No	Features Names	Field Description
1.	Patient ID	This field is used in conjunction with SEER registry to uniquely identify a person.
2.	SEER Registry	A unique code assigned to each participating SEER registry.
3.	Marital status at diagnosis	This data item identifies the patient's marital status at the time of diagnosis for the reportable tumor.
4.	Race/ethnicity	Recode which gives priority to non-white races for persons of mixed races. Note that not all codes were in effect for all years.
5.	Spanish surname or origin	This data item is used to identify patients with Spanish/Hispanic surname or of Spanish origin. Persons of Spanish or Hispanic surname/origin may be of any race.
6.	NHIA Derived Hispanic Origin	The NAACCR Hispanic Identification Algorithm (NHIA) is a computerized algorithm that uses a combination of variables to directly or indirectly classify cases as Hispanic for analytic purposes.
7.	Gender	This data item identifies the sex of the patient at diagnosis.
8.	Age at diagnosis	This data item represents the age of the patient at diagnosis for this cancer. The code is three digits and represents the patient's actual age in years.
9.	Age at diagnosis	This data item represents the age of the patient at diagnosis for this cancer. The code is three digits and represents the patient's actual age in years.
10.	Year of birth	The SEER dates on this file do not have the corresponding date flag included.

11.	Place of birth	See the following link for a list of SEER birthplace codes and definitions.
12.	Sequence number	Sequence Number-Central describes the number and sequence of all reportable malignant, in situ, benign, and borderline primary tumors, which occur over the lifetime of a patient.
13.	Month of diagnosis	The month of diagnosis is the month the tumor was first diagnosed by a recognized medical practitioner, whether clinically or microscopically confirmed. For analysis purposes, months coded to 99 (unknown) have been replaced with values 01 through 12.
14.	Year of diagnosis	The year of diagnosis is the year the tumor was first diagnosed by a recognized medical practitioner, whether clinically or microscopically confirmed.
15.	Primary Site	This data item identifies the site in which the primary tumor originated
16.	Laterality (1973+)	Laterality describes the side of a paired organ or side of the body on which the reportable tumor originated.
17.	Histology (92-00) ICD-O-2	NOTE: ALL CASES DIAGNOSED 1973-91 HAVE BEEN CONVERTED TO THE *INTERNATIONAL CLASSIFICATION OF DISEASES FOR ONCOLOGY*, SECOND EDITION (ICD-O-2, 1992).
18.	Behavior code ICD-O-2	Behavior codes are also defined in ICD-O-2, 1992.
19.	Histologic Type ICD-O-3	The data item Histologic Type describes the microscopic composition of cells and/or tissue for a specific primary.
20.	Behavior code ICD-	SEER requires registries to collect malignancies with

	O-3 (1973+)	in situ /2 and malignant /3 behavior codes as described in ICD-O-3.
21.	Grade	Grading and differentiation codes of 1-4, 9 are defined in ICD-O-2; 1992. Grade information may be incomplete for cases diagnosed before 1977.
22.	Diagnostic Confirmation	This data item records the best method used to confirm the presence of the cancer being reported.
23.	Type of Reporting Source	The Type of Reporting Source identifies the source documents used to abstract the case.
24.	EOD 10—size	This item is part of the 10-digit EOD (Item #779). It records the largest dimension of the primary tumor in millimeters.
25.	EOD 10—extent)	This item codes the farthest documented extension of tumor away from the primary site, either by contiguous extension or distant metastases. Allowable values = 00-99.
26.	EOD 10—Prostate path ext	This is an additional field for prostate cancer only to reflect information from radical prostatectomy
27.	EOD 10—nodes	This item records the highest specific lymph node chain that is involved by the tumor. Allowable values = 0-9.
28.	Regional nodes positive	Records the exact number of regional lymph nodes examined by the pathologist that were found to contain metastases.
29.	Regional nodes examined	Records the total number of regional lymph nodes that were removed and examined by the pathologist.
30.	Expanded EOD(1) – Expanded EOD	Detailed site-specific codes for EOD used by SEER for selected sites of cancer for tumors diagnosed

		(13)	1973-1982, except death-certificate-only cases.
31.	2-Digit NS EOD / 2-Digit SS EOD		Site-specific codes for EOD used by SEER for tumors diagnosed from January 1, 1973, to December 31, 1982, for cancer sites that did not have a 13-digit scheme.
32.	EOD 4 – size (first two digits)		Codes for site-specific EOD used by SEER for tumors diagnosed from January 1, 1983 to December 31, 1987 for all cancer sites.
33.	Coding system— EOD (1973-2003)		Indicates the type of SEER EOD code applied to the tumor.
34.	Tumor Marker 1 (1990-2003)		This data item records prognostic indicators for breast cases (ERA 1990-2003), prostate cases (PAP 1998-2003) and testis cases (AFP 1998-2003).
35.	Tumor Marker 2 (1990-2003)		This data item records prognostic indicators for breast cases (PRA 1990-2003), prostate cases (PSA 1998-2003), and testis cases (hCG 1998-2003).
36.	Tumor Marker 3 (1998-2003)		This data item records prognostic indicators for testis cases (LDH 1998-2003)
37.	CS tumor size (2004+)		Information on tumor size. Available for 2004+. Earlier cases may be converted and new codes added which weren't available for use prior to the current version of CS.
38.	CS extension (2004+)		Information on extension of the tumor. Available for 2004+. Earlier cases may be converted and new codes added which weren't available for use prior to the current version of CS.
39.	CS lymph nodes		Information on involvement of lymph nodes. Available for 2004+. Earlier cases may be converted

		and new codes added which weren't available for use prior to the current version of CS.
40.	CS mets at dx (2004+)	Information on distant metastasis. Available for 2004+. Earlier cases may be converted and new codes added which weren't available for use prior to the current version of CS.
41.	CS site-specific factor 1 (2004+)	Each CS site-specific factor (SSF) is schema dependent.
42.	CS site-specific factor 2 (2004+)	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case,
43.	CS site-specific factor 3 (2004+)	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case,
44.	CS site-specific factor 4 (2004+)	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information.
45.	CS site-specific factor 5 (2004+)	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information.
46.	CS site-specific factor 6 (2004+)	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information.
47.	CS site-specific factor 25 (2004+)	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to

		stage the case, clinically relevant information, or prognostic information.
48.	Derived AJCC T, 6th ed (2004+)	This is the AJCC "T" component that is derived from CS coded fields, using the CS algorithm, effective with 2004+ diagnosis.
49.	Derived AJCC N, 6th ed (2004+)	This is the AJCC "N" component that is derived from CS coded fields, using the CS algorithm, effective with 2004+ diagnosis.
50.	Derived AJCC N, 6th ed (2004+)	This is the AJCC "N" component that is derived from CS coded fields, using the CS algorithm, effective with 2004+ diagnosis.
51.	Derived AJCC M, 6th ed (2004+)	This is the AJCC "M" component that is derived from CS coded fields, using the CS algorithm, effective with 2004+ diagnosis.
52.	Derived AJCC Stage Group, 6th ed (2004+)	This is the AJCC "Stage Group" component that is derived from CS detailed site-specific codes, using the CS algorithm, effective with 2004+ diagnosis.
53.	Derived SS1977 (2004+)	This item is the derived "SEER Summary Stage 1977" from the CS algorithm, effective with 2004+ diagnosis.
54.	Derived SS2000 (2004+)	This item is the derived "SEER Summary Stage 2000" from the CS algorithm, effective with 2004+ diagnosis.
55.	Derived AJCC— Flag (2004+)	Flag to indicate whether the derived AJCC stage was derived from CS or EOD codes.
56.	Derived SS1977— Flag (2004+)	Flag to indicate whether the derived SEER Summary Stage 1977 was derived from CS or EOD codes.
57.	Derived SS2000—	Flag to indicate whether the derived SEER Summary

		Flag (2004+)	Stage 2000 was derived from CS or EOD codes.
58.		CS version input (2004+)	This item indicates the number of the version used to initially code CS fields.
59.		CS version latest (2004+)	This item indicates the number of the version of the CS used most recently to derive the CS output fields.
60.		CS version input current (2004+)	This item indicates the number of the version of the CS after input fields have been updated or recoded.
61.		RX Summ--Surg Prim Site (1998+)	Surgery of Primary Site describes a surgical procedure that removes and/or destroys tissue of the primary site performed as part of the initial work-up or first course of therapy.
62.		RX Summ--Scope Reg LN Sur (2003+)	Scope of Regional Lymph Node Surgery describes the procedure of removal, biopsy, or aspiration of regional lymph nodes performed during the initial work-up or first course of therapy at all facilities.
63.		RX Summ--Surg Oth Reg/Dis (2003+)	Surgical procedure of Other Site describes the surgical removal of distant lymph node(s) or other tissue(s) or organ(s) beyond the primary site.
64.		Num of regional lym nd exam (1998-2002)	This data item records the number of regional lymph nodes examined in conjunction with surgery performed as part of the first course of treatment at all facilities.
65.		First course of reconstruct (1998-2002)	The SEER program collects information in this field only for breast cancer and only for reconstruction begun as part of first course of treatment.
66.		Reason no cancer-directed surgery	This data item documents the reason that surgery was not performed on the primary site.
67.		Radiation	This data item indicates the method of radiation

		therapy performed as part of the first course of treatment.
68.	Radiation to Brain or CNS	This variable was only collected for years 1988-1997 for lung and leukemia cases only.
69.	Radiation sequence with surgery	This field records the order in which surgery and radiation therapies were administered for those patients who had both surgery and radiation.
70.	Site specific surgery	The actual coding schemes for individual sites and time periods can be viewed at http://seer.cancer.gov/manuals/historic/AppendD.pdf. The discussion below summarizes the information available by site and time period.
71.	Surgery of primary site	Site-specific codes for the type of surgery to the primary site performed as part of the first course of treatment at all facilities for cases diagnosed 1998-2002.
72.	Scope of reg lymph nd surg	This field describes the removal, biopsy or aspiration of regional lymph node(s) at the time of surgery of the primary site or during a separate surgical event at all facilities for cases diagnosed 1998-2002.
73.	Surgery of oth reg/dis sites	This field records the removal of distant lymph nodes or other tissue(s)/organ(s) beyond the primary site given at all facilities as part of the first course of treatment for cases diagnosed 1998-2002.
74.	Record number	The Record Number is a unique sequential number. The highest number for each patient identifies the number of records that have been submitted to SEER for that particular patient.

75.	Site-behavior override	The IF39 edit does not allow in situ cases of nonspecific sites, such as gastrointestinal tract, NOS; uterus, NOS; female genital tract, NOS; male genital organs, NOS; and others.
76.	Site-EOD-dx date override	The IF40 edit does not allow "localized" disease with non-specific sites, such as mouth, NOS; colon, NOS (except histology 8220); bone, NOS; female genital system, NOS; male genital organs, NOS; and others. This over-ride indicates that the conflict has been reviewed.
77.	Site-laterality-EOD override	The IF41 edit for paired organs does not allow EOD to be specified as in situ, localized, or regional by direct extension if laterality is coded as "bilateral, side unknown" or "laterality unknown." This over-ride indicates that the conflict has been reviewed.
78.	Site-laterality-morph override	The IF42 edit does not allow behavior code of "in situ" with non-specific laterality codes. This over-ride indicates that the conflict has been reviewed.
79.	Type of follow-up expecte	This item codes the type of follow-up expected for a SEER cas
80.	Age recode with <1 year old	The age recode variable is based on Age at Diagnosis (single-year ages). The groupings used in the age recode variable are determined by the age groupings in the population data
81.	Site recod	A recode based on Primary Site and ICD-O-3 Histology in order to make analyses of site/histology groups easier

82.	Site rec with Kaposi and mesotheliom	A recode based on Primary Site and ICD-O-3 Histology in order to make analyses of site/histology groups easier
83.	Recode ICD-O-2 to	The primary site and morphology are recoded to ICD-9
84.	Recode ICD-O-2 to 1	The primary site and morphology are converted to ICD-10
85.	ICCC site recode ICD-O-	A site/histology recode based on the International Classification of Childhood Cancer (ICCC) is mainly used to analyze data on children. The recode was applied to all cases no matter the age in order that age comparisons can be made with these groupings
86.	SEER modified ICCC site recode ICD-O-	A site/histology recode based on the International Classification of Childhood Cancer (ICCC) with slight modifications is mainly used to analyze data on children
87.	ICCC site recode ICD-O-	A site/histology recode that is mainly used to analyze data on children. The recode was applied to all cases no matter the age in order that age comparisons can be made with these groupings
88.	ICCC site recode extended ICD-O-	Based on ICD-O-3. For more information on this International Classification of Childhood Cancer (ICCC) site/histology recode
89.	Behavior recode	This recode was created so that data analyses could eliminate major groups of histologies/behaviors that weren't collected consistently over time, for example benign brain, myelodyplastic syndromes, and borderline tumors of the ovary

90.	Histology recode - broad grouping	Based on Histologic Type ICD-O-3
91.	Histology recode - Brain grouping	Based on Histologic Type ICD-O-3
92.	CS Schema – v 020	CS information is collected under the specifications of a particular schema based on site and histology
93.	Race recode (White, Black, Other	Race recode is based on the race variables and the American Indian/Native American IHS link variable
94.	Race recode (W, B, AI, API	Caution should be exercised when using this variable
95.	Origin recode NHIA (Hispanic, Non-Hisp	Caution should be exercised when using this variable
96.	SEER historic stage	Derived from Collaborative Stage (CS) for 2004+ and Extent of Disease (EOD) from 1973-2003
97.	AJCC stage 3rd edition	Derived by algorithm from extent of disease (EOD). Not available for all years or for all sites
98.	SEER modified AJCC stage 3rd (1988-2003	Derived by algorithm from extent of disease (EOD). Not available for all years or for all sites
99.	SEER summary stage	Based on SEER Extent of Disease (EOD) following a SEER algorithm. This variable is provided on the NAACCR call for data
100.	SEER summary stage	Based on SEER Extent of Disease (EOD) following a SEER algorithm. This variable is provided on the NAACCR call for data
101.	Number of	Based on the total number of tumors in SEER, not

		primaries	necessarily this database file. This value is the same across all tumors for a person
102.		First malignant primary indicato	Based on all the tumors in SEER. Tumors not reported to SEER are assumed malignant
103.		State-count	This item is a state-county combination where the first two characters represent the state FIPS code. The last three digits represent the FIPS county code
104.		Survival time recode (month) Survival time recode (year	The Survival Time Recode is calculated using the date of diagnosis and one of the following: date of death, date last known to be alive, or follow-up cutoff date used for this file (see title page for date for this file)
105.		COD to site recod	This recode was introduced to account for several newly valid ICD-10 codes and includes both cancer and non-cancer causes of death
106.		COD to site rec K	This is a recode based on underlying cause of death to designate cause of death into groups similar to the incidence site recode with KS and mesothelioma
107.		IHS Lin	Incidence files are periodically linked with Indian Health Service (IHS) files to identify Native Americans
108.		Stage	Summary Stage 2000 is derived from Collaborative Stage (CS) for 2004+ and Extent of Disease (EOD) from 1998-2003. It is a simplified version of stage: in situ, localized, regional, distant, & unknown. Used in the SEER CSR and more recent SEER publications
109.		AYA site recod	A site/histology recode that is mainly used to analyze data on adolescent and young adults

110.	Lymphoma subtype recod	A site/histology recode that is mainly used to analyze data on adolescent and young adults
111.	SEER cause-specific death classificatio	Created for use in cause-specific survival. This variable designates that the person died of their cancer for cause-specific survival
112.	SEER other cause of death classificatio	Created for use in left-truncated life table session. This variable designates that the person died of causes other than their cancer
113.	CS Tumor Size/Ext Eval	CS Tumor Size/Ext Eval
114.	CS Mets Eval	Available for 2004+, but not required for the entire timeframe. Will be blank in cases not collected
115.	Primary by International Rule	Created using IARC multiple primary rules. Did not include benign tumors or non-bladder in situ tumors in algorithm
116.	ER Status Recode Breast Cancer (1990+	Created by combining information from Tumor marker 1 (1990-2003) (NAACCR Item #=1150), with information from CS site-specific factor 1
117.	PR Status Recode Breast Cancer (1990+	Created by combining information from Tumor marker 2 (1990-2003) (NAACCR Item #=1150), with information from CS site-specific factor 2
118.	CS Schema- AJCC 6th e	CS information is collected under the specifications of a particular schema based on site and histology
119.	CS site-specific factor 8	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information. Available for varying years and schemas depending on standard setter

		requirements
120.	CS site-specific factor 10	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information. Available for varying years and schemas depending on standard setter requirements
121.	CS site-specific factor 11 (2004+	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information
122.	CS site-specific factor 13	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information. Available for varying years and schemas depending on standard setter requirements
123.	CS site-specific factor 15	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information
124.	CS site-specific factor 16 (2004+	Each CS site-specific factor (SSF) is schema dependent. They can provide information needed to stage the case, clinically relevant information, or prognostic information
125.	Lymph-vascular Invasion	LVI is required for cases originally coded under CSv2 or diagnosed 2010+ for the schemas for penis and testis only

126.	Coded	---
127.	Coded	---
128.	Coded	---
129.	Coded	---
130.	Coded	---
131.	Coded	---
132.	Coded	---
133.	Coded	---
134.	Coded	---