) Precursor cell lymphomas	2,952	1.3%
II(b.2) Mature B-cell lymphomas except Burkitt lymphoma	4,117	1.8%
II(b.3) Mature T-cell and NK-cell lymphomas	2,590	1.1%
II(b.4) Non-Hodgkin lymphomas, NOS	562	0.2%
II(c) Burkitt lymphoma	2,812	1.2%
II(d) Miscellaneous lymphoreticular neoplasms	2,669	1.2%
II(e) Unspecified lymphomas	384	0.2%
III CNS and misc intracranial and intraspinal neoplasms	50,519	22.3%
III(a) Ependymomas and choroid plexus tumor	4,145	1.8%
III(a.1) Ependymomas	3,195	1.4%
III(a.2) Choroid plexus tumor	950	0.4%
III(b) Astrocytomas	19,083	8.4%
III(c) Intracranial and intraspinal embryonal tumors	7,069	3.1%
III(c.1) Medulloblastomas	4,518	2.0%
III(c.2) PNET	1,587	0.7%
III(c.3) Medulloepithelioma	62	0.0%
III(c.4) Atypical teratoid/rhabdoid tumor	902	0.4%

Values	Frequency	Percentage
III(d) Other gliomas	6,339	2.8%
III(d.1) Oligodendrogliomas	823	0.4%
III(d.2) Mixed and unspecified gliomas	5,377	2.4%
III(d.3) Neuroepithelial glial tumors of uncertain orig	139	0.1%
III(e) Other specified intracranial/intraspinal neoplasms	12,128	5.4%
III(e.1) Pituitary adenomas and carcinomas	4,909	2.2%
III(e.2) Tumors of sellar region (craniopharyngiomas)	1,837	0.8%
III(e.3) Pineal parenchymal tumors	513	0.2%
III(e.4) Neuronal and mixed neuronal-glial tumors	3,513	1.6%
III(e.5) Meningiomas	1,356	0.6%
III(f) Unspecified intracranial and intraspinal neoplasms	1,755	0.8%
IV Neuroblastoma and other peripheral nervous cell tumors	9,870	4.4%
IV(a) Neuroblastoma and ganglioneuroblastoma	9,588	4.2%
IV(b) Other peripheral nervous cell tumors	282	0.1%
V Retinoblastoma	3,659	1.6%
VI Renal tumors	^1	^1
VI(a) Nephroblastoma and other nonepithelial renal tumors	^1	^1
VI(a.1) Nephroblastoma	6,717	3.0%
VI(a.2) Rhabdoid renal tumor	181	0.1%
VI(a.3) Kidney sarcomas	229	0.1%
VI(a.4) pPNET of kidney	^2	^2
VI(b) Renal carcinomas	744	0.3%
VI(c) Unspecified malignant renal tumors	28	0.0%
VII Hepatic tumors	2,613	1.2%
VII(a) Hepatoblastoma	1,908	0.8%
VII(b) Hepatic carcinomas	681	0.3%
VII(c) Unspecified malignant hepatic tumors	24	0.0%
VIII Malignant bone tumors	10,362	4.6%
VIII(a) Osteosarcomas	5,845	2.6%
VIII(b) Chondrosarcomas	400	0.2%
VIII(c) Ewing tumor and related sarcomas of bone	3,399	1.5%
VIII(c.1) Ewing tumor and Askin tumor of bone	3,268	1.4%
VIII(c.2) pPNET of bone	131	0.1%
VIII(d) Other specified malignant bone tumors	513	0.2%
VIII(d.1) Malignant fibrous neoplasms of bone	49	0.0%
VIII(d.2) Malignant chordomas	238	0.1%
VIII(d.3) Odontogenic malignant tumors	65	0.0%
VIII(d.4) Miscellaneous malignant bone tumors	161	0.1%
VIII(e) Unspecified malignant bone tumors	205	0.1%
IX Soft tissue and other extraosseous sarcomas	13,970	6.2%
IX(a) Rhabdomyosarcomas	5,454	2.4%
IX(b) Fibrosarcomas, peripheral nerve & other fibrous	1,519	0.7%
IX(b.1) Fibroblastic and myofibroblastic tumors	796	0.4%
IX(b.2) Nerve sheath tumors	702	0.3%
IX(b.3) Other fibromatous neoplasms	21	0.0%
IX(c) Kaposi sarcomae	55	0.0%
IX(d) Other specified soft tissue sarcomas	5,448	2.4%
IX(d.1) Ewing tumor and Askin tumor of soft tissue	699	0.3%
IX(d.2) pPNET of soft tissue	336	0.1%
IX(d.3) Extrarenal rhabdoid tumor	282	0.1%
IX(d.4) Liposarcomas	313	0.1%

Values	Frequency	Percentage
IX(d.5) Fibrohistiocytic tumors	1,363	0.6%
IX(d.6) Leiomyosarcomas	227	0.1%
IX(d.7) Synovial sarcomas	1,220	0.5%
IX(d.8) Blood vessel tumors	209	0.1%
IX(d.9) Osseous & chondromatous neoplasms of soft tissue	117	0.1%
IX(d.10) Alveolar soft parts sarcoma	185	0.1%
IX(d.11) Miscellaneous soft tissue sarcomas	497	0.2%
IX(e) Unspecified soft tissue sarcomas	1,494	0.7%
X Germ cell & trophoblastic tumors & neoplasms of gonads	13,349	5.9%
X(a) Intracranial & intraspinal germ cell tumors	2,342	1.0%
X(a.1) Intracranial & intraspinal germinomas	1,466	0.6%
X(a.2) Intracranial & intraspinal teratomas	592	0.3%
X(a.3) Intracranial & intraspinal embryonal carcinomas	25	0.0%
X(a.4) Intracranial & intraspinal yolk sac tumor	34	0.0%
X(a.5) Intracranial & intraspinal choriocarcinoma	16	0.0%
X(a.6) Intracranial & intraspinal tumors of mixed forms	209	0.1%
X(b) Extracranial & extragonadal germ cell tumors	1,689	0.7%
X(b.1) Germinomas: extracranial/extragonadal	174	0.1%
X(b.2) Malignant teratomas: extracranial/extragonadal	660	0.3%
X(b.3) Embryonal carcinomas: extracranial/extragonadal	17	0.0%
X(b.4) Yolk sac tumor: extracranial/extragonadal	386	0.2%
X(b.5) Choriocarcinomas: extracranial/extragonadal	204	0.1%
X(b.6) Other mixed germ cell: extracranial/extragonadal	248	0.1%
X(c) Malignant gonadal germ cell tumors	8,447	3.7%
X(c.1) Malignant gonadal germinomas	1,764	0.8%
X(c.2) Malignant gonadal teratomas	1,501	0.7%
X(c.3) Gonadal embryonal carcinomas	856	0.4%
X(c.4) Gonadal yolk sac tumor	813	0.4%
X(c.5) Gonadal choriocarcinoma	82	0.0%
X(c.6) Malignant gonadal tumors of mixed forms	3,431	1.5%
X(d) Gonadal carcinomas	493	0.2%
X(e) Other and unspecified malignant gonadal tumors	378	0.2%
XI Other malignant epithelial neoplasms and melanomas	21,375	9.4%
XI(a) Adrenocortical carcinomas	253	0.1%
XI(b) Thyroid carcinomas	9,104	4.0%
XI(c) Nasopharyngeal carcinomas	660	0.3%
XI(d) Malignant melanomas	6,043	2.7%
XI(e) Skin carcinomas	101	0.0%
XI(f) Other and unspecified carcinomas	5,214	2.3%
XI(f.1) Carcinomas of salivary glands	1,032	0.5%
XI(f.2) Carcinomas of colon and rectum	678	0.3%
XI(f.3) Carcinomas of appendix	704	0.3%
XI(f.4) Carcinomas of lung	529	0.2%
XI(f.5) Carcinomas of thymus	78	0.0%
XI(f.6) Carcinomas of breast	234	0.1%
XI(f.7) Carcinomas of cervix uteri	178	0.1%
XI(f.8) Carcinomas of bladder	350	0.2%
XI(f.9) Carcinomas of eye	35	0.0%
XI(f.10) Carcinomas of other specified sites	1,204	0.5%
XI(f.11) Carcinomas of unspecified site	192	0.1%

XII Other and unspecified malignant neoplasms	^[1]	^[1]
XII(a) Other specified malignant tumors	^[1]	^[1]
XII(a.1) Gastrointestinal stromal tumor	107	0.0%
XII(a.2) Pancreatoblastoma	38	0.0%
XII(a.3) Pulmonary blastoma and pleuropulmonary blastoma	211	0.1%
XII(a.4) Other complex mixed and stromal neoplasms	52	0.0%
XII(a.5) Mesothelioma	46	0.0%
XII(a.6) Other specified malignant tumors	^[2]	^[2]
XII(b) Other unspecified malignant tumors	426	0.2%
Not classified by ICCC or *in situ*	7,975	3.5%

[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	24,121	6.2%
1.1 Acute lymphoid leukemia	9,358	2.4%
1.2 Acute myeloid leukemia	9,194	2.4%
1.3 Chronic myeloid leukemia	3,732	1.0%
1.4 Other and unspecified leukemia	1,837	0.5%
2 Lymphomas	56,929	14.6%
2.1 Non-Hodgkin lymphoma	21,975	5.7%
2.2 Hodgkin lymphoma	34,954	9.0%
3 CNS and Oth Intracranial and Intraspinal Neo (all behav)	34,583	8.9%
3.1. Astrocytoma	11,728	3.0%
3.1.1 Specified low-grade astrocytic tumors	4,696	1.2%
3.1.2 Glioblastoma and anaplastic astrocytoma	4,363	1.1%
3.1.3 Astrocytoma, NOS	2,669	0.7%
3.2 Other glioma	6,278	1.6%
3.3 Ependymoma	2,285	0.6%
3.4. Medulloblastoma and other PNET	2,026	0.5%
3.4.1 Medulloblastoma	1,131	0.3%
3.4.2 Supratentorial PNET	895	0.2%
3.5 Other specified intracranial and intraspinal neoplasms	10,372	2.7%
3.6 Unspecified intracranial and intraspinal neoplasms	1,894	0.5%
3.6.1 Unspec malignant intracranial and intraspinal neo	360	0.1%
3.6.2 Unspec ben/border intracran. and intraspin neo	1,534	0.4%
4 Osseous & Chondromatous Neoplasms	10,136	2.6%
4.1 Osteosarcoma	4,418	1.1%
4.2 Chondrosarcoma	1,367	0.4%
4.3 Ewing tumor	3,411	0.9%
4.4 Other specified and unspecified bone tumors	940	0.2%
5 Soft Tissue Sarcomas	16,608	4.3%
5.1 Fibromatous neoplasms	4,157	1.1%
5.2 Rhabdomyosarcoma	1,871	0.5%
5.3 Other soft tissue sarcoma	10,580	2.7%

Values	Frequency	Percentage
5.3.1 Specified soft tissue sarcoma	8,196	2.1%
5.3.1.1 Specified (excluding Kaposi sarcoma)	6,609	1.7%
5.3.1.2 Kaposi sarcomae	1,587	0.4%
5.3.2 Unspecified soft tissue sarcoma	2,384	0.6%
6 Germ Cell and Trophoblastic Neoplasms	45,962	11.8%
6.1 Germ cell and trophoblastic neoplasms of gonads	41,960	10.8%
6.2 Germ cell and trophoblastic neo of nongonadal sites	4,002	1.0%
6.2.1 Intracranial (all behaviors)	1,509	0.4%
6.2.2 Other nongonadal	2,493	0.6%
7 Melanoma and Skin Carcinomas	36,488	9.4%
7.1 Melanoma	36,183	9.3%
7.2 Skin carcinomas	305	0.1%
8 Carcinomas	112,980	29.1%
8.1 Thyroid carcinoma	48,726	12.5%
8.2 Other carcinoma of head and neck	6,486	1.7%
8.2.1 Nasopharyngeal carcinoma	1,186	0.3%
8.2.2 Other sites in lip, oral cavity and pharynx	4,685	1.2%
8.2.3 Nasal cav,mid ear,sinus,larynx,ill-def head/neck	615	0.2%
8.3 Carcinoma of trachea,bronchus, and lung	2,731	0.7%
8.4 Carcinoma of breast	14,239	3.7%
8.5 Carcinoma of genitourinary tract	22,723	5.8%
8.5.1 Carcinoma of kidney	4,602	1.2%
8.5.2 Carcinoma of bladder	1,998	0.5%
8.5.3 Carcinoma of gonads	3,475	0.9%
8.5.4 Carcinoma of cervix and uterus	12,025	3.1%
8.5.5 Carc of oth and ill-defined sites	623	0.2%
8.6 Carcinoma of gastrointestinal tract	15,958	4.1%
8.6.1 Carcinoma of colon and rectum	10,574	2.7%
8.6.2 Carcinoma of stomach	1,604	0.4%
8.6.3 Carcinoma of liver and intrahepatic bile ducts	1,544	0.4%
8.6.4 Carcinoma of pancreas	1,120	0.3%
8.6.5 Carc oth and ill-def sites, gastrointestinal tract	1,116	0.3%
8.7 Carcinoma of other and ill-defined sites	2,117	0.5%
8.7.1 Adrenocortical carcinoma	331	0.1%
8.7.2 Carcinoma of other and ill-defined sites, NOS	1,786	0.5%
9 Miscellaneous specified neoplasms, NOS	8,744	2.2%
9.1 Other pediatric and embryonal tumors, NOS	898	0.2%
9.1.1 Wilms tumor	177	0.0%
9.1.2 Neuroblastoma	261	0.1%
9.1.3 Other pediatric and embryonal tumors, NOS	460	0.1%
9.2 Other specified and embryonal tumors, NOS	7,846	2.0%
9.2.1 Paraganglioma and glomus tumors	328	0.1%
9.2.2 Other specified gonadal tumors	665	0.2%
9.2.3 Myeloma, mast cell, misc lymphoreticular neo, NOS	1,374	0.4%
9.2.4 Other specified neoplasms, NOS	5,479	1.4%
10 Unspecified Malignant Neoplasms	2,084	0.5%
Unclassified and Non-Malignant	40,037	10.3%

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	1,549,856	6.8%
1 Hodgkin Lymphoma	117,836	0.5%
1(a) Classical Hodgkin lymphoma	111,831	0.5%
1(a)1 Lymphocyte-rich/mixed cell/lymphocyte depleted	19,334	0.1%
1(a)1.1 Lymphocyte-rich	4,143	0.0%
1(a)1.2 Mixed cellularity	13,688	0.1%
1(a)1.3 Lymphocyte-depleted	1,503	0.0%
1(a)2 Nodular sclerosis	64,575	0.3%
1(a)3 Classical Hodgkin lymphoma, NOS	27,922	0.1%
1(b) Nodular lymphocyte predominant Hodgkin lymphoma	6,005	0.0%
2 Non-Hodgkin lymphoma	1,385,539	6.1%
2(a) Non-Hodgkin lymphoma, B-cell	1,274,970	5.6%
2(a)1 Precursor Non-Hodgkin lymphoma, B-cell	49,934	0.2%
2(a)2 Mature Non-Hodgkin lymphoma, B-cell	1,159,797	5.1%
2(a)2.1 Chronic/Sm/Prolymphocytic/Mantle B-cell NHL	289,630	1.3%
2(a)2.1.1 Chronic/Small lymphocytic leuk/lymph	253,354	1.1%
2(a)2.1.2 Prolymphocytic leukemia, B-cell	1,172	0.0%
2(a)2.1.3 Mantle-cell lymphoma	35,104	0.2%
2(a)2.2 Lymphoplasmacytic lymphoma/Waldenstrom	26,257	0.1%
2(a)2.2.1 Lymphoplasmacytic lymphoma	11,279	0.0%
2(a)2.2.2 Waldenstrom macroglubulinemia	14,978	0.1%
2(a)2.3 Diffuse large B-cell lymphoma (DLBCL)	299,096	1.3%
2(a)2.3.1 DLBCL, NOS	296,199	1.3%
2(a)2.3.2 Intravascular large B-cell lymphoma	518	0.0%
2(a)2.3.3 Primary effusion lymphoma	371	0.0%
2(a)2.3.4 Mediastinal large B-cell lymphoma	2,008	0.0%
2(a)2.4 Burkitt lymphoma/leukemia	17,084	0.1%

Values	Frequency	Percentage
2(a)2.5 Marginal-zone lymphoma (MZL)	79,147	0.3%
2(a)2.5.1 Splenic MZL	7,114	0.0%
2(a)2.5.2 Extranodal MZL, MALT type	47,747	0.2%
2(a)2.5.3 Nodal MZL	24,286	0.1%
2(a)2.6 Follicular lymphoma	164,025	0.7%
2(a)2.7 Hairy-cell leukemia	12,704	0.1%
2(a)2.8 Plasma cell neoplasms	271,626	1.2%
2(a)2.8.1 Plasmacytoma	18,066	0.1%
2(a)2.8.2 Multiple myeloma/plasma-cell leuk	253,560	1.1%
2(a)2.9 Heavy chain disease	228	0.0%
2(a)3 Non-Hodgkin lymphoma, B-cell, NOS	65,239	0.3%
2(b) Non-Hodgkin lymphoma, T-cell	89,281	0.4%
2(b)1 Precursor Non-Hodgkin lymphoma, T-cell	4,742	0.0%
2(b)2 Mature Non-Hodgkin lymphoma, T-cell	83,993	0.4%
2(b)2.1 Mycosis fungoides/Sezary syndrome	18,177	0.1%
2(b)2.1.1 Mycosis fungoides	17,504	0.1%
2(b)2.1.2 Sezary syndrome	673	0.0%
2(b)2.2 Peripheral T-cell lymphoma	49,313	0.2%
2(b)2.2.1 Peripheral T-cell lymphoma, NOS	17,271	0.1%
2(b)2.2.2 Angioimmunoblastic T-cell lymphoma	5,319	0.0%
2(b)2.2.3 Subcutan panniculitis-like T-cell lymph	450	0.0%
2(b)2.2.4 Anaplastic lar cell lymph, T-/Null-cell	11,271	0.0%
2(b)2.2.5 Hepatosplenic T-cell lymphoma	370	0.0%
2(b)2.2.6 Enteropathy-type T-cell lymphoma	575	0.0%
2(b)2.2.7 Cutaneous T-cell lymphoma, NOS	10,186	0.0%
2(b)2.2.8 Prim cutaneous anaplastic lar cell lymph	3,871	0.0%
2(b)2.3 Adult T-cell leukemia/lymphoma	8,690	0.0%
2(b)2.4 NK/T-cell lymph, nasal-type/aggres NK leuk	2,612	0.0%
2(b)2.5 T-cell large granular lymphocytic leukemia	3,666	0.0%
2(b)2.6 Prolymphocytic leukemia, T-cell	1,535	0.0%
2(b)3 Non-Hodgkin lymphoma, NOS, T-cell	546	0.0%
2(c) Non-Hodgkin lymphoma, unknown lineage	21,288	0.1%
2(c)1 Precursor lymphoblastic leuk/lymph, unk lineage	8,214	0.0%
2(c)2 Prolymphocytic leukemia, unknown lineage	621	0.0%
2(c)3 Non-Hodgkin lymphoma, NOS, unknown lineage	12,453	0.1%
3 Composite Hodgkin lymphoma and NHL	1,927	0.0%
4 Lymphoid neoplasm, NOS	44,554	0.2%
Unclassified	21,228,399	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 USCS official federal cancer statistics.
- For more information, please see SEER coding manual at http://seer.cancer.gov/icd-o-3.

Values	Frequency	Percentage
Benign	475,456	2.1%
Borderline malignancy	47,479	0.2%
In situ	1,549,291	6.8%
Malignant	20,378,654	89.5%
Only malignant in ICD-O-3	311,621	1.4%
Only malignant 2010+	15,754	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 and 759

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 2004–2013. This stage variable can be used for diagnosis years 2001–2013.

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 2013, then the Derived Summary Stage 2000 (Derived SS2000) variable was used.
  - If the Derived Summary Stage 2000 variable was blank and a valid value was available for the Summary Stage 2000 variable, that value was used to populate the merged variable. For example, if the case was diagnosed in 2013 and Derived Summary Stage was blank, but Summary Stage had a value of local, then the merged variable was coded as local stage. Otherwise, the merged variable would be left blank.
- For more information about SEER Summary Stage 2000 and Derived SS2000 variables, please review https://cancerstaging.org/cstage/Pages/default.aspx.

Values		
In situ	1,922,809	8.4%
Localized only	9,359,082	41.1%
Regional, direct extension only	1,517,920	6.7%
Regional, regional lymph nodes only	1,606,892	7.1%
Regional, direct extension and regional lymph nodes	942,916	4.1%
Regional, NOS	234,658	1.0%
Distant site(s)/node(s) involved	4,949,557	21.7%
Not applicable	522,822	2.3%
Unknown/unstaged/unspecified-	1,721,599	7.6%

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		
Not a paired site	12,578,729	55.2%
Right - origin of primary	4,936,907	21.7%
Left - origin of primary	4,568,911	20.1%
Only one side - side unspecified	48,034	0.2%
Bilateral, single primary	177,902	0.8%
Paired site: midline tumor	41,112	0.2%
Paired site, but no information concerning laterality	426,660	1.9%

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	16,536,849	^[1]
1st of 2 or more primaries	1,861,853	^[1]
2nd of 2 or more primaries	3,206,717	^[1]
3rd of 3 or more primaries	531,371	^[1]
4th of 4 or more primaries	91,290	^[1]
5th of 5 or more primaries	18,324	^[1]
6th or more primaries[2]	8,490	^[1]
Only one state registry-defined neoplasm	501,105	^[1]
1st of 2 or more state registry-defined neoplasms	9,654	^[1]
2nd of 2 or more state registry-defined neoplasms	10,958	^[1]
3rd of 3 or more state registry-defined neoplasms	748	^[1]
4th of 4 or more state registry-defined neoplasms	187	^[1]
5th of 5 or more state registry-defined neoplasms	86	^[1]
6th or more state registry-defined neoplasms[1]	37	^[1]
Carcinoma *in situ* of the Cervix diagnosed 1/1/1996 or later	^[3]	^[3]

Values	Frequency	Percentage
Unknown sequence number - federally required *in situ* or malignant tumors	115	^[1]
Unknown sequence number - state registry-defined neoplasms	304	^[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. For age-adjusted rates by ethnicity, the variable "Origin recode NHIA (Hispanic/Non-Hispanic)" should be used.
- 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 Arkansas, 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 Arkansas, Delaware, Kansas, Kentucky, Massachusetts, and Pennsylvania.

Values	Frequency	Percentage
Non-Spanish-Hispanic-Latino	21,272,041	93.4%
Mexican	349,043	1.5%
Puerto Rican	80,325	0.4%
Cuban	64,510	0.3%
South or Central American excluding Brazil	115,677	0.5%
Other specified Spanish/Hispanic origin including Europe	36,451	0.2%
Spanish/Hispanic/Latino, NOS	622,550	2.7%
NHIA surname match only	207,320	0.9%
Dominican Republic	27,044	0.1%
Invalid Value(s)	3,294	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 the variable.

Values	Frequency	Percentage
1890	^1	^1
1891	^1	^1
1892	^1	^1
1893	^1	^1
1894	16	^2
1895	29	^2
1896	43	^2
1897	69	^2
1898	145	^2
1899	218	^2
1900	554	^2
1901	873	^2
1902	1,277	^2
1903	1,887	^2
1904	2,943	^2
1905	4,546	^2
1906	6,526	^2
1907	9,984	^2
1908	14,336	^2
1909	19,395	^2
1910	26,987	^2
1911	34,996	^2
1912	48,208	^2
1913	60,956	^2
1914	78,775	^2
1915	94,648	^2
1916	114,759	^2
1917	139,328	^2
1918	169,225	^2

Values	Frequency	Percentage
1919	191,814	^2
1920	238,284	^2
1921	279,147	^2
1922	303,091	^2
1923	336,667	^2
1924	375,209	^2
1925	399,124	^2
1926	424,036	^2
1927	457,635	^2
1928	470,885	^2
1929	480,667	^2
1930	507,663	^2
1931	501,606	^2
1932	511,908	^2
1933	498,674	^2
1934	525,434	^2
1935	536,080	^2
1936	537,383	^2
1937	546,106	^2
1938	559,416	^2
1939	548,375	^2
1940	553,934	^2
1941	564,615	^2
1942	605,735	^2
1943	601,556	^2
1944	550,183	^2
1945	517,644	^2
1946	583,542	^2
1947	616,115	^2
1948	552,075	^2
1949	516,714	^2
1950	482,687	^2
1951	472,788	^2
1952	456,726	^2
1953	432,031	^2
1954	417,880	^2
1955	392,801	^2
1956	376,627	^2
1957	357,462	^2
1958	327,660	^2
1959	306,156	^2
1960	284,013	^2
1961	260,797	^2

Values	Frequency	Percentage
1962	235,450	^2
1963	213,799	^2
1964	190,906	^2
1965	164,729	^2
1966	146,416	^2
1967	130,582	^2
1968	120,198	^2
1969	112,681	^2
1970	106,561	^2
1971	93,665	^2
1972	80,610	^2
1973	70,155	^2
1974	64,941	^2
1975	59,111	^2
1976	54,762	^2
1977	51,799	^2
1978	47,613	^2
1979	45,505	^2
1980	42,116	^2
1981	38,919	^2
1982	35,877	^2
1983	31,781	^2
1984	29,287	^2
1985	26,628	^2
1986	24,083	^2
1987	21,954	^2
1988	20,376	^2
1989	18,620	^2
1990	17,000	^2
1991	15,445	^2
1992	13,716	^2
1993	12,265	^2
1994	11,308	^2
1995	10,188	^2
1996	9,684	^2
1997	9,340	^2
1998	9,797	^2
1999	9,529	^2
2000	9,632	^2
2001	9,588	^2
2002	8,951	^2
2003	8,524	^2
2004	8,083	^2

Values	Frequency	Percentage
2005	7,513	^2
2006	7,240	^2
2007	6,712	^2
2008	6,319	^2
2009	5,401	^2
2010	4,619	^2
2011	3,583	^2
2012	2,597	^2
2013	1,731	^2
2014	674	^2
Blank(s)	^1	^1

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

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

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

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

Description

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

Considerations for use

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

Values	Frequency	Percentage
January	2,010,762	8.8%
February	1,772,064	7.8%
March	1,938,063	8.5%
April	1,907,919	8.4%
May	1,924,743	8.4%
June	1,956,833	8.6%
July	1,871,941	8.2%
August	1,921,971	8.4%
September	1,827,205	8.0%
October	1,955,950	8.6%
November	1,783,979	7.8%
December	1,755,818	7.7%
Blank(s)	151,007	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	19,949,297	87.6%
Radiation treatment or medical oncology center (2006+)	438,456	1.9%
Laboratory only (hospital or private)	595,705	2.6%
Physician's office/private medical practitioner (LMD)	969,749	4.3%
Nursing/convalescent home/hospice	29,775	0.1%
Other hospital outpatient unit or surgery center (2006+)	795,272	3.5%

Additional Resources

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

Abbreviations

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