



Texas Department of State Health Services  
October 2014



*Texas Cancer Registry*

Cancer Epidemiology and Surveillance Branch  
Publication No. E10-11701

## **Cancer Research File Data Dictionary**

For Cancers Diagnosed in 1995 through 2011



*Texas Cancer Registry*

**Texas Cancer Registry (TCR)  
Cancer Epidemiology and Surveillance Branch  
Texas Department of State Health Services (DSHS)**

**TCR Incidence File Descriptions**

**Introduction:**

This document is a description of the TCR data items that are available to customers for analysis and research. However, TCR will only release data fields that have been submitted to the NAACCR and/or NPCR. This ensures that any data released has met a protocol of quality standards.

All 1995-2011 incidence files are based on the 1995-2011 Call for Data cancer incidence submission records sent by the Texas Cancer Registry (TCR) to the National Program of Cancer Registries (NPCR) and the North American Association of Central Cancer Registries (NAACCR) in November 2013. This file has complete (defined as at least 95%) statewide data for neoplasms diagnosed in 1995 through 2011. More recent data are available on a limited basis for certain types of studies. Although TCR collects records of all cancers diagnosed in the state of Texas (both residents and non-residents), the incidence files only contain records for Texas residents. Individual-level non-resident data is not released to researchers. Additionally, individual-level data reported by a Department of Veteran's Affairs facility, or by another state's central cancer registry that does not participate in the NAACCR National Interstate Data Exchange Agreement, will not be released to researchers.

Requested TCR data can be provided in a variety of formats. SEER\*Stat files contain population data and permit easy calculation of incidence rates. Other file formats available include SAS, Excel, and fixed-width or tab-delimited ASCII text; format provided may depend on the variables requested and the size of the final dataset. Files will be sent by secure file transfer protocol (TCR's WebPlus server). A TCR-created tumor identification number (MedRefID) will be included in all datasets.

For additional information on all data fields reported by TCR to NAACCR, see the Version 13 data dictionary and reporting standards (Standards for Cancer Registries, Vol II: Data Standards and Data Dictionary) at <http://www.naacr.org/StandardsandRegistryOperations/VolumeII.aspx>. NAACCR develops uniform cancer data reporting standards for consistency among all cancer registries in North America. For additional information on NPCR, visit <http://www.cdc.gov/cancer/npcr/>.

### Available TCR Incidence Files and Contents:

The following files contain 1,604,654 Texas resident malignant and in-situ cancers (excluding malignant non-genital basal and squamous cancers of the skin, and in situ cervix cases) diagnosed from 1995–2011. Benign and borderline brain and central nervous system cases are included for 2004–2011 only. Persons treated for cancer in Texas but who were residents of another state or country are not included.

#### SEER\*Stat files (permit incidence calculations)

1. Limited-Use Incidence - Texas, 1995–2011, Cut-off 11/30/2013, SEER\*Prep 2.5.2.
2. Incidence - Texas, 1995–2011, Cut-off 11/30/2013, SEER\*Prep 2.5.2 (Confidential). *[Contains confidential\* data—requires DSHS-Institutional Review Board (IRB) approval].*

#### SAS (can be converted to other formats as described above)

1. Limited\_use\_final\_epi\_analysis\_2011.sas7bdat, Cut-off 11/30/2013.
2. final\_epi\_analysis\_2011.sas7bdat, Cut-off 11/30/2013. *[Contains confidential\* data —requires DSHS-Institutional Review Board (IRB) approval].*

### Acknowledgment of TCR Incidence Data:

The TCR requests that any person or organization reporting results or analyses using TCR incidence data include the following statement of acknowledgement in the text or frontispiece of the analysis, presentation, report, or publication:

“Cancer incidence data have been provided by the Texas Cancer Registry, Cancer Epidemiology and Surveillance Branch, Texas Department of State Health Services, 11 + j<sup>o</sup> th Street, Austin, TX 787 , <http://www.dshs.state.tx.us/tcr/default.shtm>, or (512) - 0.”

### Acknowledgement of CDC Support:

We acknowledge the National Program of Cancer Registries (NPCR), Centers for Disease Control and Prevention (CDC) for the funds that helped support the availability of these data.

-----

\*Confidential variables are only available for release with prior Texas Department of State Health Services (DSHS) Institutional Review Board (IRB) approval. Variables that may be considered by the TCR as confidential are noted with an asterisk. Instructions for applying for confidential TCR data are available at <http://www.dshs.state.tx.us/tcr/irb.shtm>. All day components of dates are considered confidential.

**Important Notes:**

1. One of the great strengths of the Limited-Use SEER\*Stat Incidence database is the inclusion of “site recodes” which are generated by SEER\*Prep and take into account the anatomic site and cell type of the tumor. These recodes, which are based on SEER definitions, are extremely useful when analyzing non-solid and/or childhood cancers.
2. **SEER\*Stat User-Specified** variables *should not* be used to generate any rates due to **NOT** being linked to correct population data.

(variables in order by category, then by NAACCR item #, then by Short Name)

Availability in complete, edited analysis files with 1995-2011 cases, based on Call for Data files submitted November 2013

Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	NAACCR	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files	Confiden-	In Confidential	In Confidential	In Limited Use	In Limited Use
		Version 13					tial?	SEER*Stat	SAS incidence	SEER*Stat	SAS incidence
		Item No, if applicable					(yes=X)	incidence file	file	incidence file	file
<b>Patient and Demographics</b>											
Patient ID Number	Sequential identification number assigned to each new patient with a record in the TCR database. Unique for each patient (not for each tumor, because patients may have more than 1 tumor in the database). In combination with Sequence Number - Central [380], can be used to identify multiple tumors in same patient.	20	PatientID	1995	present	Sequentially assigned number; unique per person.		X	X	X	X
Race 1	Patient's reported race.	160	Race1	1995	present	01=White; 02=African American; 03=American Indian, Aleutian, Eskimo; 98=Other; 99=Unknown. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail.		X	X	X	X
Race 2	Patient's reported race, with additional detail for Asians and Pacific Islanders.	161	Race2	2001	present	01=White; 02=Black; 03=American Indian, Aleutian, Eskimo; codes 04-97 (not sequential) each indicate specific Asian and Pacific Islander races; 98=other; 99=unknown; blank=not coded. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail.		X	X	X	X
Spanish/Hispanic Origin	Identifies persons of Spanish or Hispanic origin.	190	Hispanic	1995	present	0=Non-Spanish/non-Hispanic; 1=Mexican; 2=Puerto Rican; 3=Cuban; 4=South or Central American (except Brazil); 5=Other specified Spanish/Hispanic (includes European); 6=Spanish, NOS/Hispanic, NOS/Latino; NOS; 7=Spanish surname only; 8=Dominican Republic (date of diagnosis: 2005-forward); 9=unknown. See NAACCR standards manual for more information.		X	X	X	X
NHIA Derived Hisp Origin	Classifies Hispanic ethnicity based on NAACCR algorithm. Only available for records obtained from complete research file, not for records obtained from production database. Available for all diagnosis years.	191	NHIA			Same codes as item #190.		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
IHS Link	Race based on linkage with Indian Health Service patient database. Linkage updated each year. Available for all diagnosis years.	192	IHSLink			0=match; 1=no match; Blank=link not attempted		X	X	X	X
Race--NAPIIA(derived API)	NAACCR Asian Pacific Islander Identification Algorithm. Derived API that recodes some single-race cases with a Race1 code of 96 to a more specific Asian race category, based on an algorithm that makes use of the birthplace and name fields (first, last, and maiden names). Only available for records obtained from complete research file, not for records obtained from production database. Available for all diagnosis years.	193	RaceNAPIIA			See NAACCR standards manual for more information. See white paper at <a href="http://naaccr.org/LinkClick.aspx?fileticket=3HnBhlmhkBs%3D&amp;tabid=92&amp;mid=432">http://naaccr.org/LinkClick.aspx?fileticket=3HnBhlmhkBs%3D&amp;tabid=92&amp;mid=432</a>		X	X	X	X
Sex	Patient's sex	220	Sex	1995	present	1=Male; 2=Female; 3=Other (Hermaphrodite); 4=Transsexual; 9=Not Stated/Unknown		X	X	X	X
Date of Birth	Patient's date of birth. Researchers may request year of birth without IRB approval. Often blank.	240	BirthDate	1995	present	YYYYMMDD, no dashes or slashes; 999 or 9999=Unknown. In SEER*Stat file this is 'Year of Birth' and ranges from 1886-2010.	X	X (as month and year)	X	X (as month and year)	X (as month and year)
Date of Birth Flag	If there is no appropriate value stored in Date of Birth[240] field, this code explains why.	241	BirthDateFlag	2010		12=Birth date is unknown; blank=A valid date is provided in Date of Birth [240], or date was not expected to have been transmitted.			X		X
Birthplace--State	State in which patient was born. This NAACCR item and #254 were converted from the single NAACCR item 250 in v12.2.	252	BPState	1998	present	See NAACCR Data Standards for Cancer Registries, Version 13 Appendix B for geocodes ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
Birthplace--Country	Country in which patient was born. This NAACCR item and #252 were computed from the single NAACCR item 250 in v12.2. NOTE: This field is usually blank.	254	BPCountry	1998	present	See NAACCR Data Standards for Cancer Registries, Version 13 Appendix B for geocodes ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
Name--Last	Patient's last name	2230	LastName	1995	present	Free entry	X	X	X		
Name--First	Patient's first name	2240	FirstName	1995	present	Free entry	X	X	X		
Name--Middle	Patient's middle name	2250	MiddleName	1995	present	Free entry	X	X	X		
Name--Suffix	Patient's name suffix (Jr, III, etc.) NOTE: This field is usually blank.	2270	NameSuffix	1995	present	Free entry	X	X	X		

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
Name--Alias	Records an alternate name or "Also Known As" used by the patient. Should not include Maiden Name, which is entered into Name--Maiden [2390]. NOTE: This is usually blank.	2280	Alias	1995 2006	2003	Free entry	X	X	X	
Social Security Number	Patient's Social Security Number. NOTE: SSN will not be released but can be used for linkage purposes.	2320	SocSec	1995	present	Nine-digit number, no dashes or slashes. 999999999=Unknown.	X	X	X	
Name--Maiden	Patient's maiden name if female	2390	MaidName	1995	present	Free entry	X		X	
Age recode	18 Age Groups that correspond to the population data. (old age recode; ages 0-5 is lowest)	SEER*Stat recode				in categories of years: 00, 01-04, 05-09, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+ years		X		X
Age recode with <1 year olds	19 Age Groups that correspond to the population data.	SEER*Stat recode	SEER*Stat recode			Years of age		X		X
Age recode with <1 year olds and 100+	22 Age groups that correspond to the population data.	SEER*Stat recode				in categories of years: 00, 01-04, 05-09, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85-89, 90-94, 95-99, 100+		X		X
Age recode with single ages and 85+	Single ages up to 85+	SEER*Stat recode				00 years, 01 years ... sequentially through 84 years; 85+		X		X
Day of birth		SEER*Stat recode						X		
Month of birth		SEER*Stat recode						X		X
Origin recode SEER (Hispanic, Non-Hisp)	See population documentation on which race variable to use when merging with different versions of population data.	SEER*Stat recode				Spanish-Hispanic-Latino; Non-Spanish-Hispanic-Latino; Unknown		X		X
Race recode (W, B, AI, API)	For detailed information regarding major changes to this field in 2010, please see the following website: <a href="http://seer.cancer.gov/seerstat/variables/seer/race_ethnicity/">http://seer.cancer.gov/seerstat/variables/seer/race_ethnicity/</a>	SEER*Stat recode				Allowable values: White; Black, American Indian/Alaska Native; Asian or Pacific Islander; Other unspecified; Unknown		X		X
Race Recode (White, Black, Other)		SEER*Stat recode						X		X
Race Recode (White, Non-White)		SEER*Stat recode						X		X
Year of birth		SEER*Stat recode						X		X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
CustomRaceTCR	Creates a merged variable which shows rates by race (with non-Hispanic and Hispanic designation for whites), and by Hispanic ethnicity (of any race), and includes persons of Other and Unknown race (it is not possible to generate a rate for persons of unknown race due to lack of denominator.)	SEER*Stat User Defined Merged				White (White, Hispanic; White, Non-Hispanic); Black; Asian/Pacific Islander; American Indian/Alaska Native; Hispanic (of any race); Other & Unknown Race		X		X
<b>Tumor and patient characteristics at time of diagnosis</b>										
Address at DX--City	City where patient lived at time of diagnosis	70	DxCity	1995	present	Free entry	X		X	
Address at DX--State	Patient's state of residence at date of diagnosis of reportable tumor.	80	DxState	1995	present	USPS standard abbreviations--TX, CA, NM, etc.Complete, edited incidence files contain only tumors diagnosed in persons reported to be TX residents at the time of diagnosis.		X	X	X X
County at DX	FIPS code of county where it was reported that the patient lived at date of diagnosis. In TCR database, this is updated annually with FIPS code of county of geocoded address, if that is different from reported FIPS	90	DxCounty	1995	present	FIPS code of reported TX county (001-999), based on FIPS publication "Counties and Equivalent Entities of the United States, Its Possessions, and Associated Areas." FIPS code of TX county assigned by geocoding is also available if requested by researcher.	X	X	X	X X
Address at DX--Postal Code	Patient's ZIP Code of residence at date of diagnosis	100	DxPostalZip	1995	present	United States Postal Codes, no hyphens, ZIP or ZIP+4 (99999=Unknown)	X	X	X	
Census Tract 1970/80/90	Census tract of patient's residence at date of diagnosis based on the 1970/80/90 Census designations.	110	CenTract708090			Codes: Census Tract Codes 000100-949999; BNA Codes: 950100-998999; 000000 area not census-tract; 999999 area census-tract, census tract not available; blank= Census Tract 1970/80/90 not coded.	X	X	X	
Census Tract 2000	Census tract of patient's residence at date of diagnosis based on the 2000 Census designations. Data have been geocoded using specialized software by DSHS Center for Health Statistics. Calculated 1995-present.	130	CenTract2000			Census 2000 tract codes--see Census 2000 <a href="http://www.census.gov/main/www/cen2000.html">http://www.census.gov/main/www/cen2000.html</a>	X	X	X	
Census Tract 2010	Census tract of patient's residence at date of diagnosis based on the 2010 Census designations. Data have been geocoded using specialized software by DSHS Center for Health Statistics. Calculated 1995-present.	135	CenTract2010			Census 2010 tract codes--see Census 2010 <a href="http://www.census.gov/2010census/">http://www.census.gov/2010census/</a>	X	X	X	

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Census Tract Poverty Indicator	A 4-level categorical variable calculating the poverty level of the census tract in which the patient was diagnosed. For cases diagnosed 1995-2005, based on 2000 US Census for Texas; for cases diagnosed 2006-2011, based on 2010 US Census for Texas. SAS program to create variable provided by NAACCR <a href="http://www.naacr.org/Research/DataAnalysisTools.aspx">http://www.naacr.org/Research/DataAnalysisTools.aspx</a>	145	PovertyInd			1<=5%; 2=5-9.9%; 3=10-19.9%; 4>=20%; 5=Unknown.		X	X	X	X
Marital Status at DX	Code for patient's marital status at time of diagnosis for the reportable tumor. If patient has multiple tumors, marital status may be different for each tumor.	150	Marital			1=single (never married), 2=married (including common law); 3=separated; 4=divorced; 5=widowed; 6=unmarried or domestic partner (same sex or opposite sex, registered or unregistered, other than common law marriage); 9=unknown			X		
Age at Diagnosis	Age of patient at time of diagnosis or age at time of death for death certificate only cases. 5 or 10 year age groups are available upon request.	230	AgeDX			Number with leading zeros; 999=Unknown.		X	X	X	X
Census Occ Code 1970-2000	Census Occupation Code 1970-2000. Majority are missing.	270	OccCen			3-digit code. See more info at: <a href="http://www.census.gov/people/io/">http://www.census.gov/people/io/</a>			X		X
Census Ind Code 1970-2000	Census Industry Code 1970-2000. Majority are missing.	280	IndCen			3-digit code. See more info at: <a href="http://www.census.gov/people/io/">http://www.census.gov/people/io/</a>			X		X
Occupation Source	Code that best describes source of occupation information provided on this patient. Applied by central cancer registry.	290	OccSrc			0=Unknown occupation; 1=reporting facility; 2=death certificate; 3=interview; 7=other source; 8=N/A age < 14 at dx; 9=unknown source; blank=not collected			X		X
Industry Source	Code that best describes source of industry occupation information provided on this patient. Applied by central cancer registry.	300	IndSrc			0=Unknown industry; 1=reporting facility; 2=death certificate; 3=interview; 7=other source; 8=N/A age < 14 at dx; 9=unknown source; blank=not collected			X		X
Text--Usual Occupation	Patient's usual job or type of work. NOTE: This field has not undergone any edits and must be coded by the researcher to be analyzed appropriately. This field is frequently from death certificate information. Majority are missing.	310	TxUsualOcc	2010		Free entry	X		X		

		NAACCR Version 13									
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files	Confidential? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Text--Usual Industry	Patient's usual type of business or industry. NOTE: This field has not undergone any edits and must be coded by the researcher to be analyzed appropriately. This field is frequently from death certificate information. Majority are missing.	320	TxUsualInd	2010		Free entry	X		X		
Census/Occ/Ind Sys 70-00	Census Occupation/Industry Coding System 1970-2000	330	OcclndCodSys			1=1970 Census; 2=1980 Census; 3=1990 Census; 4=2000 Census; 5=2010 Census; 6=Other coding system; 9=Unknown coding system; blank=not collected			X		X
Census Tr Cert 1970/80/90	Describes the basis of assignment of census tract for the record. Calculated 1995-present.	364	CenTractCer708090			1=Street address; 2-4=ZIP Code Only; 5=PO Box ZIP; 9=Unable to code		X	X	X	X
Census Tr Certainty 2000	Describes the basis of assignment of census tract for the record. Calculated 1995-present.	365	CenTractCer2000			1=Street address; 2-4=ZIP Code Only; 5=PO Box ZIP; 9=Unable to code		X	X	X	X
Census Tr Certainty 2010	Describes the basis of assignment of census tract for the record. Calculated 1995-present.	367	CenTractCer2010			1=Street address; 2-4=ZIP Code Only; 5=PO Box ZIP; 9=Unable to code		X	X	X	X
Sequence Number--Central	Sequence of all reportable neoplasms over the lifetime of the person, with 'reportable' defined by the Central Registry. In combination with TCR Patient ID [20], can be used to identify multiple tumors in same patient.	380	SeqNumCentr	1995	present	00=one primary only; 01=first of two or more primaries; 02=second of two or more primaries; etc. through 59; 60=one nonmalignant tumor; 61 first of 2 or more nonmalignant tumors, etc.; 88=unspecified nonmalignant; 99=unspecified.		X	X	X	X
Date of Diagnosis	Date of initial cancer diagnosis; earliest date the primary cancer was diagnosed and clinically or microscopically confirmed by a recognized medical practitioner, regardless of whether the diagnosis was made at the reporting facility or elsewhere. By NAACCR/CDC standards, for cases where information is derived solely from death certificate, the date of diagnosis is the date of death. Researchers may request month and year of diagnosis without IRB approval.	390	DxDate	1995	present	YYYYMMDD Format (no slashes). In SEER*Stat this is stored as Month of Diagnosis with possible values are January-December; Unknown; and Year of Diagnosis.	X	X (as month and year)	X	X (as month and year)	X (as month and year)
Date of Diagnosis Flag	If Date of Diagnosis [390] has no appropriate value, this flag explains why.	391	DxDateFlag	2010		12=Date of diagnosis is unknown; blank=a valid date is provided in Date of Diagnosis [390], or the date was not expected to have been transmitted.			X		X

		NAACCR Version 13				Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Primary Site	Topography; anatomical site of neoplasm; ICD-O-3 code for primary site of the reported tumor.	400	PSite	1995	present	Codes 000-809 from International Classification of Diseases for Oncology, Third Edition (ICD-O-3), Topography Section. The decimal point is eliminated.		X	X	X	X
Laterality	Location of primary tumor (left, right, bilateral, etc.)	410	Lateral	1995	present	0=Not a paired site; 1=Right origin of primary; 2=Left origin of primary; 3=Only one side involved, left or right not stated; 4=Bilateral involvement, lateral origin unknown; 9=Paired site, no information concerning laterality.		X	X	X	X
Grade	The degree of abnormality of cancer cells (degree of differentiation). Grading and differentiation codes of 1-4, 9 are defined in ICD-O-2, 1992. Because reporting requirements and medical terminology have changed over time, care should be exercised when analyzing this information.	440	Grade	1995	present	ICD-O-3, page 67: 1=Well differentiated; 2=Moderately differentiated; 3=Poorly differentiated; 4=Undifferentiated; 5=T cell; 6=B cell, pre-B, B-precursor; 7=Null cell, Non T-non B; 8=Natural Killer (NK) cell; 9=Unknown/Undetermined, not stated or not applicable.		X	X	X	X
Site Coding Sys--Current	Code that best describes how the primary site currently is coded. If converted, this field shows the system to which it is converted.	450	SiteCodSysCur			5=ICD-O, Third Edition; 6=ICD-10. See NAACCR Data Standards for Cancer Registries, Version 13 for additional codes.			X		X
Morph Coding Sys--Current	Code that best describes how morphology is currently coded. If converted, this field shows the system it is converted to.	470	MorphCodSysCur			7=ICD-O, Third Edition; 8=ICD-O, Third Edition, plus 2008 WHO hematopoietic/lymphoid new terms used for conditions diagnosed 1/1/2010 and later. See NAACCR Data Standards for Cancer Registries, Version 13 for additional codes.			X		X
Diagnostic Confirmation	How a cancer diagnosis was diagnostically confirmed.	490	DxConf	1995	present	1=histology; 2=cytology, 3=histology plus immunophenotyping and/or genetic studies (only for hematopoietic & lymphoid neoplasms M-9590/3-9992/3); 4=microscopic, NOS, 5=laboratory test/marker study, 6=direct visualization w/o microscopic confirmation, 7=radiography and/or other imaging w/o microscopic confirmation, 8=clinical diagnosis only (other than 5, 6, or 7), 9=unknown whether or not microscopically confirmed; death certificate only		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Type of Reporting Source	What type of source documents were used to abstract the reported tumor. Death certificate only cases may be missing more information than cases from other sources and are frequently excluded from analysis.	500	TypeRepSrc	1995	present	1=Hospital/Facility; 2=Radiation Treatment Center or Medical Oncology Center; 3=Laboratory; 4=Physician's Office/private medical practitioner; 5=Nursing/Convalescent Home/Hospice; 6=Autopsy Only; 7=Death Certificate Only; 8=Other		X	X	X	X
Histologic Type ICD-O-3	ICD-O-3 code for the histologic type of reported tumor. ICD-O-2 codes from 1995-2000 converted to ICD-O-3.	522	HistTypeICDO3	1995	present	Four-digit codes 8000-9989 from International Classification of Diseases for Oncology, Third Edition (ICD-O-3), Morphology Section.		X	X	X	X
Behavior Code ICD-O-3	ICD-O-3 code for behavior of reported tumor. ICD-O-2 codes from 1995-2000 converted to ICD-O-3. SEER requires registries to collect malignancies with in situ/2 and malignant/3 behavior codes as described by ICD-O-3. SEER requires registries to collect benign/0 and borderline/1 intracranial and CNS tumors for cases diagnosed on or after 1/1/2004. Behavior is the fifth digit of the morphology code [NAACCR item 522] after the slash. See ICD-O-3 for discussion of the behavior code.	523	BehaviorICDO3	1995	present	0=Benign (reportable only for intracranial and CNS sites) ; 1=Borderline (reportable only for intracranial and CNS sites); 2=In Situ; 3=Malignant. All ICD-O-2 behaviors for 1995-2000 were converted to ICD-O-3.		X	X	X	X
Reporting Facility	Identifies the facility that reported the tumor. In the case of multiple reports from multiple facilities, the first reporter, or all reporters, are available. In SAS incidence file, the RepHosp_Best variable is the facility providing the report with the best class of case.	540	RepHosp	1995		10-digit numeric code assigned by TCR to identify the reporting facility. Hierarchy of class of case codes, from best to worst: 12, 14, 11, 13, 10, 22, 20, 21, 00, 30-99. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail and code definitions: <a href="http://www.naacccr.org/StandardsandRegistryOperations/Volumell.aspx">http://www.naacccr.org/StandardsandRegistryOperations/Volumell.aspx</a>	X		X		

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Class of Case	Class of case identifies the role that the facility plays in the patient's diagnosis and treatment. This variable is generally used by reporting facilities for analysis of treatment and survival for their reported cases. (Provided as both 'Analytic' and 'ClassCase' variables.)	610	ClassCase	1998		Provided as both a binary variable (named 'Analytic') indicating either '1' (analytic) or '0' (non-analytic), and as a Best class of case variable (named 'ClassCase_Best') containing 2-digit class of case code. Analytic cases (initial diagnosis and/or some or all of treatment was administered at the facility) are defined as cases having any report with class of case codes 00-22; non-analytic cases (neither initial diagnosis nor first course treatment was provided at the reporting facility) are coded 30-99. Hierarchy of class of case codes, from best to worst: 12, 14, 11, 13, 10, 22, 20, 21, 00, 30-99. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail and code definitions: <a href="http://www.naacr.org/StandardsandRegistryOperations/Volumell.aspx">http://www.naacr.org/StandardsandRegistryOperations/Volumell.aspx</a> .		X	X	X	X
Primary Payer at DX	Primary payer/insurance carrier at time of initial diagnosis and/or treatment.	630	PriPayerDx	2007	present	01=not insured; 02=not insured, self-pay; 10=insurance, NOS; 20-21: types of private insurance; 31, 35=types of Medicaid; 60-64=types of Medicare; 65=TRICARE; 66=Military. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail.		X	X	X	X
Addr at DX--No & Street	Patient's street address at date of diagnosis	2330	DxNumAndSt	1995	present	Free text, with USPS Postal Addressing Standards format and abbreviations.	X		X		
Latitude	Latitude of the patient's residence at the date of diagnosis.	2352	Latitude			Data units are in decimal degrees, not degrees, minutes and seconds. Data have been geocoded using specialized software by DSHS Center for Health Statistics, with up to 6 digits after the decimal.	X		X		

Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	NAACCR Version 13		First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files	Confidential? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
		Item No, if applicable	SAS variable name (if applicable)								
Longitude	Longitude of the patient's residence at the date of diagnosis.	2354	Longitude			Data units are in decimal degrees, not degrees, minutes and seconds. Data have been geocoded using specialized software by DSHS Center for Health Statistics, with up to 6 digits after the decimal.	X		X		
Physician--Managing	Texas Medical Board license number of the managing physician who is responsible for overall management of patient during diagnosis and/or treatment for this cancer.. NOTE: This variable is often not populated.	2460	PhysManaging	2006	2009	one capital letter followed by four numbers; 99999999=unknown. Lookup for license numbers available at <a href="http://www.tmb.state.tx.us/page/look-up-a-license">http://www.tmb.state.tx.us/page/look-up-a-license</a>	X		X		
Physician--Follow-Up	Texas Medical Board license number of the follow-up physician who is currently responsible for the patient's medical care.	2470	PhysFup	2006	present	one capital letter followed by four numbers; 99999999=unknown. Lookup for license numbers available at <a href="http://www.tmb.state.tx.us/page/look-up-a-license">http://www.tmb.state.tx.us/page/look-up-a-license</a>	X		X		
Comorbid/Complication 1	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3110	ComorbidComp1	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X		
Comorbid/Complication 2	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3120	ComorbidComp2	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X		
Comorbid/Complication 3	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3130	ComorbidComp3	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X		
Comorbid/Complication 4	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3140	ComorbidComp4	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X		
Comorbid/Complication 5	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3150	ComorbidComp5	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X		

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
Comorbid/Complication 6	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3160	ComorbidComp6	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X	
Comorbid/Complication 7	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3161	ComorbidComp7	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X	
Comorbid/Complication 8	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3162	ComorbidComp8	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X	
Comorbid/Complication 9	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3163	ComorbidComp9	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X	
Comorbid/Complication 10	ICD-9-CM code for patient comorbidity or complication present at time of diagnosis; up to 10 collected. Began collecting in 2011 and data completeness and quality are being evaluated.	3164	ComorbidComp10	2011	present	ICD-9-CM classification. See <a href="http://www.cdc.gov/nchs/icd/icd9cm.htm">http://www.cdc.gov/nchs/icd/icd9cm.htm</a> .			X	
Rural-Urban Continuum/Beale Code 1993	1993 Beale code based on county of residence on diagnosis date	3300	RuralUrbanCon1993			Codes 0-3 = Metro, Codes 4-9 = Nonmetro. For more information on codes, consult <a href="http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx">http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx</a>		X	X	X X
Rural-Urban Continuum/Beale Code 2003	2003 Beale code based on county of residence on diagnosis date	3310	RuralUrbanCon2003			Codes 0-3 = Metro, Codes 4-9 = Nonmetro. For more information on codes, consult <a href="http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx">http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx</a>		X	X	X X
Height	Patient height in inches on date of dx. Began collecting in 2011 and data completeness and quality are being evaluated.	9960	Height	2011		2-digit numbers measured in inches (1 foot=12 inches). Code '98' for 98 inches or greater. Code '99' for unknown height.			X	
Weight	Patient weight in pounds on date of dx. Began collecting in 2011 and data completeness and quality are being evaluated.	9961	Weight	2011		3-digit numbers measured in pounds (1 kg=2.2 pounds). Code '999' for unknown weight.			X	

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
Tobacco Use Cigarettes	Indicator of patient cigarette smoking on date of dx. Began collecting in 2011 and data completeness and quality are being evaluated.	9965	TobaccoUseCig	2011		0=never used; 1=current user on date of dx; 2=former user, quit within 1 year of date of dx; 3=former user, quit >1 year prior to date of dx; 4=former user; unknown quit date; 9=unknown, not stated, or no specifics provided.			X	
Tobacco Use Other Smoke	Indicator of patient use of non-cigarette tobacco smoking products (e.g. pipes, cigars, kreteks) on date of dx. Began collecting in 2011 and data completeness and quality are being evaluated.	9966	TobaccoUseOther	2011		0=never used; 1=current user on date of dx; 2=former user, quit within 1 year of date of dx; 3=former user, quit >1 year prior to date of dx; 4=former user; unknown quit date; 9=unknown, not stated, or no specifics provided.			X	
Tobacco Use Smokeless	Indicator of patient use of smokeless tobacco products (e.g. chewing tobacco, snuff, etc.) on date of dx. Began collecting in 2011 and data completeness and quality are being evaluated.	9967	TobaccoUseSmokels	2011		0=never used; 1=current user on date of dx; 2=former user, quit within 1 year of date of dx; 3=former user, quit >1 year prior to date of dx; 4=former user; unknown quit date; 9=unknown, not stated, or no specifics provided.			X	
Tobacco Use NOS	Indicator of tobacco use, Not Otherwise Specified, on date of dx. Began collecting in 2011 and data completeness and quality are being evaluated.	9968	TobaccoUseNOS	2011		0=never used; 1=current user on date of dx; 2=former user, quit within 1 year of date of dx; 3=former user, quit >1 year prior to date of dx; 4=former user; unknown quit date; 9=unknown, not stated, or no specifics provided.			X	
Source Comorbidity	Data source from which comorbidities/ complications were collected. Began collecting comorbidities in 2011 and data completeness and quality are being evaluated.	9970	SourceComorbidity	2011		0=no comorbid condition or complication identified/Not Applicable; 1=facility face sheet; 2=linkage to facility/hospital discharge data set; 3=linkage to Medicare/Medicaid data set; 4=linkage with another claims data set; 5=combination of 2 or more sources above; 6=other source			X	
EDP MDE Link Variable	Indicates whether the tumor record linked to a tumor recorded in the DSHS Breast and Cervical Cancer Services (BCCS) program database (funded by federal NBCCEDP program)	9980	MDELink	2011		0=record sent for linkage, not match for this cancer with Texas BCCP data; 1=record sent for linkage, match for this cancer with Texas BCCP; blank=record not sent for linkage or linkage result pending.		X	X	

		NAACCR Version 13				Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
EDP MDE Link Date	Date the tumor record was linked to a tumor record recorded in the DSHS Breast and Cervical Cancer Services (BCCS) program database (funded by federal NBCCEDP program)	9981	MDDate	2011		YYYYMMDD if record linked; blank if it did not link or linkage result pending.		X		
MedRefID	Sequential numeric identification number assigned to each tumor in the TCR database; unique to each tumor.	10015	MedRefID				X	X	X	X
GEO_UNIQUE	Unique numeric identifier assigned to each tumor in the registry through May 2012. Obsolete in Registry Plus.	N/A	SC_GeoUniq					X		X
AYA site recode	This is a text variable with site names printed out. A site recode variable for cancers in adolescents and young adults (ages 15-39) that is derived from the variables Primary Site, Histologic type (ICD-O-3) and Behavior code (ICD-O-3).	SEER*Stat recode				See <a href="http://seer.cancer.gov/ayarecode/index.html">http://seer.cancer.gov/ayarecode/index.html</a> for SEER variable definition table.	X		X	
AYA site recode/WHO 2008	Uses SEER*stat coding scheme as shown: <a href="http://seer.cancer.gov/ayarecode/aya-who2008.html">http://seer.cancer.gov/ayarecode/aya-who2008.html</a> These codes are updated for Hematopoietic codes based on WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008). These codes are subject to change based on evolving ICD-O-3 coding rules.	SEER*Stat recode					X		X	
Day of diagnosis		SEER*Stat recode					X			
ICCC site rec extended ICD-O-3/WHO 2008	Uses SEER*stat coding scheme as shown: <a href="http://seer.cancer.gov/iccc/iccc-who2008.html">http://seer.cancer.gov/iccc/iccc-who2008.html</a> These codes are updated for Hematopoietic codes based on WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008). These codes are subject to change based on evolving ICD-O-3 coding rules. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, Third Edition. Cancer 2005;103:1457-67.	SEER*Stat recode					X		X	

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
ICCC site recode extended ICD-O-3	For publications, the International Classification of Childhood Cancer has been defined based on primary site and ICD-O-3 morphology.	SEER*Stat recode				See <a href="http://seer.cancer.gov/ICCC/iccc3_ext.html">http://seer.cancer.gov/ICCC/iccc3_ext.html</a> for ICC-3 main site recode definitions.		X		X
ICCC site recode ICD-O-3	For publications, the International Classification of Childhood Cancer has been defined based on primary site and ICD-O-3 morphology.	SEER*Stat recode				See <a href="http://seer.cancer.gov/iccc/iccc3.html">http://seer.cancer.gov/iccc/iccc3.html</a> for ICC-3 main site recode definitions.		X		X
ICCC site recode ICD-O-3/WHO 2008	Uses SEER*stat coding scheme as shown: <a href="http://seer.cancer.gov/iccc/iccc-who2008.html">http://seer.cancer.gov/iccc/iccc-who2008.html</a> These codes are updated for Hematopoietic codes based on WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008). These codes are subject to change based on evolving ICD-O-3 coding rules.	SEER*Stat recode						X		X
ICD-O-3 Hist/behavior, labeled	This includes each ICD-O-3 Histology code including the behavior code (/code), and the respective name of that histology and behavior. See the International Classification of Diseases for Oncology, Third edition (ICD-O-3) for histology and and behavior codes.	SEER*Stat recode				This is a text variable with histology and behavior names printed out. Ordered by individual ICD-O-3-codes, with label.		X		X
ICD-O-3 Hist/behavior, malig, labeled	This includes each ICD-O-3 Histology code including the behavior code (/code), and the respective name of that histology and behavior. See the International Classification of Diseases for Oncology, Third edition (ICD-O-3) for histology and and behavior codes.	SEER*Stat recode				This is a text variable with histology and behavior names printed out. Ordered by individual ICD-O-3-codes and labels only for malignant neoplasms (behavior code /3).		X		X

NAACCR Version 13							Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Lymphoma subtype recode	International Lymphoma Epidemiology Consortium (InterLymph) Pathology Working Group published a proposed nested classification of lymphoid neoplasms1 based on the World Health Organization classification of lymphoid neoplasms and the International Classification of Diseases-Oncology, Third Edition (ICD-O-3). This classification scheme was designed to facilitate epidemiologic studies of lymphoma subtypes. See: <a href="http://seer.cancer.gov/lymphomarecode/">http://seer.cancer.gov/lymphomarecode/</a>	SEER*Stat recode						X		X	
Lymphoma subtype recode/WHO 2008	Uses SEER*stat coding scheme as shown: <a href="http://seer.cancer.gov/lymphomarecode/lymphoma-who2008.html">http://seer.cancer.gov/lymphomarecode/lymphoma-who2008.html</a> for Lymphoma subtypes. These codes are updated for Hematopoietic codes based on WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (2008). These codes are subject to change based on evolving ICD-O-3 coding rules.	SEER*Stat recode						X		X	
Month Initiated RX--SEER		SEER*Stat recode						X		X	
Month of diagnosis		SEER*Stat recode						X		X	
Month of diagnosis recode	Derived field used internally by SEER*Stat for calculating survival time, and should not be used for generating true month of diagnosis. It generates a month of diagnosis when month of diagnosis is unknown.	SEER*Stat recode						X		X	
Primary Site - labeled	This is a text variable with site names listed. Ordered by individual ICD-O-3 site codes to 1 decimal place, with site name listed.	SEER*Stat recode				Primary site (NAACCR item [400]) with text label added.		X		X	
Site rec with Kaposi and mesothelioma	This is a text variable with site names printed out. New site recode using ICD-O-3, including Kaposi sarcoma and mesothelioma.	SEER*Stat recode				See <a href="http://seer.cancer.gov/siterecode/icdo3_d01272003/">http://seer.cancer.gov/siterecode/icdo3_d01272003/</a> for SEER Site recode with Kaposi sarcoma and mesothelioma definitions.		X		X	

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Site recode	This is a text variable with site names listed. Old site recode using ICD-O-2.	SEER*Stat recode				See <a href="http://seer.cancer.gov/siterecode/icdo2_d01272003/">http://seer.cancer.gov/siterecode/icdo2_d01272003/</a> for SEER Site Recode definitions.		X		X	
State-county	This item is a state-county FIPS code combination.	SEER*Stat recode				The first two digits represent the state FIPS code. The last two digits represent the FIPS county code.		X		X	
Year Initial RX--SEER		SEER*Stat recode						X		X	
Council of Government	COG variable: Texas Council of Government geographic area where the patient lived at time of diagnosis. Created from the counties included in each COG, and in SEER*Stat may be used to generate rates by COG.	TCR only				00-24, each representing a Texas geographic region. Coded as COG1-COG24 in SEER*Stat file. See <a href="http://www.txregionalcouncil.org/display.php?page=regions_map.php">http://www.txregionalcouncil.org/display.php?page=regions_map.php</a> .		X	X	X	X
Public Health Region	HSR variable: Texas Health Service Region (HSR) where the patient lived at time of diagnosis. Defined by the counties included in each of the 11 Health Services Regions. In SEER*Stat may be used to generate rates by Health Service Region.	TCR only				01-11, each representing a Texas geographic region. Coded as Health Service Region 1 thru Health Service Region 11 in SEER*Stat file. See <a href="http://www.dshs.state.tx.us/regions/">http://www.dshs.state.tx.us/regions/</a>		X	X	X	X
Site recode ICD-O-3/WHO 2008	Numeric 5-digit cancer site recode, per SEER. These codes are subject to change based on evolving ICD-O-3 coding rules.	TCR only	N/A			Codes listed at <a href="http://seer.cancer.gov/siterecode/icdo3_dwhohome/index.html">http://seer.cancer.gov/siterecode/icdo3_dwhohome/index.html</a>		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
<b>Stage and Prognostic Factors</b>											
SEER Summary Stage 2000	The degree of advancement of a malignant tumor. For date of diagnosis 2001-2003. SEER Summary Staging, per 2000 Staging Manual: <a href="http://seer.cancer.gov/tools/ssm/">http://seer.cancer.gov/tools/ssm/</a>	759	SEERSumStg2000	2001	2003	Coded only for 2001-2004 cases. 0=In situ; 1=Localized; 2=Regional by direct extension only; 3=Regional to regional lymph nodes only; 4=Regional (direct extension and regional lymph nodes); 5=Regional, NOS; 7=Distant metastasis or systemic disease (leukemia, multiple myeloma); 8=Not applicable; 9=Unstaged, Unknown, Unspecified		X	X	X	X
SEER Summary Stage 1977	The degree of advancement of a malignant tumor. For date of diagnosis 1995-2000. SEER Summary Staging, per 1977 Staging Manual: <a href="http://seer.cancer.gov/manuals/historic/ssm_1977.pdf">http://seer.cancer.gov/manuals/historic/ssm_1977.pdf</a>	760	SEERSumStg1977	1995	2000	0=In situ; 1=Localized; 2=Regional by direct extension only; 3=Regional by regional lymph nodes only; 4=Regional (direct extension and lymph nodes); 5=Regional, NOS; 7=Distant metastasis or systemic disease (leukemia, multiple myeloma); 8=Not applicable; 9=Unstaged, Unknown, Unspecified		X	X	X	X
EOD--Tumor Size	Tumor size prior to 2004. Largest dimension or the diameter of the primary tumor recorded in <i>millimeters</i> .	780	EODTumSize	1998	2003	Number in millimeters with leading zeros. Only available for 1998-2003 (variable name 'EODTumSize_r')			X		X
Regional Nodes Positive	Number of regional lymph nodes that were positive for cancer.	820	RegNodPos	1998	present	Number with leading zeros. Only collected for colorectal cancer.			X		X
Regional Nodes Examined	Number of regional lymph nodes that were examined.	830	RegNodExam	1998	present	Number with leading zeros. Only collected for colorectal cancer.			X		X
CS Tumor Size	Records the largest dimension of the reported tumor in <i>millimeter</i> .	2800	CSTumorSize	2004	present	Size in millimeters with leading zeros. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X
CS Extension	Identifies contiguous growth (extension) of the primary tumor within the organ of origin or extension into neighboring organs. For certain sites such as ovary, discontinuous metastasis is coded in CS Extension	2810	CSExt	2004	present	See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X
CS Tumor Size/Ext Eval	Records how the codes for the two items CS Tumor Size [2800] and CS Extension [2810] were determined, based on the diagnostic methods employed.	2820	CSSizeExtEval	2008		See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
CS Lymph Nodes	Identifies regional lymph nodes involved.	2830	CSLymphNodes	2004	present	See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X
CS Lymph Nodes Eval	Records how the code for CS Lymph Nodes [2830] was determined, based on the diagnostic methods employed.	2840	CSNodesEval	2011		See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X
CS Mets at DX	Identifies the distant site(s) of metastatic involvement at diagnosis	2850	CSMetsDX	2004	present	See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X
CS Mets Eval	Records how the code for CS Mets at Dx [2850] was determined based on the diagnostic methods employed.	2860	CSMetsEval	2011		See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.		X	X	X	X
CS Site-Specific Factor 15	For breast (C50). Identifies additional information needed to generate stage, or prognostic factors that have an effect on stage or survival. NOTE: Some American College of Surgeons facilities began providing this for cases diagnosed in 2010; TCR required it beginning with cases diagnosed in 2011.	2869	CSSSF15	2011	present	The information recorded differs for each anatomic site. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.			X		X
CS Site-Specific Factor 25	Identifies additional information needed to generate stage, or prognostic factors that have an effect on stage or survival. Used to discriminate between CS staging schema or between AJCC chapters where site and histology alone are insufficient to identify the tumor type or location to identify the applicable staging method. Collected 2004-present; required 2010 & later as a Schema Discriminator.	2879	CSSSF25	2004	present	The information recorded differs for each anatomic site. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.			X		X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
CS Site-Specific Factor 1 for: C50 (breast) C70.0-C70.9, C71.0-C71.9, C72.0-C72.9, C75.1- C75.3 (brain, CNS) C34 (lung) C384 (pleura) C692 with 9510/3, 9511/3, 9512/3, 9513/3 (retinoblastoma)	Identifies additional information needed to generate stage, or prognostic factors that have an effect on stage or survival. Collected 2004+ for pleura C38.4 only; 2010+ for breast, lung, retinoblastoma; 2011+ for brain/CNS.	2880	CSSSF1	2004	present	The information recorded differs for each anatomic site. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.			X		X
CS Site-Specific Factor 2 for: C50 (breast) C54, C55 (corpus uteri)	Identifies additional information needed to generate stage, or prognostic factors that have an effect on stage or survival.	2890	CSSSF2	2010	present	The information recorded differs for each anatomic site. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures.			X		X
CS Site-Specific Factor 3 for: C619 (prostate)	Identifies additional information needed to generate stage, or prognostic factors that have an effect on stage or survival.	2900	CSSSF3	2004	present	The information recorded differs for each anatomic site. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures. Collected 2004-present for prostate only.			X		X
CS Version Input Original	Indicates the number of the version initially used to code Collaborative Staging (CS) fields. The CS version number is returned as part of the output of the CS algorithm. Because input codes and CS instructions may change over time, this identifies the correct interpretation of input CS items.	2935	CSVerInputOrig			6-digit code (e.g., 010100). First two digits represent the major version number; the 2nd two digits represent minor version changes; and, the last two digits represent even less significant changes that do not affect coding or derivation of results See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures. .		X	X	X	X

		NAACCR Version 13									
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files	Confidential? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
CS Version Derived	Recorded the first time the CS output fields are derived and is updated each time CS Derived items are recomputed. The CS version number is returned as part of the output of the CS algorithm.	2936	CSVerDerived			6-digit code (e.g., 010100). First two digits represent the major version number; the 2nd two digits represent minor version changes; and, the last two digits represent even less significant changes that do not affect coding or derivation of results Should not be blank of CS Derived items contain values; should be blank of CS Derived items are empty or CS algorithm has not been applied. See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures. .		X	X	X	X
CS Version Input Current	Indicates the version of CS input fields after updating or recoding. Recorded the first time the CS input fields are entered and updated each time CS input fields are modified. Because input codes and CS instructions may change over time, this identifies the correct interpretation of input CS items.	2937	CSVerInputCurrent			6-digit code (e.g., 010100). First two digits represent the major version number; the 2nd two digits represent minor version changes; and, the last two digits represent even less significant changes that do not affect coding or derivation of results See the most current version of the "Collaborative Stage Data Collection System" ( <a href="http://cancerstaging.org">http://cancerstaging.org</a> ) for rules and site-specific codes and coding structures. .		X	X	X	X
Derived SS2000	The degree of advancement of a malignant tumor. For date of diagnosis 2004 and later. This is the derived SEER Summary Stage 2000 from the Collaborative Staging algorithm: <a href="https://cancerstaging.org/cstage">https://cancerstaging.org/cstage</a>	3020	DerivedSS2000	2004	present	0/IS=In situ; 1/L=Localized; 2/RE=Regional by direct extension; 3/RN=Regional to lymph nodes; 4/RE+RN=Regional (direct extension and lymph nodes); 5/RNOS=Regional, NOS; 7/D=Distant metastasis or systemic disease (leukemia, multiple myeloma); 8/NA=Not applicable; 9/U=Unstaged, Unknown, Unspecified		X	X	X	X
Derived SS2000--Flag	Flag to indicate whether the derived SEER Summary Stage 2000 was derived from CS or EOD codes.	3050	DerivedSS2000Flag	2004	present	1=SS2000 derived from Collaborative Stage; 2=SS2000 derived from EOD (prior to 2004); blank=not derived		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Summary Stage (Combined)	Combines the various Summary Stage coding schemes in effect across time (SEER Summary Stage 1977 in effect for 1995-2000 diagnoses, SEER Summary Stage 2000 in effect for 2001-2003 diagnoses, and Derived SS 2000 in effect for diagnoses 2004 and forward).	SEER*Stat User Defined Merged				In situ; Localized; Regional; Distant; Unknown Stage; Blank(s)		X		X	
<b>Treatment - First Course</b>											
Date Initial RX SEER	Date of initiation of first course of therapy. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1260	FirstRxDateSEER	1995	present	YYYYMMDD Format; 99999999=Unknown if any treatment; 00000000=No therapy. In SEER*Stat this is stored as Year of Initial Rx--SEER; and a second variable Month of Initial Rx--SEER.		X (as month and year)	X	X (as month and year)	X (as month and year)
Date Initial RX SEER Flag	If there is no appropriate value in the Date of Initial Rx--SEER [1260] field, this code explains why.	1261	FirstRxDateSEERFlag	2010		10=No information whatsoever can be inferred from this exceptional value; 11=No proper value is applicable in this context; 12=A proper value is applicable but not known; blank=a valid date value is stored in Date of Initial Rx--SEER [1260], or date was not expected to have been transmitted.		X	X	X	X
Rx Summ--Surg Prim Site	Type of surgery performed as first course of treatment. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1290	RxSumSurgPSite	1995	present	00=None; 10-19=Site-specific code (tumor destruction); 20-80=Site-specific codes (resection); 90=Surgery, NOS; 98=Site-specific codes (special); 99=Unknown. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail.		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Rx Summ--Scope Reg LN Sur	Scope of regional lymph node surgery. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1292	RxSumScopeRegLN	2001	present	0=None; 1=Biopsy or aspiration of regional lymph node, NOS; 2=Sentinel lymph node biopsy; 3=# of regional lymph nodes removed unknown; 4=1 to 3 regional lymph nodes removed; 5=4 or more regional lymph nodes removed; 6=sentinel node bopsy and code 3, 4, or 5 at same time or timing not noted; 7=sentinel node biopsy and code 3, 4, or 5 at different times; 9=unknown or not applicable.		X	X	X	X
Rx Summ--Surg Oth Reg/Dis	Records surgical removal of distant lymph nodes or other tissue(s)/organ(s) beyond the primary site. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1294	RxSumSurgOthReg	1998	present	0=None; diagnosed at autopsy; 1=non-primary surgical procedure; 2=non-primary surgical procedure to other regional sites; 3=non-primary surgical procedure to distant lymph node(s); 4=non-primary surgical procedure to distant site; 5=any combination of codes 2, 3 or 4; 9=unknown; death certificate only		X	X	X	X
Reason for No Surgery	Reason surgery was not performed on the primary site. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1340	ReasonNoSurg	1998 2006	2002 present	1 of 8 possible codes, 0-2, 5-9. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
RX Summ--Radiation	Codes for the type of radiation therapy performed as part of first course of treatment.. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1360	RxSumRad	1998 2012	2002 present	Only available for 1998-2003 (variane name "RxSumRad_r"). Codes: 0=none; 1=beam radiation; 2=radioactive implants; 3=radioisotopes; 4=combination of 1 with 2 or 3; 5=radiation, NOS; 6=only for historical cases; 7=patient or their guardian refused; 8=radiation recommended, unknown if administered; 9=unknown if radiation administered		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
RX Summ--Surg/Rad Seq	Indicates sequencing of radiation and surgery given as part of first course of treatment. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1380	RxSumSurgRadSeq	2004	present	0=no radiation and/or surgery, or unknown if given; 2=radiation before surgery; 3=radiation after surgery; 4=radiation both before and after surgery; 5=intraoperative radiation; 6=intraoperative radiation with other radiation given before and/or after the surgery; 7=surgery both before and after radiation; 9=sequence unknown, but both surgery and radiation given		X	X	X	X
Rx Summ--Chemo	Codes for chemotherapy given as part of first course of treatment or reason chemotherapy was not given.. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1390	RxSumChemo	1995	present	Codes 00-03, 82, 85-88, 99. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
Rx Summ--Hormone	Records whether systemic hormonal agents were administered as first-course treatment at any facility, or the reason they were not given. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1400	RxSumHorm	1995	present	Codes 00-01, 82, 85-88, 99. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Rx Summ--BRM	Records whether immunotherapeutic (biologic response modifiers) agents were administered as first-course treatment at all facilities or the reason they were not given.. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1410	RxSumBRM	1995	present	Codes 00-01, 82, 85-88, 99. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
Rx Summ--Other	Identifies other treatment given at all facilities that cannot be defined as surgery, radiation, or systemic therapy. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1420	RxSumOth	1995	present	0=none; 1=other; 2=other experimental; 3=other - double-blind; 4=other - unproven; 7=refusal; 8=recommended; 9=unknown; unkwno if administered. See NAACCR Data Standards for Cancer Registries, Version 13 for additional information: ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
Reason for No Radiation	Reason for no regional radiation therapy as part of first-course treatment. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1430	ReasonNoRad	1998 2011	2002 present	Codes 0-2, 5-9. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions: ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Rad--Regional Rx Modality	Records dominant modality of radiation therapy used to deliver the clinically most significant regional dose to the primary volume of interest during first course of treatment. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1570	RadRegModal	2003	present	Codes 00, 20-32, 40-43, 50-55, 60-62, 80, 85, 98-99. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions: ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
RX Summ--Systemic/Sur Seq	Records sequencing of systemic therapy. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	1639	RxSumSysSurSeq	2006		0=No systemic therapy and/or surgical procedures; unknown if surgery and/or systemic therapy given; 2=systemic therapy before surgery; 3=systemic therapy after surgery; 4=systemic therapy both before and after surgery; 5=intraoperative systemic therapy; 6=intraoperative systemic therapy with other therapy administered before and/or after surgery; 7=surgery both before and after systemic therapy; 8=sequence unknown, but both surgery and systemic therapy given.		X	X	X	X
Rx Summ--Transplant/Endocr	Identifies systemic therapeutic hematologic, transplant, and endocrine procedures at first-course treatment, or reason they were not performed. NOTE: For all treatment variables, only coding accuracy has been assessed; these data most often represent the first treatment data available, may not represent complete treatment, and are often blank. Great care should be exercised when analyzing and/or interpreting these data.	3250	RxSumTransEndo	2003	present	Codes 00, 10-12, 20, 30, 40, 82, 85-88, 99. See NAACCR Data Standards for Cancer Registries, Version 13 for code definitions ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X

		NAACCR Version 13				Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
<b>Follow-Up</b>											
Date of Last Contact	Date of last contact with the patient, or date of death if the patient has died. Researchers may request month and year of diagnosis without IRB approval.	1750	DateLastContact	1995	present	YYYYMMDD, no dashes or slashes. 99 or 9999=Unknown. In SEER*Stat this is Year of Last Contact which ranges from 1995-2011, or unknown; and a separate variable Month of Last Contact which ranges from January-December; Unknown.	X	X (as month and year)	X	X (as month and year)	X (as month and year)
Date of Last Contact Flag	If there is no appropriate value in Date of Last Contact [1750] field, this explains why.	1751	DateLastContFlag	2010		12=a proper value is applicable but not known. This event occurred, but the date is unknown (e.g. date of last contact is unknown); blank=A valid date value is provided in Date of Last Contact [1750], or the date was not expected to have been transmitted.		X	X	X	X
Vital Status	Vital status at last date of follow-up	1760	VitalStatus	1995	present	0=Dead; 1=Alive		X	X	X	X
Follow-Up Source Central	Records the source of Vital Status/Date of Death information	1791	FupSourceCentral	1995	present	Possible codes: 00 follow-up not performed for this patient; 01-29 file linkages; 30-39 hospitals and treatment facilities; 40-49 physicians; 50-59 patient; 60-98 other. See NAACCR Data Standards for Cancer Registries, Version 13 for additional detail.		X	X	X	X
Unusual Follow-Up Method	Indicator of whether the case was only reported by a Veterans Administration facility or another state's central cancer registry; or whether the person has requested they not be contacted by researchers. CRS Production database updated monthly with fresh calculation.	1850	FupUnusual			00=record was not reported to TCR solely by a VA or other state, and person has not opted out of being contacted for research studies; 01=record was reported only by another state's central cancer registry (OOS); 02=patient opted out of being contacted by researchers (HIPAA); 04=record was reported only by a Veterans Administration facility (VA)	X		X		

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
Cause of Death	Cause of death using ICD-9 (1995-1998) and ICD-10 (1999-forward) from Texas resident death certificate data	1910	DthCause	1995	present	For deaths in 1995-1998 use ICD-9; possible values are 0084-7718, 7806-9881; State death certificate not available; State death certificate available but COD not coded; Alive at last contact. For deaths in 1995+ use ICD-10; possible values are codes A021-Y883; State death certificate not available; State death certificate available but COD not coded; Alive at least contact. Also see SEER mortality recodes <a href="http://seer.cancer.gov/codrecode/1969+_d09172004/index.html">http://seer.cancer.gov/codrecode/1969+_d09172004/index.html</a>		X	X	X	X
ICD Revision Number	Indicator for the coding scheme used to code the cause of death.	1920	ICDRevNum	1995	present	0=Patient alive at last follow-up; 1=ICD-10; 7=ICD-7; 8=ICDA-8; 9=ICD_9		X	X	X	X
Place of Death	Patient's state of death from death certificate data. NOTE: This field is usually blank.	1940	DthPlace	1995	present	077=Texas; 997=Not applicable, patient alive; 999=Unknown. See NAACCR Data Standards for Cancer Registries, Version 13 Appendix B for geocodes ( <a href="http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#">http://naaccr.org/StandardsandRegistryOperations/Volumell.aspx#</a> )		X	X	X	X
DC State File Number	Death certificate number for the patient. Usually only available for Texas residents at the time of death.	2380	DCStateFileNum			N/A			X		
Cause of death (ICD-10)	Official cause of death as coded from death certificate in valid ICD-10 codes. (Deaths in 1999+).	SEER*Stat recode						X		X	
Cause of death (ICD-7, 8 or 9)	Official cause of death as coded from death certificate in valid ICD-9 codes. (Deaths in 1995-1998).	SEER*Stat recode						X		X	
Cause of death recode	The cause of death recode variable records the cause of death including both cancer and non-cancer causes. This is the official cause of death based on ICD 8-10. Does not include separate categories for Kaposi sarcoma and mesothelioma.	SEER*Stat recode				Text variable with each cancer site cause of death, and non-cancer causes of death. See <a href="http://seer.cancer.gov/codrecode/1969+_d09172004/">http://seer.cancer.gov/codrecode/1969+_d09172004/</a> for SEER cause of death recode.				X	

NAACCR Version 13						Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
COD recode with Kaposi and mesothelioma	The cause of death recode variable records the cause of death including both cancer and non-cancer causes. This is the official cause of death based on ICD 8-10. This version includes the separate categories for Kaposi sarcoma and mesothelioma.	SEER*Stat recode				Text variable with each cancer site cause of death, and non-cancer causes of death. See <a href="http://seer.cancer.gov/codrecode/1969+_d09172004/">http://seer.cancer.gov/codrecode/1969+_d09172004/</a> for SEER cause of death recode.			X	
Date of Last Contact Flag		SEER*Stat recode					X		X	
Day of last contact		SEER*Stat recode					X			
Month of last contact		SEER*Stat recode					X		X	
Month of last contact recode	Derived field used internally by SEER*Stat for calculating survival time, and should not be used for generating true month of last contact. It generates a month of last contact when month of last contact is unknown.	SEER*Stat recode					X		X	
SEER cause-specific death classification	This field is generated by the Texas Cancer Registry. Using a program from SEER we classify the vital status of the case for cause-specific survival analysis, by determining whether, based on the coded cause of death, this death would be considered to be caused by the cancer diagnosis.	SEER*Stat recode				Three values: Alive or dead of other cause; Dead; N/A not first tumor	X		X	
Vital Status (Study Cutoff used)		SEER*Stat recode					X		X	
Year of last contact	Derived field used internally by SEER*Stat for calculating survival time, and should not be used for generating true year of last contact. It generates a year of last contact appropriate for the study cutoff years, for example if the true year of last contact is beyond the study cutoff year.	SEER*Stat recode					X		X	
Year of last contact recode		SEER*Stat recode					X		X	

		NAACCR Version 13				Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file	
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files					
<b>Miscellaneous</b>											
Over-Ride Age/Site/Morph	Flag indicating that the record has been reviewed and, while, unusual, is correct	1990	OvrdAgeSiteMorph			1=reviewed & confirmed age/ site/ histology combo correct as reported; 2=reviewed & confirmed case diagnosed in utero; 3=reviewed & combined that conditions 1 & 2 apply; blank=not reviewed or reviewed and corrected.		X	X	X	X
Over-Ride SeqNo/DxConf	Flag indicating that the record has been reviewed and, while, unusual, is correct	2000	OvrdSqDxCnf			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Site/Lat/SeqNo	Flag indicating that the record has been reviewed and, while, unusual, is correct	2010	OvrdSitLatSq			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Surg/Dxconf	Flag indicating that the record has been reviewed and, while, unusual, is correct	2020	OvrdSurgDxCf			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Site/Type	Flag indicating that the record has been reviewed and, while, unusual, is correct	2030	OvrdSiteType			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Histology	Flag indicating that the record has been reviewed and, while, unusual, is correct	2040	OvrdHist			1=reviewed & confirmed pathologist states primary to be in situ or malignant although behavior code is benign or uncertain; 2=reviewed & confirmed behavior code is in situ but case not microscopically confirmed; 3=reviewed & confirmed conditions 1 & 2 both apply; blank=not reviewed or reviewed and corrected.		X	X	X	X
Over-Ride Report Source	Flag indicating that the record has been reviewed and, while, unusual, is correct	2050	OvrdRepSrc			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Ill-define Site	Flag indicating that the record has been reviewed and, while, unusual, is correct	2060	OvrdIldfSite			1=reviewed & confirmed as reported; a second or subsequent primary reported with ill-defined primary site has been reviewed & is an independent primary; blank=not reviewed or reviewed and corrected.		X	X	X	X
Over-Ride Leuk, Lymphoma	Flag indicating that the record has been reviewed and, while, unusual, is correct	2070	OvrdLeukLym			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Site/Behavior	Flag indicating that the record has been reviewed and, while, unusual, is correct	2071	OvrdSiteBeh			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X
Over-Ride Site/Lat/Morph	Flag indicating that the record has been reviewed and, while, unusual, is correct	2074	OvrdSiteLatMorph			blank=not reviewed or reviewed and corrected; 1=reviewed and confirmed as reported		X	X	X	X

		NAACCR Version 13				Confiden- tial? (yes=X)	In Confidential SEER*Stat incidence file	In Confidential SAS incidence file	In Limited Use SEER*Stat incidence file	In Limited Use SAS incidence file
Short name (NAACCR Description or SEER*Stat name)	Description in complete, edited epi analysis files	Item No, if applicable	SAS variable name (if applicable)	First year TCR collected	Last year TCR collected	Coding scheme in complete, edited epi analysis files				
Date Tumor Record Availabl	Date the demographic and tumor identification information on a primary/reportable neoplasm, compiled from one or more source records, from one or more facilities, is available in the central cancer registry database to be counted as an incident tumor. Cancer identification information includes, at a minimum, site, histology, laterality, behavior, and date of diagnosis. ALWAYS USED TO DETERMINE CUT-OFF DATE.	2113	DtTumorRecAvail			YYYYMMDD.		X		
ICD-O-3 Conversion Flag	Specifies how the conversion of ICD-O-2 to ICD-O-3 was accomplished	2116	ICDO3ConvFlag			0=Originally coded ICD-O-3; 1=Converted, no review; 2=Converted, review	X	X	X	X
SEER Type of Follow-Up	Codes for the type of follow-up expected for a SEER case.	2180	SEERTypeFup			1=Autopsy-Only or Death-Certificate-Only case; 2=Active follow-up case; 3=in situ cancer of the cervix uteri only. (In SEER*Stat files, always coded as '2' (non-SEER Registry) to facilitate survival calculations.)	X		X	
Over-ride CS 20	Flag to identify cases directly coded using SEER Summary Stage 2000 [759].	3769	OvrdCS20			1=Directly coded SEER Summary Stage 2000 [759] used to report Summary Stage; blank=Derived Summary Stage 2000 [3020] reported using Collaborative State Data Collection System or case diagnosed prior to 2012.	X	X	X	X

## Technical Notes

### Source of Data

The Texas Cancer Registry (TCR) is a population based cancer surveillance (reporting) system that includes incident reports of certain benign, borderline, in-situ, and malignant neoplasms occurring in Texas among state residents. The TCR was first established in 1979, but statewide, population-based reporting of newly diagnosed cancer cases was not fully implemented until 1995. Regional offices cover the entire state and assist with data collection and record processing.

Texas hospitals and cancer treatment centers are the primary sources of case reporting. Reports also are received from outpatient clinics, free-standing pathology labs, and other state central cancer registries when a Texas resident is diagnosed or treated at a facility outside of Texas. The data provided in these reports were primarily abstracted from medical records and pathology reports.

Population data included in SEER\*Stat files are based on the U.S. Census Bureau's Population Estimates Program, and modified by the National Cancer Institute using the special processing procedures for counties affected by Hurricanes Rita and Katrina. This is described on the SEER website: <http://seer.cancer.gov/popdata/>. The population data by county, age, sex, race, and ethnicity are determined by the U.S. Census, and modified by the SEER program for intercensal years for use by state cancer registries. Users should be aware that there are different estimates of the Texas population available from other organizations, but that, like other cancer registries throughout the United States, we use the official estimates by the SEER program to calculate cancer incidence and mortality rates.

### Confidentiality

Protecting the confidentiality of persons whose cancers are reported to the TCR is the highest priority of the Registry in all aspects of operations, and required by state law and rule (Health and Safety Code, §82.009; Texas Administrative Code, Title 25, Part 1, Chapter 91, Subchapter A).

***No data presented in these data files are intended to be used to identify individuals who have been diagnosed with cancer.***

### Case Definition

A "case" is a primary cancer, and the anatomic site recorded is the site of tumor origin. Individuals can have more than one primary cancer and each primary cancer counts as a case. Additional tumors resulting from metastasis, or spread of cancer from the original tumor site to other organs, are not counted as cases.

### **Classification by Anatomic Site**

Primary anatomic site and histologic type were coded for each cancer incident case using the International Classification of Diseases for Oncology (ICD-O). For cases diagnosed from 1995–2000, the second edition was used (ICD-O-2)<sup>1</sup> and cases were then recoded to ICD-O-3 for analysis. For cases diagnosed from 2001–2011, the third edition was used (ICD-O-3).<sup>2</sup> Cases were then recoded into SEER program site recode groups for classifying types of cancer, using SeerPrep version 2.5.2 software. The SEER site recodes can be found at [http://seer.cancer.gov/siterecode/icdo3\\_d01272003/](http://seer.cancer.gov/siterecode/icdo3_d01272003/).

### **Classification by Race and Ethnicity**

Race and ethnicity information for cancer cases is based primarily on information contained in the patient’s medical record. This information may be supplied directly by the patient, may be determined by admissions staff or other medical personnel, and/or can be based on last name, race and ethnicity of parents, birthplace, or maiden name. The reporting of race and ethnicity may be influenced by the race and ethnic distribution of the local population, by local interpretation of data collection guidelines, and other factors.

The race and ethnicity of each cancer patient is classified according to the categories defined in the North American Association of Central Cancer Registries (NAACCR) coding manual.<sup>3</sup> The race and ethnic groups most often used by the TCR for generating incidence and mortality rates include the following categories: white, black, American Indian/Alaskan Native, Asian or Pacific Islander, and Hispanics of any race (a double-count).

The Hispanic designation can therefore be of any race, but in 1995–2011, 98.8 percent of Texas Hispanics diagnosed with cancer were of the white race. Unless persons of unknown race are coded as Hispanic (only 0.7% of all persons with cancer in 1995-2009), they are not included in any of the race or ethnic-specific categories, but are included in All Races. Therefore, the race ethnic sub-categories will not sum exactly to the total for All Races.

We used the field “Race1” to identify American Indians and Alaska Natives. However, we also added a variable to identify all persons who were linked from our cancer database with the U.S. Indian Health Service Inpatient database, since only persons recognized as American Indian/Alaska Native are included in that database.

### **Data Quality**

The Texas Cancer Registry employs multiple procedures to assure the quality of incoming data, and these are described in the Texas Cancer Registry Cancer Reporting Handbook, distributed to all cancer reporters in the state.<sup>4</sup> The Texas Cancer Registry data currently meet Centers for Disease Control and Prevention national high quality data standards, as

well as gold certification from the North American Association of Central Cancer Registries. Numerous quality assurance procedures were applied to the data based on SEER, NPCR, NAACCR, and TCR standards. Quality control included both internal and external processes to insure the reliability, completeness, consistency, and comparability of TCR data. Examples of internal consolidation and quality assurance processes include 1) a review of multiple abstracts on the same patient for multiple primaries, 2) identifying possible duplicate records, 3) correcting unacceptable codes or inter-field inconsistencies, and 4) reviewing unusual code combinations for site/sex, age/site, age/morphology or site/morphology. Inconsistencies and unknown values for date of birth, race, ethnicity, sex, county of residence, date of diagnosis, site, and histologic type were rectified to the greatest extent possible. External procedures included training of reporting facility staff, on-site case-finding, and re-abstracting studies.

Cancer death certificate files were also matched against reported incident cases for an additional reporting completeness check. To further assist in identifying any cancer cases not reported to the TCR, information on all death certificates with the underlying cause of death due to a malignant neoplasm were obtained from the Texas Department of State Health Services, Center for Health Statistics. Institutions listed on the death certificates as the place of death were queried for additional cancer case information. Missed cases not identified from any institution were added to the cancer database as “death certificate only” (DCO) cases. These DCO cases for which the only available information is from the death certificate, are included in these data and can be identified by the field “type of reporting source.”

---

#### References

<sup>1</sup> Percy D, Van Holten V, and Muir C (eds). International Classification of Diseases for Oncology, Second Edition, Geneva: World Health Organization, 1990.

<sup>2</sup> Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin D, Whelan S (eds). International Classification of Diseases for Oncology, Third Edition, Geneva: World Health Organization, 2000.

<sup>3</sup> Thorton, Monica, editor. *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary, Sixteenth Edition, Record Layout Version 13*. Springfield, IL: North American Association of Central Cancer Registries, implemented January 1, 2013, available at <http://www.naaccr.org/StandardsandRegistryOperations/VolumellArchive.aspx>.

<sup>4</sup> Texas Cancer Registry, Texas Department of State Health Services, 2013. Cancer Reporting Handbook. Available online <http://www.dshs.state.tx.us/tcr/CancerReporting/2013-Cancer-Reporting-Handbook.aspx>. Publication Number E10-10677.