



California Cancer Registry Data Dictionary

A:

ACCNO: Accession number assigned to the patient by the reporting hospital according to American College of Surgeons (ACOS) definitions.

ACOS: Flag which indicates whether or not the reporting facility has an ACOS-approved cancer program:

Values:

- 1: Cancer program approved
- 2: Cancer program not approved
- Blank: Cases diagnosed before 1999

ADMISSION_ID: CCR unique ID for each admission.

AGE: Age of the patient in years at time of diagnosis.

AGECAT: Age group where AGE falls for statistical purposes:

Values:

- | | |
|----------------|-----------------|
| 1: 0-4 years | 10: 45-49 years |
| 2: 5-9 years | 11: 50-54 years |
| 3: 10-14 years | 12: 55-59 years |
| 4: 15-19 years | 13: 60-64 years |
| 5: 20-24 years | 14: 65-69 years |
| 6: 25-29 years | 15: 70-74 years |
| 7: 30-34 years | 16: 75-79 years |
| 8: 35-39 years | 17: 80-84 years |
| 9: 40-44 years | 18: 85+ years |
| | 99: Unknown age |

AJCC_C and AJCC_P: A two-digit summary code of the (first) clinical and (second) pathological TNM classification entered for this case (see AJCC Manual for Staging of Cancer):

Values:

- 0+: Stage 0
- 1*: Stage I, IA, or IB
- 2*: Stage II, IIA, or IIB
- 3*: Stage III, IIIA, or IIIB
- 4*: Stage IV, IVA, or IVB
- 88: Not applicable
- 99: Recurrent, unstaged, unknown, or Stage X
- +Second character may be A, I, S, or blank
- *Second character may be A, B, C, E, S, or blank

B:

BEHAVIOR: SEER behavior recode for analysis:

- 0 = Benign in both ICD-O-2 and ICD-O-3
- 1 = Borderline in both ICD-O-2 and ICD-O-3
- 2 = In situ in both ICD-O-2 and ICD-O-3
- 3 = Malignant in both ICD-O-2 and ICD-O-3
- 4 = Only malignant in ICD-O-3
- 5 = No longer reportable in ICD-O-3
- 9 = Unknown

BELOW200: Census variable- percent of block group below 200% of the Federal poverty level.

BIRTHPL: Place of birth for the patient- state or country (please refer to Appendix D of Volume I for codes):

http://ccrcal.org/Vol_1_May_2008_appendix_html/Vol_I_CA_Append.htm

BLKGRP00: Block groups are subdivisions of Census tracts. A block group is a set of blocks sharing the same first digit of the four-digit block with a Census tract. For example, block group 3 within a particular Census tract would include any blocks numbered between 301 and 399. This variable uses the 2000 Census.

BLKGRP90: Block groups using the 1990 Census.

BLUECOLR: Census variable- proportion of the block group 16 years of age or older with a blue collar job.

C:

CASEFIND: Method used to first identify the case by this reporting source (source of case finding):

Values:

Case first identified in a cancer-reporting facility:

- 10: Reporting hospital, NOS
- 20: Pathology department review
- 21: Daily discharge review
- 22: Disease index review
- 23: Radiation therapy department/center
- 24: Laboratory reports
- 25: Outpatient chemotherapy
- 26: Diagnostic imaging/radiology
- 27: Tumor board
- 28: Hospital rehabilitation service or clinic
- 29: Other hospital reporting source, including clinic, NOS

Case first identified by source other than cancer-reporting facility:

- 30: Physician-initiated case
- 40: Consultation-only or pathology-only report
- 50: Private pathology lab report
- 60: Nursing home initiated case
- 70: Coroner's Office records review
- 75: Managed care organization (MCO) or insurance records
- 80: Death certificate follow-back
- 85: Out-of-state case sharing
- 90: Other non-reporting hospital source
- 95: Quality control review
- 99: Unknown

CDEATH: Underlying cause of death on the death certificate - please refer to ICD-9 (1988-1999 year of death) or ICD-10 (2000 and beyond year of death):

CCR values:

- 0000: Patient not dead
- 7777: Death certificate not yet available

CDEATH_RECODE: Underlying cause of death as of December 31, 2005 on the death certificate - please refer to ICD-9 (1988-1999 year of death) or ICD-10 (2000 and beyond year of death):

CCR values:

- 0000: Patient not dead
- 7777: Death certificate not yet available

CDP2000: Census variable- Census designated place, 2000 Census.

CENSUS00: Census tract of usual residence when this tumor was first diagnosed, using year 2000

Census tracting:

CCR values:

999993: Unknown street address or city, but state is CA

999994: PO Boxes for CA residents only

999995: Neither manual nor machine tracting has been successful

999996: Not yet submitted for tracting

999997: Submitted for tracting once- unsuccessfully

999998: California case- not machine tractable

999999: Non-California resident or unknown residence

CENSUS90: Census tract of usual residence when this tumor was first diagnosed, using 1990

Census tracting:

CCR values:

999993: Unknown street address or city, but state is CA

999994: PO Boxes for CA residents only

999995: Neither manual nor machine tracting has been successful

999996: Not yet submitted for tracting

999997: Submitted for tracting once- unsuccessfully

999998: California case- not machine tractable

999999: Non-California resident or unknown residence

CENTRD00: Flag which indicates the basis (certainty) of assignment of the census tract, 2000

Census:

Values:

1: Census tract based on complete and valid street address

2: Census tract based on residence zip+4

3: Census tract based on residence zip+2

4: Census tract based on residence zip only

5: Census tract based on zip of post office box

9: Unable to assign census tract based on available information

Blank: Census tracting not yet attempted/not applicable

CENTRD90: Flag which indicates the basis (certainty) of assignment of the census tract, 1990

Census:

Values:

1: Census tract based on complete and valid street address

2: Census tract based on residence zip+4

3: Census tract based on residence zip+2

4: Census tract based on residence zip only

5: Census tract based on zip of post office box

9: Unable to assign census tract based on available information

Blank: Census tracting not yet attempted/not applicable

CHEMO_HOSP: Identifies the type of chemotherapy given at the reporting hospital. If chemo not given, codes are provided to record that reason:

Values:

- 00: None
- 01: Chemotherapy, NOS
- 02: Single agent
- 03: Multiple agents
- 82: Contraindicated
- 85: Patient died
- 86: Recommended, not given
- 87: Refused
- 88: Recommended, unknown if given
- 99: Unknown or death certificate only case (DCO)

CHEMOSUM: Identifies the type of chemotherapy given as first course of treatment at this and all other facilities. If chemo not given, codes are provided to record that reason:

Values:

- 00: None
- 01: Chemotherapy, NOS
- 02: Single agent
- 03: Multiple agents
- 82: Contraindicated
- 85: Patient died
- 86: Recommended, not given
- 87: Refused
- 88: Recommended, unknown if given
- 99: Unknown or death certificate only case (DCO)

CHMPPC: Designation of rural/urban for a MSSA by the California Health Manpower Policy Commission. The codes are:

- R = rural
- blank = urban
- 9 = unknown

CLASS: Distinguishes cases that are usually included in a hospital's treatment and survival statistics (0-2) from those that are not (3-9):

Values:

- 0: Diagnosis only here
- 1: Diagnosis and treatment here
- 2: Treatment here
- 3: Diagnosis and treatment elsewhere
- 4: Diagnosis and/or treatment prior to hospital reference date
- 5: Diagnosis at autopsy
- 6: Staff physician
- 7: Pathology report only
- 8: Death certificate only
- 9: Unknown

CNTYFIPS: Census variable- California county FIPS code- to get the FIPS County code, take the COUNTY variable and then multiply by 2 and subtract by 1. For example, Los Angeles County is COUNTY=19, CNTYFIPS would= $19*2=38-1=37$ (CNTYFIPS=37).

CODER_C and CODER_P: Identifies who did any of the TNM coding. If a physician did any coding, it is coded to physician:

- 0 = Not staged
- 1 = Managing physician
- 2 = Pathologist
- 3 = Other physician
- 4 = Any combination of 1, 2, or 3.
- 5 = Registrar
- 6 = Any combination of 5 with 1, 2, or 3.
- 7 = Other
- 8 = Staged, individual not specified
- 9 = Unknown if staged
- Blank = Not abstracted.

CODPROC: Designates the set of rules used to code this case

Values:

- 03-09: Coded prior to May 1978 (Region 8 only)
- 10: Coded May 1978- Dec. 1983 (Region 8 only)
- 11: Coded Jan. 1984- Apr. 1988
- 12: Coded May 1988 and later
- 13: Coded June 1991 and later
- 14: Coded Jan. 1, 1992 and later
- 15: Coded during or after summer 1993
- 16: Coded beginning in 1996
- 17: Coded beginning in 1997
- 18: Coded beginning in 1998
- 19: Coded beginning in 2000
- 20: Coded beginning in 2001
- 21: Coded beginning in 2003
- 22: Coded beginning in 2004
- 23: Coded beginning in 2005
- 24: Coded beginning in 2006

COLLEGE: Census variable- proportion of block group 25 years of age and older with a college degree.

CONSTATE: Contact state.

CONZIP: Contact zip code.

COUNTY: County of residence at diagnosis:

Values:

000: California, NOS	030: Orange
001: Alameda	031: Placer
002: Alpine	032: Plumas
003: Amador	033: Riverside
004: Butte	034: Sacramento
005: Calaveras	035: San Benito
006: Colusa	036: San Bernardino
007: Contra Costa	037: San Diego
008: Del Norte	038: San Francisco
009: El Dorado	039: San Joaquin
010: Fresno	040: San Luis Obispo
011: Glenn	041: San Mateo
012: Humboldt	042: Santa Barbara
013: Imperial	043: Santa Clara
014: Inyo	044: Santa Cruz
015: Kern	045: Shasta
016: Kings	046: Sierra
017: Lake	047: Siskiyou
018: Lassen	048: Solano
019: Los Angeles	049: Sonoma
020: Madera	050: Stanislaus
021: Marin	051: Sutter
022: Mariposa	052: Tehama
023: Mendocino	053: Trinity
024: Merced	054: Tulare
025: Modoc	055: Tuolumne
026: Mono	056: Ventura
027: Monterey	057: Yolo
028: Napa	058: Yuba
029: Nevada	
100-725: Non-USA (refer to country codes in Appendix D of Vol. I)	
999: Unknown	

***For the following Collaborative Staging variables, please refer to the “Collaborative Staging Manual and Coding Instructions” (currently version 1.0):**

<http://www.cancerstaging.org/cstage/index.html>

CS_EXT: Part of the Collaborative Staging (CS) data item set and is only valid for cases diagnosed 2004 forward. It is based on and replaces the extension variable (DIREXTTU), and identifies the primary tumor growth within the organ of origin or its extension into neighboring organs. This modification for CS is collapsible into AJCC T-code. Values are site-specific.

CS_LN: Part of the Collaborative Staging (CS) data item set and is only valid for cases dx 2004 forward. This data item is based on and replaces the number of nodes involved variable (LNSUMTU) and identifies the regional lymph nodes involved. This modification for CS is collapsible into AJCC N-code. Values are site-specific.

CS_METS_DX: Part of the Collaborative Staging (CS) data item set and is only valid for cases dx 2004 forward. This variable identifies the site(s) of metastatic involvement at the time of diagnosis. Values are site-specific.

CS_METS_EVAL: Part of the Collaborative Staging (CS) data item set and is only valid for cases dx 2004 forward. Records how codes for CS_METS were determined- based on the diagnostic method employed. It is used in CS to identify whether the M-code (of AJCC TNM) was clinically or pathologically diagnosed.

CS_REG_LN_EVAL: Part of the Collaborative Staging (CS) data set and valid for cases diagnosed 2004 forward. Records how the codes for CS_LN were determined based on the diagnostic methods employed. It is used in CS to identify whether the N (of AJCC TNM) was clinically or pathologically diagnosed.

CS_SITE_SPEC_F1, F2, F3, F4, F5, and F6: Part of the Collaborative Staging (CS) data item set and is valid for cases diagnosed 2004 and forward. These 6 variables are used to code additional site-specific information needed to derive TNM or AJCC stage, or to code prognostic factors that have an effect on stage or survival. Values are site-specific.

CS_TUM_SIZE: Part of the Collaborative Staging (CS) data item set and is valid for cases diagnosed 2004 forward. Based on and replaces tumor size (TSIZETU), and for most sites records the largest dimension or diameter of the primary tumor in millimeters. For many sites, the CS algorithm uses this data item to derive the DERIVED_AJCC_T.

CS_TS_EXT_EVAL: Part of the Collaborative Staging (CS) data item set and is valid for cases diagnosed 2004 forward. Records how the codes for CS_TUM_SIZE and CS_EXT were determined based on the diagnostic methods employed. This data item is used in CS to identify whether the T (of AJCC TNM) was clinically or pathologically diagnosed.

CS_VERSION_1ST: Indicates the version number used to initially code CS fields.

CS_VERSION_LATEST: Indicates the version number of CS used most frequently to derive the output fields.

D:

DATEADM: Date patient was first seen at this hospital for evaluation and/or treatment of this tumor. This date is independent of whether the patient was an inpatient or an outpatient (MMDDCCYY format).

DATECOMP: Date this abstract was first completed by this abstractor. Used to monitor abstracting progress (MMDDCCYY format).

DATEDIS: Date of the inpatient discharge for the most definitive surgery, other cancer-directed therapy, or diagnostic evaluation. This date corresponds with the date of inpatient admission field (MMDDCCYY format).

DATEDX: Date of diagnosis (MMDDCCYY format).

DATEFSNT: Date this tumor was first sent to CCR (MMDDYYCC format).

DATETRAN: Date that the case was transmitted by the hospital to the Regional Registry (MMDDCCYY format).

DC_FATHERS_SURNAME: Surname of patient's father on the death certificate.

DC_FATHERS_SURNAME_UPDT: Surname of patient's father on the death certificate updated from older death certificates.

DCBIRTHPL: State or country of the patient's birth as stated on the death certificate.

DCBIRTHPL_UPDT: State or country of the patient's birth as stated on the death certificate updated from older death certificates.

DCRACE: Race/ethnicity of the patient as stated on the death certificate.

DCRACE_UPDT: Race/ethnicity of the patient as stated on the death certificate updated from older death certificates.

DCSPANISH: Used to denote those persons of Spanish origin as stated on the death certificate.

DCSPANISH_UPDT: Used to denote those persons of Spanish origin as stated on the death certificate updated from older death certificates.

DCSTATE: Postal code for the state that issued the death certificate.

DCSTATE_UPDT: Postal code for the state that issued the death certificate updated from older death certificates.

DCUPDT: Was the death certificate information updated from older death certificates? 0 = No; 1 = Yes.

DEFINIT: Designation of rural/urban for a MSSA by the California Health Manpower Policy Commission. The codes are:

F = Frontier: A frontier Medical Service Study Area is the same as a rural MSSA adopted by the California Health Manpower Policy Commission, but with a smaller population density of less than 11 persons per square mile.

R = Rural: Rural MSSA is defined as a geographic area with a population density of less than 250 persons per square mile, and has no census-defined place within the area with a population in excess of 50,000. Only the population that is located within the MSSA is counted in determining the population of the census defined place.

U = Urban: Not frontier or rural

9 = Unknown

DEN: Density of the MSSA. The codes are:

H = High density: population density over 250 per square mile

L = Low density: population density under 250 per square mile

9 = Unknown

DERIVED_AJCC_FLAG: Flag to indicate whether AJCC was coded directly or was derived from CS or EOD codes:

Values:

1: AJCC 6th Edition derived from Collaborative Staging Manual and Coding Instructions, Vol. 1

2: AJCC 6th Edition derived from Extent of Disease (dx prior to 2004)

DERIVED_AJCC_M: The AJCC M-code that is derived from CS coded fields using the CS algorithm- valid for cases dx 2004 and later:

Values:

00: M0

10: M1

11: M1a

12: M1b

13: M1c

19: M1, NOS

88: Not applicable

99: MX

DERIVED_AJCC_M_DESC: The AJCC M-code descriptor component that is derived from coded fields, using the CS algorithm:

Values:

C: Clinical stage

P: Pathologic stage

A: Autopsy stage

Y: Pathologic examination of metastatic tissue performed after pre-surgical systemic treatment and extension based on pathologic evidence

N: Not applicable

Blank: Date of diagnosis before Jan. 2004

DERIVED_AJCC_N: The AJCC N-code that is derived from coded fields using the CS algorithm, for cases diagnosed 2004 forward:

Values:

00: N0

01: N0 (i-)

02: N0 (i+)

03: N0 (mol-)

04: N0 (mol+)

09: N0, NOS

10: N1

11: N1a

12: N1b

13: N1c

18: N1mi

19: N1, NOS

20: N2

21: N2a

22: N2b

23: N2c

29: N2, NOS

30: N3

31: N3a

32: N3b

33: N3c

39: N3, NOS

88: Not applicable

99: NX

DERIVED_AJCC_N_DESC: The AJCC N-code descriptor component that is derived from coded fields using the CS algorithm:

Values:

C: Clinical stage

P: Pathologic stage

A: Autopsy stage

Y: Lymph nodes removed for examination after pre-surgical systemic treatment or radiation and lymph nodes evaluation based on pathological evidence

N: Not applicable

Blank: Date of diagnosis after Jan. 2004

DERIVED_AJCC_STG_GRP: The AJCC stage group component that is derived from the CS detailed site-specific codes, using the CS from the CS algorithm- valid for cases diagnosed 2004 and forward:

Values:

00: 0	39: IISB
01: 0a	40: IIS
02: 0is	41: IIESA
10: I	42: IIESB
11: I, NOS	43: IIES
12: IA	50: III
13: IA1	51: III, NOS
14: IA2	52: IIIA
15: IB	53: IIIB
16: IB1	54: IIIC
17: IB2	55: IIIEA
18: IC	56: IIIEB
19: IS	57: IIIE
20: IEA	58: IIISA
21: IEB	59: IIISB
22: IE	60: IIIS
23: ISA	61: IIIESA
24: ISB	62: IIIESB
30: II	63: IIIES
31: II, NOS	70: IV
32: IIA	71: IV, NOS
33: IIB	72: IVA
34: IIC	73: IVB
35: IIEA	74: IVC
36: IIEB	88: Not Applicable
37: IIE	90: Occult
38: IISA	99: Unknown

Blank: Dx before Jan 2004

DERIVED_AJCC_T: The AJCC T-code that is derived from CS coded fields, using the CS algorithm- valid for cases dx 2004 forward:

Values:

00: T0

01: Ta

05: Tis

06: Tispu

07: Tispd

10: T1

11: T1mic

12: T1a

13: T1a1

14: T1a2

15: T1b

16: T1b1

17: T1b2

18: T1c

19: T1, NOS

20: T2

21: T2a

22: T2b

23: T2c

29: T2, NOS

30: T3

31: T3a

32: T3b

33: T3c

39: T3, NOS

40: T4

41: T4a

42: T4b

43: T4c

44: T4d

49: T4, NOS

88: Not applicable

99: TX

Blank: Date of diagnosis is before Jan. 2004

DERIVED_AJCC_T_DESC: The AJCC T-code descriptor component that is derived from CS coded fields, using the CS algorithm:

Values:

C: Clinical stage

P: Pathologic stage

A: Autopsy stage

Y: Surgical resection performed after pre-surgical systemic treatment or radiation; tumor size/extension based on pathologic evidence

N: Not applicable

DERIVED_SS1977: Derived “SEER Summary Stage 1977” from the CS algorithm (or EOD codes); effective with cases dx 2004 forward:

Values:

0: In situ

1: Localized

2: Regional by direct extension

3: Regional by lymph nodes

4: Regional by extension and nodes

5: Regional, NOS

7: Distant

8: Not applicable

9: Unknown/unstaged

DERIVED_SS1977_FLAG: Flag to indicate whether SEER Summary Stage 1977 was coded directly or was derived from CS or EOD codes:

Values:

1: SS1977 derived from Collaborative Staging Manual and Coding Instructions; Version 1.0

2: SS1977 derived from EOD (prior to 2004)

Blank: Not derived and date of diagnosis before 2004

DERIVED_SS2000: Derived “SEER Summary Stage 2000” from the CS algorithm (or EOD codes); effective with cases dx 2004 forward

Values:

0: In situ

1: Localized

2: Regional by direct extension

3: Regional by lymph nodes

4: Regional by extension and nodes

5: Regional, NOS

7: Distant

8: Not applicable

9: Unknown/unstaged

DERIVED_SS2000_FLAG: Flag to indicate whether SEER Summary Stage 2000 was coded directly or was derived from CS or EOD codes:

Values:

1: SS2000 derived from Collaborative Staging Manual and Coding Instructions; Version 1.0

2: SS2000 derived from EOD (prior to 2004)

Blank: Not derived and date of diagnosis before 2004

DIFFERN2: Sixth digit of ICD-O-3; designates the grade or differentiation of the tumor:

Values:

1: Grade I or well differentiated

2: Grade II or moderately well differentiated

3: Grade III or poorly differentiated

4: Grade IV or undifferentiated/anaplastic

5: T-cell

6: B-cell

7: Null cell

8: Natural Killer (NK) cell

9: Grade and differentiation not stated

DIREXTTU: Two-digit extension code that represents the growth of the primary tumor within the organ of origin, its extension into neighboring organs, or its metastasis to distant structures. Only valid for cases diagnosed prior to Jan. 2004- values are site-specific.

DOB: Date of birth of the patient (MMDDCCYY format).

DRATTEND: California license number of the attending/managing physician at the time of this admission for this tumor. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DRFOLLOW: California license number of the physician who will be seeing the patient after discharge from the hospital and who could be contacted for follow-up information. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DRMEDONC: California license number of the medical oncologist. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DROTHER1 & 2: California license number of the physicians other than the attending and following physicians. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DRRADONC: California license number of the radiation oncologist. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DRREF: California license number of the referring physician. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DRSURG: California license number of the surgeon. Out-of-state physicians' license numbers may be entered here with the first character being an "X".

DRXTTU_P: Two-digit code for prostate cancers only (and used only after prostatectomy within 4 months of diagnoses) to code for the extension of the tumor.

DTDEFSURG: Records the date of the most definitive surgical resection of the primary site performed as the first course of treatment (MMDDCCYY format):

All 0's: No surgery or unknown if ever treated

All 9's: DC only case or definitive surgery performed, but the date is unknown

DTSYSTEMIC: Records the date of initiation for systemic therapy that is part of the first course of treatment. Systemic therapy is considered to be: chemotherapy agents, hormonal agents, biological response modifiers, bone marrow transplants, stem cell harvests, and surgical and/or radiation endocrine therapy (MMDDCCYY format).

DTTRANSP: Date that the transplant/endocrine procedure was done (MMDDCCYY format).

DXCONF: Indicates whether at any time during the patient's medical history there was microscopic confirmation of this cancer:

Values:

1: Histology, including hematology

2: Cytology

4: Microscopically confirmed, but type of specimen unknown

5: Positive lab test or marker study

6: Direct observation

7: Radiography

8: Clinical, other

9: Unknown

DXSTATE: Address at diagnosis, state- 2 character abbreviation.

DXZIP: Address at diagnosis, zip code.

E:

EDINDEX: Census variable- education index.

F:

FUDATE: Date of last contact/known to be alive if patient is alive/not known to be dead (if VITSTAT=1) or date of death if patient is dead (VITSTAT=0); MMDDCCYY format.

FUTYPPAT: Source of last follow-up information, whether that follow-up was specifically looking for patient or tumor information:

Values:

HOSPITAL:

- 00: Admission being reported
- 01: Re-admission to reporting hospital
- 02: Follow-up report from physician
- 03: Follow-up report from patient
- 04: Follow-up report from relative
- 05: Obituary
- 06: Follow-up report from Social Security Admin or Medicare
- 07: Follow-up report from hospice
- 08: Follow-up report from other hospital
- 09: Other source
- 11: Telephone call to any source
- 12: Special studies
- 13: Equifax
- 14: AIDS Registry System
- 15: Computer match with discharge data (OSHDP)

REGIONAL REGISTRY:

- 20: Letter to a physician
- 21: Computer match with DMV
- 22: Computer match with Medicare or Medicaid file
- 23: Computer match with HMO file
- 24: Computer match with voter registration file
- 25: National Death Index
- 26: Computer match with state death tape
- 27: Social Security, Death Master file
- 29: Computer match, other or NOS
- 30: Other source
- 31: Telephone call to any source
- 32: Special studies
- 33: Equifax
- 34: AIDS Registry System
- 35: Computer match with discharge data (OSHDP)
- 36: Obituary
- 37: Computer match using address service
- 38: TRW Credit
- 39: Regional Registry follow-up listing

CENTRAL REGISTRY:

- 40: Letter to a physician
- 41: Telephone call to any source
- 50: CMS (Center for Medicare & Medicaid Services)

- 51: DMV
- 52: California Medical Review, Inc. (CMRI)
- 53: HMO file
- 54: California Voter Registration file
- 55: National Death Index
- 56: State death tape- death clearance (StatMaster)
- 57: Medi-cal/Medicaid eligibility
- 58: Social Security- deaths
- 59: Computer match, other or NOS
- 60: Other source
- 61: Social Security- SSN
- 62: Special studies
- 65: OSHPD hospital discharge data
- 66: National change of address (NCOA)
- 67: Social Security Admin- epidemiological vital status
- 68: Property tax linkage
- 69: State death tape- incremental death clearance

HOSPITAL, SUPPLEMENTAL:

- 73: Computer match with HMO file
- 76: Computer match with state death tape

REGIONAL REGISTRY, SUPPLEMENTAL:

- 80: Social Security Admin- epidemiological vital status
- 81: Property tax linkage
- 82: Probe360
- 83: SSDI- internet
- 84: E-path
- 85: Path labs
- 86: Patient
- 87: Relative

UNKNOWN SOURCE:

- 99: Source unknown

FUTYPTUM: Source of last tumor follow-up information which gave a specific tumor status:

*Values are the same as for FUTYPPAT above

H:

HISPSURN: Is the Hispanic surname on the 1980 Census list?

HISTO_M2: Fifth digit of the ICD-O-2 morphology code- indicates the malignancy or behavior of this tumor:

Values:

2: In situ

3: Malignant/invasive

HISTO_M3: Fifth digit of the ICD-O-3 morphology code- indicates the malignancy or behavior of this tumor:

Values:

0: Benign

1: Uncertain/borderline

2: In situ

3: Malignant/invasive

HISTO_T2 & _T3: The first four digits of the ICD-O-2 and 3 morphology code, respectively- indicates the histology/cell type of this tumor.

*Refer to ICD-O-2 or -3 for values.

HMNGSURN: Hmong surname, 1 = yes.

HORMSUM: Records whether systemic hormonal agents were given as first course of treatment at any facility, or the reason why they were not given:

Values:

00: None

01: Hormones given

82: Contraindicated

85: Patient died

86: Recommended, not given

87: Refused

88: Recommended, unknown if given

99: Unknown or death certificate only (DCO) case

HOSPNO: Unique 10-digit number assigned by CCR for the hospital or other facility reporting this case to the CCR.

*Please refer to the following link on the CCR web page for values:

<http://www.ccrca.org/PDF-DSQC/CAHospLabels-Vers-1.8.0.8,5-14-08-Code.pdf>

I:

ICCCMDG: ICCC (for childhood cancers) recode from ICD-O-3.

IHSLINK: Did patient match to Indian Health Services file? 0 = No; 1 = Yes.

IMMUSUM: Records whether immunotherapy given as first course of treatment at this and all other facilities. If it was not given, codes are provided that record the reason:

Values:

- 00: None, or diagnosed at autopsy
- 01: Immunotherapy given
- 82: Contraindicated
- 85: Patient died
- 86: Recommended, not given
- 87: Refused
- 88: Recommended, unknown if given
- 99: Unknown or death certificate only (DCO) case

INDUS80 & 90: Identifies the type of industry or business associated with the longest-held occupation at time of diagnosis. Coding scheme was used by the Census Bureau in 1980 and 1990.

*Refer to Volume III, Appendix 14 for codes:

http://www.ccrca.org/VOL3/Vol_3_CA.htm

L:

LATERAL: For some specific primary sites, the side of the body in which the tumor originated:

Values:

- 0: Not a paired site
- 1: Right
- 2: Left
- 3: Unilateral, NOS
- 4: Bilateral
- 9: Paired site, no information

LATITUDE: The distance, north or south, measured in degrees along a meridian as provided by the geocoding vendor.

LNSUMTU: Lymph node involvement:

Values:

- 0: No nodes involved
- 1-8: Values vary by site
- 9: Unknown whether nodes involved

LONGITUDE: The distance, east or west of the prime meridian at Greenwich, England, measured in degrees as provided by the geocoding vendor.

M:

MAIDSURN: Is Hispanic maiden name on 1980 Census list?

MARKER1: Field to record estrogen receptor for breast cancer, acid phosphatase for prostate cancer, alphafetoprotein for testicular cancer, carcinoembryonic antigen (CEA) for colorectal cancer, and carbohydrate antigen (CA-125) for ovarian cancer. If a hospital is collecting tumor markers for other cancer sites, they may use this field:

Values:

- 0: Test not done (includes cases dx at autopsy)
- 1: Test done, results positive/elevated
- 2: Test done, results negative/normal
- 3: Test done, results borderline or undetermined whether positive or negative
- 4: Range 1
- 5: Range 2
- 6: Range 3
- 8: Test ordered, results not in chart
- 9: Unknown if test done or ordered/no info

MARKER2: Field to record progesterone receptor for breast cancer, prostate specific antigen (PSA) for prostate cancer, and HCG for testicular cancer. If a hospital is collecting tumor markers for other cancer sites, they may use this field:

Values:

- 0: Test not done (includes cases dx at autopsy)
- 1: Test done, results positive/elevated
- 2: Test done, results negative/normal
- 3: Test done, results borderline or undetermined whether positive or negative
- 4: Range 1- <5000 mIU/ml
- 5: Range 2- 5000-50,000 mIU/ml
- 6: Range 3- >50,000 mIU/ml
- 8: Test ordered, results not in chart
- 9: Unknown if test done or ordered/no info

MARKER3: Field to record LDH for testicular cancer:

Values:

- 0: Test not done (includes cases dx at autopsy)
- 1: Test done, results positive/elevated
- 2: Test done, results negative/normal
- 3: Test done, results borderline or undetermined whether positive or negative
- 4: Range 1- <1.5*N
- 5: Range 2- 1.5-10*N
- 6: Range 3- >10*N
- 8: Test ordered, results not in chart
- 9: Unknown if test done or ordered/no info

MARKERCA: Breast cancer tumor marker for California- 1: Her2/neu (also known as c-erbB2 or ERBB2):

Values:

- 0: Not done
- 1: Positive
- 2: Negative
- 3: Borderline
- 8: Ordered, results not in chart
- 9: Unknown

MARSTAT: Marital status when the patient was first diagnosed:

Values:

- 1: Single, never married
- 2: Married
- 3: Separated
- 4: Divorced
- 5: Widowed
- 9: Unknown

MCODE: EOD recode of M code for TNM.

MCODE_C & _P: The M-code that describes the absence or presence of distant metastasis, as staged clinically (first) and pathologically (second):

Values:

- 0: No distant metastasis (M0)
- 1: Distant metastasis present (M1, M1a, M1b, or M1c)
- 2: Visceral metastasis (for melanoma)
- 88: Not applicable
- X: Not assessed (MX)

MDINC: Census variable- median household income for the block group.

MEDRENT: Census variable- median gross rent for the block group, renter-occupied houses.

MEDVALUE: Census variable- median value of owner occupied houses in block group.

MSSAID: Census variable- MSSA identification.

MSSANAME: Census variable- MSSA name.

N:

NCODE: EOD recode of N-code for TNM.

NCODE_C & _P: The N-code that represents the regional lymph node involvement, as staged clinically (first) and pathologically (second).

*Definitions for these codes vary by site- refer to AJCC Manual for Staging

NHIA: The NAACCR Hispanic Identification Algorithm (NHIA) uses a combination of NAACCR variables to directly or indirectly classify cases as Hispanic for analytical purposes. The algorithm uses the following NAACCR standard variables: Spanish/Hispanic origin (190), Name-Last (2230), Name-Maiden (2390), Birthplace (250), Race 1 (160), and Sex (220). The CCR generates this value by examining the primary last name, all alias last names, all maiden names, and the death certificate father's surnames:

Values:

0: Non-Hispanic

1: Mexican

2: Puerto Rican

3: Cuban

4: South or Central American (except Brazilian)

5: Other specified Spanish/Hispanic origin (includes European, excludes Dominican Republican)

6: Spanish, NOS/Hispanic, NOS/Latino, NOS

7: NHIA surname match only

8: Dominican Republican

Blank: Algorithm has not been run

NNODES: Number of regional lymph nodes identified in the pathology report during surgical procedure- this variable is only valid for cases diagnosed prior to Jan. 2004:

Values:

00: No regional nodes examined

01: One regional node examined

02-89: Actual number examined

90: Ninety or more nodes examined

95: No regional lymph nodes removed but aspiration of regional lymph nodes was performed

96: Regional lymph node removal documented as a sampling and number of lymph nodes unknown/not stated

97: Regional lymph node removal documented as dissection and number of lymph nodes unknown/not stated

98: Regional lymph nodes surgically removed, but number of lymph nodes unknown/not stated and not documented as sampling or dissection

99: Unknown/not stated/death certificate only

NOCONTACT: A flag to indicate that the patient does not want to be contacted regarding research:

Values:

- 0: No flag
- 1: Hospital first notified
- 2: Region first notified
- 3: CCR first notified
- 4: Out of state case, not for research

NOEDUC: Census variable- proportion of block group \geq 25 years of age without a high school education.

NOJOB: Census variable- proportion of block group \geq 16 years of age in labor force that is unemployed.

NONRURAL: Census variable- from 1990 MSSA, 0 = rural, 1 = urban.

NORAD: Reason why the first course of treatment did not include radiation:

Values:

- 0: Radiation was performed
- 1: Not recommended
- 2: Contraindicated due to other conditions
- 6: Reason unknown
- 7: Refused
- 8: Recommended, unknown if given
- 9: Unknown if radiation either recommended or given

NOSURG: Reason why the first course of treatment did not include surgery:

Values:

- 0: Surgery was performed
- 1: Not recommended
- 2: Contraindicated due to other conditions
- 6: Reason unknown
- 7: Refused
- 8: Recommended, unknown if given
- 9: Unknown if surgery either recommended or given

O:

OCCUP80 & 90: The code for the patient's longest-held occupation at the time of diagnosis. The coding scheme is that used by the Census Bureau in 1980 and 1990, respectively

*Refer to Volume III, Appendix 14 for values

http://www.ccrca.org/VOL3/Vol_3_CA.htm

Override flags: Indicates that the unusual combination of codes in different fields have been reviewed and are correct. The following is a list of the CCR override fields (values- 0: not reviewed, 1: reviewed):

ORADDX: Date of diagnosis and date of admission discrepancy.

ORAGST: Unusual occurrence of SITE_02 and/or HISTO_T3 for a given age at diagnosis.

ORDCSQ: DC-only cases with multiple primaries (other than lymphoma and leukemia).

ORHIST: HISTO_T3 code is one that ICD-O classifies as benign or uncertain behavior.

ORILDEF: Multiple primaries involving ill-defined sites.

ORLATEOD: Site/laterality/EOD conflict.

ORLATHIS: Site/laterality/histology conflict.

ORLLDX: Lymphoma or leukemia with DXCONF=6 (direct observation) or 8 (clinical, other).

ORNMSX: First name and sex conflict.

ORRACE: Race, Spanish, and birth place discrepancy.

ORSPAN: Spanish and birthplace conflict.

ORSQDX: Multiple primaries where at least one tumor was not microscopically confirmed.

ORSTBEH: Site/behavior conflict.

ORSTSQ: Questionable multiple primaries.

ORSTSTG: Site/stage conflict.

ORSTTY: Unusual SITE_02/HISTO_T3 combination.

ORSUDX: Surgically treated case that was not microscopically confirmed.

OTHREGID: Other Region ID.

OTHREGPT: Other Region patient number.

OTHREGTU: Other Region tumor number.

OTHSUM: Indicates whether the first course of treatment included other types of therapy:

Values:

0: None

1: Other

2: Experimental

3: Double-blind study

6: Unproven methods

7: Refused

8: Planned

9: Unknown

P:

PATIENT_ID: CCR unique ID at the patient level (statewide).

PAYER: Primary source of payment to the hospital:

Values:

- 01: Not insured
- 02: Not insured, self-pay
- 10: Insurance, NOS
- 20: Managed care/HMO/PPO
- 28: HMO
- 29: PPO
- 31: Medicaid
- 35: Medicaid through Managed Care
- 36: Medicaid w/Medicare supplement
- 60: Medicare/Medicaid, NOS
- 61: Medicare w/supplement, NOS
- 62: Medicare- administered through a Managed Care plan
- 63: Medicare with private supplementation
- 64: Medicare with Medicaid eligibility
- 65: TRICARE
- 66: Military
- 67: Veterans Affairs
- 68: Indian/Public Health Service
- 89: County funded, NOS
- 99: Unknown

PAYER2: Secondary source of payment to the hospital.

*Values are the same as for PAYER above

PEDCODER: Identifies the person who documented the pediatric staging:

Values (for pediatric cases only):

- 0: Not staged
- 1: Managing physician
- 2: Pathologist
- 3: Other physician
- 4: Any combination of 1, 2, and 3
- 5: Registrar
- 6: Any combination of 5 with 1, 2, or 3
- 7: Other
- 8: Staged, individual not specified
- 9: Unknown if staged

PEDSTAGE: The pediatric stage as specified in the pediatric staging system selected (so values vary based on the staging system used):

Allowable values:

- 1 (2nd character may be blank, A, or B)
- 2 (2nd character may be blank, A, B, or C)
- 3 (2nd character may be blank, A, B, C, D, or E)
- 4 (2nd character may be blank, A, B, or S)
- 5
- A
- B
- C
- D
- DS
- 88: Not applicable (not a pediatric case)
- 99: Unstaged/unknown (for pediatric cases only)

PEDSYS: Identifies the type of staging system used for the pediatric staging:

Values:

- 00: None
- 01: American Joint Committee on Cancer (AJCC)
- 02: Ann Arbor
- 03: Children's Cancer Group (CCG)
- 04: Evans
- 05: General Summary
- 06: Intergroup Ewings
- 07: Intergroup Hepatoblastoma
- 08: Intergroup Rhabdomyosarcoma
- 09: International System
- 10: Murphy
- 11: National Cancer Institute (Pediatric oncology)
- 12: National Wilms Tumor Study
- 13: Pediatric Oncology Group (POG)
- 14: Reesse-Ellsworth
- 15: SEER Extent of Disease
- 88: Not applicable
- 97: Other
- 99: Unknown

PLDEATH: State or country where the patient died.

*Refer to Volume I, Appendix D for values

http://ccrcal.org/Vol_1_May_2008_appendix_html/Vol_I_CA_Append.htm

PLDEATH_UPDT: State or country where the patient died updated from older death certificates.

PNODETU: Number of regional lymph nodes with evidence of involvement (positive):

Values:

- 00: All nodes examined negative
- 01-89: Actual number of nodes positive
- 90: Ninety or more nodes positive
- 95: Positive aspiration of lymph nodes
- 97: Positive nodes, but number unknown
- 98: No nodes examined (none removed)
- 99: Unknown if nodes are positive or negative, or unknown if examined

PROTOCOL: Code indicating agency or group that established the protocol in which the patient is participating:

Values:

- 00: Not applicable
- 01: NSABP
- 02: GOG
- 03: RTOG
- 04: SWOG
- 05: ECOG
- 06: POG
- 07: CCG
- 08: CALGB
- 09: NCI
- 10: ACS
- 11: National protocol, NOS
- 12: ACOS-OG
- 13: VA (Veterans Admin)
- 14: COG (Children's Oncology Group)
- 15: CTSU (Clinical Trials Support Unit)
- 16-50: National Trials
- 51-79: Locally defined
- 80: Pharmaceutical
- 81-84: Locally defined
- 85: In-house trial
- 86-88: Locally defined
- 89: Other
- 90-98: Locally defined
- 99: Unknown

Q:

QME: Name on Middle Eastern surname list? 1 = Yes.

QUINYOST: SES quintile of YOSTSCL score- based on Census variables, at the block group level. Ranges from 1 (low SES) to 5 (high SES).

R:

RACE1-RACE5: Race of patient- allows for patients to be multiple races (to coincide with the 2000 Census). Does not include Hispanic ethnicity, but does include Asian sub-groups:

Values for all 5 variables:

- 01: White
- 02: Black
- 03: American Indian
- 04: Chinese
- 05: Japanese
- 06: Filipino
- 07: Hawaiian
- 08: Korean
- 09: Asian Indian, Pakistani, Sri Lankan, Nepalese, Sikkimese, Bhutanese, or Bangladeshi
- 10: Vietnamese
- 11: Laotian
- 12: Hmong
- 13: Cambodian/Kampuchean
- 14: Thai
- 20: Micronesian, NOS
- 21: Chamorran
- 22: Guamanian, NOS
- 25: Polynesian, NOS
- 26: Tahitian
- 27: Samoan
- 28: Tongan
- 30: Melanesian, NOS
- 31: Fiji Islander
- 32: New Guinean
- 96: Other Asian, including Burmese, Indonesian, and Asian, NOS
- 97: Pacific Islander, NOS
- 98: Other
- 99: Unknown

RACE99: Race/ethnicity of patient- includes Hispanic ethnicity (Hispanics may be of any race):

Values:

- 1: Non-Hispanic White
- 2: Non-Hispanic Black
- 3: Hispanic
- 4: Asian/Pacific Islander
- 5: Non-Hispanic American Indian
- 9: Other/Unknown

RADBSTMOD: Records the dominant modality of radiation therapy used to deliver the most clinically significant **boost** dose to the primary volume of interest during the first course of treatment:

Values:

- 00: None, diagnosed at autopsy
- 20: External beam, NOS
- 21: Orthovoltage
- 22: Cobolt-60 & Cesium-137
- 23: Photons (2-5 MV)
- 24: Photons (6-10 MV)
- 25: Photons (11-19 MV)
- 26: Photons (>19 MV)
- 27: Photons (mixed energies)
- 28: Electrons
- 29: Photons and electrons mixed
- 30: Neutrons with or w/o photons/electrons
- 31: IMRT
- 32: Conformal or 3-D therapy
- 40: Protons
- 41: Stereotactic radiosurgery, NOS
- 42: Linac radiosurgery
- 43: Gamma knife
- 50: Brachytherapy, NOS
- 51: Brachytherapy, intracavitary, LDR
- 52: Brachytherapy, intracavitary, HDR
- 53: Brachytherapy, interstitial, LDR
- 54: Brachytherapy, interstitial, HDR
- 55: Radium
- 60: Radioisotopes, NOS
- 61: Strontium-89
- 62: Strontium-90
- 98: Other, NOS
- 99: Unknown

RADREGMOD: Records the dominant modality of radiation therapy used to deliver the most clinically significant **regional** dose to the primary volume of interest during the first course of treatment.

*Values are the same as for RADBSTMOD above.

RADSEQ: Indicates the sequence of radiation therapy with surgery (pre-op, post-op, etc.) during the first course of treatment:

Values:

- 0: Not applicable
- 2: Radiation before surgery
- 3: Radiation after surgery
- 4: Radiation before and after surgery
- 5: Intraoperative
- 6: Intraoperative + before and/or after
- 9: Both given, but sequence unknown

RADSUM: Did the first course of treatment include radiation therapy?

Values:

- 0: None
- 1: Beam
- 2: Implants
- 3: Isotopes
- 4: Combination of 1 with 2 or 3
- 5: Radiation, NOS
- 6: Implants/isotopes, NOS
- 9: Unknown

RANDOM: Was the census tract and block group randomly assigned?

RBCNSYS: SEER variable- treatment summary- radiation to CNS.

REGIONID: Regional Registry ID number (assigned by CCR):

Values:

- 1: Monterey, San Benito, Santa Clara, and Santa Cruz counties.
- 2: Fresno, Kern, Kings, Madera, Mariposa, Merced, Stanislaus, Tulare, and Tuolumne counties.
- 3: Alpine, Amador, Calaveras, El Dorado, Nevada, Placer, Sacramento, San Joaquin, Sierra, Solano, Sutter, Yolo, and Yuba counties.
- 4: San Luis Obispo, Santa Barbara, and Ventura counties.
- 5: Inyo, Mono, Riverside, and San Bernardino counties.
- 6: Butte, Colusa, Del Norte, Glenn, Humboldt, Lake, Lassen, Mendocino, Modoc, Napa, Plumas, Shasta, Siskiyou, Sonoma, Tehama, and Trinity counties.
- 7: San Diego and Imperial counties.
- 8: Alameda, Contra Costa, Marin, San Francisco, and San Mateo counties.
- 9: Los Angeles County.
- 10: Orange County.

REGPATNO: Unique patient ID assigned by Regional Registries- no longer in use (replaced by statewide PATIENT_ID).

REGTUMNO: An independent number assigned by Regional Registries to each tumor entered into the Region's database for a patient. No longer in use (replaced by statewide TUMOR_ID).

RELIGION: Patient's religion at time of diagnosis.

*Refer to Volume III, Appendix 19 for values:

http://www.ccrca.org/VOL3/Vol_3_CA.htm

RESPONSIBLE_REGION: Region responsible for the case.

RESREG: Region of residence for patient.

RUCA: Census variable- RUCA code.

RXCALC: SEER variable- first course calculated method.

RXDATE: Date any therapy started (MMDDCCYY format).

RXDATEC: Date chemotherapy started (MMDDCCYY format).

RXDATEH: Date hormone therapy started (MMDDCCYY format).

RXDATEI: Date immunotherapy started (MMDDCCYY format).

RXDATEO: Date other therapy started (MMDDCCYY format).

RXDATER: Date radiation therapy started (MMDDCCYY format).

RXDATESN: Date of diagnostic or staging procedure (MMDDCCYY format).

S:

SASDOB: Date of birth as a SAS date.

SASDX: Date of diagnosis as a SAS date.

SCOPE: Records surgery removing regional lymph nodes during the first course of treatment:

1. For 1988-2002 DX see Appendix Q of Volume I:

(http://www.ccrca.org/Vol_1/Appendices/Vol_I_CA_Append.htm)

2. For 2003+ DX see section VI.2.2 of Volume I:

(http://www.ccrca.org/Vol_1/Calif_Vol_1.htm)

SEER: SEER site recode- site group used for statistical reporting according to SEER's standards. Separates leukemias and lymphomas from site-specific cancers.

*Refer to SEER site for values:

http://seer.cancer.gov/siterecode/ICD-O3_d01272003/

SEQNOCEN: Indicates the sequence of all reportable neoplasms during the patient's lifetime as determined by the central registry. When two or more tumors are diagnosed simultaneously, the tumor with the worse prognosis is assigned the lowest sequence number:

Values:

- 00: This tumor is the only primary tumor for this patient
- 01: When a 2nd tumor is diagnosed, the 00 value from above becomes 01
- 02: 2nd tumor diagnosed in patient's lifetime
- 03-35: Actual number of this primary
- 99: Unspecified sequence number or unknown

State Registry defined:

- 60: Only one state registry defined neoplasm
- 61: First of two or more neoplasms (60 becomes a 61 in this case)
- 62: Second of two or more neoplasms
- 63-87: Actual number of state registry defined neoplasms
- 88: Unspecified number of state registry defined neoplasms

SEQNOHSP: Chronological order of this tumor among all the reportable tumors diagnosed during the patient's lifetime determined by the reporting facility. When two or more tumors are diagnosed simultaneously, the tumor with the worst prognosis is assigned the lowest sequence number:

Values:

- 00: This tumor is the only primary tumor for this patient
- 01: When a 2nd tumor is diagnosed, the 00 value from above becomes 01
- 02: 2nd tumor diagnosed in patient's lifetime
- 03-35: Actual number of this primary
- 99: Unspecified sequence number or unknown

State Registry/Cancer Committee Reportable:

- 60: Only one neoplasm required by the state or the cancer committee
- 61: First of two or more neoplasms in this category
- 62: Second of two or more neoplasms in this category
- 63-87: Actual number of neoplasms in this category
- 88: Unspecified number of neoplasms in this category

SEX: Sex of the patient

Values:

- 1: Male
- 2: Female
- 3: Hermaphrodite
- 4: Transsexual
- 9: Unknown

SITE_02: Location where this tumor originated in as much detail as is known and for which a code is provided in ICD-O-2 (please refer to ICD-O-2 for values).

SOURCE00: Field which tracks how and where year 2000 Census was performed (Census tract and block group). Two-digit field with codes as follows:

First digit (how geocoding was performed):

- 1: Batch computerized geocoding (automatic, non-interactive)
- 2: Interactive computerized geocoding (done with a software program, but an individual also makes decisions)
- 3: Manual (using source other than computer software program, such as maps, contacting local planning departments, etc)
- 9: Year 2000 geocodes not assigned

Second digit (where performed):

- 1: GDT (current vendor)
- 2: Teale (secondary vendor)
- 3: CCR
- 4: Region
- 9: Year 2000 geocodes not assigned

SPANISH: Used to denote those persons of Spanish origin- may be of any race:

Values:

- 0: Non-Spanish
- 1: Mexican
- 2: Puerto Rican
- 3: Cuban
- 4: Central or South American (except Brazil)
- 5: Other Spanish
- 6: Spanish, NOS
- 7: Spanish surname
- 8: Dominican Republic (starting with cases diagnosed 1/1/2005)
- 9: Unknown whether Spanish or not

SPANSURN: Used to denote those persons with Spanish surnames as recognized by the computer from tables of Spanish names. It conforms to the codes used by NAACCR for “computer-derived ethnicity”:

Values:

- 0: No match was run
- 1: Non-Hispanic last name and non-Hispanic maiden name
- 2: Non-Hispanic last name, but didn’t check maiden name (or male)
- 3: Non-Hispanic last name and missing maiden name
- 4: Hispanic last name and non-Hispanic maiden name
- 5: Hispanic last name, but didn’t check maiden name (or male)
- 6: Hispanic last name and missing maiden name
- 7: Hispanic maiden name (females only), regardless of Surname

SRCSUR: Source of Middle Eastern Surname:

Values:

- 1: Male surname
- 2: Female surname, no maiden name
- 3: Female maiden name

SSNOYEAR: Year of issue of SSNO.

STAGE: EOD for 1994-2003 diagnoses or collaborative stage for 2004+ diagnoses converted to AJCC stage:

Values:

- | | |
|------------------|--------------------|
| 00: 00 (in situ) | 30: III |
| 10: I | 31: IIIA |
| 11: IA | 32: IIIB |
| 12: IB | 33: IIIC |
| 13: IC | 39: III, NOS |
| 19: I, NOS | 40: IV |
| 20: II | 41: IVA |
| 21: IIA | 42: IVB |
| 22: IIB | 49: IV, NOS |
| 23: IIC | 90: Unstaged |
| 29: II, NOS | 98: Not applicable |
| | 99: Error |

STAGE_AJCC: Extent of Disease recoded to AJCC stage (1994-2004 cases, except for breast cancer, which is 1988-2004). Only available for a select number of cancer types:

Values:

00: 00 (in situ)	30: III
10: I	31: IIIA
11: IA	32: IIIB
12: IB	33: IIIC
13: IC	39: III, NOS
19: I, NOS	40: IV
20: II	41: IVA
21: IIA	42: IVB
22: IIB	49: IV, NOS
23: IIC	90: Unstaged
29: II, NOS	98: Not applicable
	99: Error

STAGE_AJCCSEER: EOD recoded to AJCC stage, modified by SEER.

*Criteria and values same as above.

STAGE_SEER: EOD for 1994-2003 diagnoses or collaborative stage for 2004+ diagnoses converted to AJCC stage as modified by SEER:

Values:

00: 00 (in situ)	30: III
10: I	31: IIIA
11: IA	32: IIIB
12: IB	33: IIIC
13: IC	39: III, NOS
19: I, NOS	40: IV
20: II	41: IVA
21: IIA	42: IVB
22: IIB	49: IV, NOS
23: IIC	90: Unstaged
29: II, NOS	98: Not applicable
	99: Error

STYPEFU: Type of follow-up, SEER.

SUMSTAGE: Summary stage at diagnosis- is a combination of:

1988-1993: SUMSTAGE77
1994-2000: SUMSTAGE77_SEER
2001-2003: SUMSTAGE00_SEER
2004-forward: DERIVED_SS2000

Values:

0: In situ
1: Localized
2: Regional by direct extension
3: Regional by lymph nodes
4: Regional by direct extension and lymph nodes
5: Regional, NOS
7: Remote
8: Not abstracted
9: Unknown or not specified

SUMSTAGE00: SEER summary stage at diagnosis, 2000 criteria (only valid for cases diagnosed 2001-2003). Values are the same as for SUMSTAGE.

SUMSTAGE00_SEER: EOD recode to SEER summary stage, 2000 criteria (only valid for cases diagnosed 2001-2003). Values are the same as for SUMSTAGE.

SUMSTAGE77: SEER summary stage at diagnosis, 1977 criteria (only valid for cases diagnosed 1988-2000). Values are the same as for SUMSTAGE.

SUMSTAGE77_SEER: EOD recode to SEER summary stage, 1977 criteria (only valid for cases diagnosed 1988-2000). Values are the same as for SUMSTAGE.

SURG_HOSP: Definitive surgery performed at reporting hospital:

Values:

00: No surgery
10-70: Definitive surgery to the primary site (the higher the value, the more definitive the surgery); values differ by cancer site
80: Definitive surgery to a metastatic site (not primary site)
90: Definitive surgery, unknown whether to a primary or metastatic site
99: Unknown whether or not definitive surgery was done

SURGDATE: Date definitive surgery was first performed (MMDDCCYY format).

SURGOTH: Surgical removal of tissue other than the primary tumor or organ of origin. See Appendix Q of Volume I:

http://ccrcal.org/Vol_1_May_2008_appendix_html/Vol_I_CA_Append.htm

SURGPRIM: Most extensive type of surgery performed during first course of treatment.
See Appendix Q of Volume I:

http://ccrcal.org/Vol_1_May_2008_appendix_html/Vol_I_CA_Append.htm

SURGRCON: Most extensive reconstructive surgery performed during first course of treatment.
See Appendix Q of Volume I:

http://ccrcal.org/Vol_1_May_2008_appendix_html/Vol_I_CA_Append.htm

SURGTYPE: Treatment summary- surgery type, SEER.

SURV_TIME: Survival time in months as computed by SEER based on a study end date of December 31, 2005.

T:

TCODE: EOD recode of T code for TNM.

TCODE_C: The T-code represents the primary tumor- clinical staging; definitions for values vary by cancer site (see AJCC Manual for Staging of Cancer).

TCODE_P: Pathological staging- definitions for values vary by cancer site (see AJCC Manual for Staging of Cancer).

TOTPOP: Census variable- total population in the census block group.

TRANSSUM: Identifies systemic therapeutic procedures given as first course of treatment at this facility and all other facilities or the reason why they were not used:

Values:

- 00: None; diagnosed at autopsy
- 10: Bone marrow transplant, NOS
- 11: Bone marrow transplant, autologous
- 12: Bone marrow transplant, allogenic
- 20: Stem cell harvest and infusion
- 30: Endocrine surgery and/or endocrine radiation therapy
- 40: Code 30 in combination with 10, 11, 12, or 20
- 82: Contraindicated
- 85: Patient died
- 86: Recommended, not given
- 87: Refused
- 88: Recommended, unknown if given
- 99: Unknown; death certificate only

TSIZETU: Extent of disease variable- size of tumor in millimeters.

TUMOR_ID: CCR unique ID at the tumor level (statewide).

TYPEADM: Type of admission to hospital:

Values:

- 1: Inpatient
- 2: Outpatient
- 3: Tumor Board
- 4: Path lab
- 5: Inpatient and outpatient
- 6: Inpatient and tumor board
- 7: Outpatient and tumor board
- 8: Inpatient, outpatient, and tumor board

TYPERP: Best source of information to prepare this case abstract:

Values:

- 1: Hospital (inpatient) or clinic
- 2: Radiation treatment centers or medical oncology centers (for cases diagnosed 2006 and later)
- 3: Laboratory only
- 4: Physician's office/private medical practitioner
- 5: Nursing/convalescent home/hospice
- 6: Autopsy only
- 7: Death certificate only
- 8: Other hospital outpatient units/surgery centers (for cases diagnosed 2006 and later)

V, X, and Y:

VIETSURN: Vietnamese surname? (1 = yes)

VITSTAT: Vital status (0 = dead and 1 = alive).

VITSTAT_REC: Vital status (0 = dead and 1 = alive) as of December 31, 2005.

XNODETU: Number of regional lymph nodes examined by the pathologist

Values:

00: No regional lymph nodes examined

01-89: Actual number examined

90: Ninety or more examined

95: No regional lymph nodes removed, but aspiration of regional lymph nodes was performed

96: Regional lymph node removal documented as a sampling and number of lymph nodes unknown/not stated

97: Regional lymph node removal documented as dissection and number of lymph nodes unknown/not stated

98: Regional lymph nodes surgically removed but number of lymph nodes unknown/not stated and not documented as sampling or dissection

99: Unknown; not stated; death certificate only

YEARDX: Year of diagnosis.

YOSTSCL: Derived from Census variables- an SES scale from principal components analysis.