FILENAME: CAHPS.pedsf.cancer.txt.gz
LRECL: 4856

CAHPS Survey cohort 1997-2015

PEDSF for SEER-CAHPS created June 5, 2019

SEER Cases diagnosed 1973-2013
SEER data extracted from November, 2016 SEER Submission

ABSTRACT: The PEDSF file contains one record per person for individuals in the SEER
                 Program data base who have been matched with Medicare enrollment records.
                 Of persons who were reported by the SEER Registries to have been diagnosed
                 with cancer at age 65 or older, 94 percent were matched with Medicare
                 enrollment records. For persons appearing in the PEDSF file, basic SEER
                 Program diagnostic information is available for up to 10 diagnosed cancer
                 occurrences. Data also include Medicare entitlement and utilization data from
                 1996-2015.

PROGRAMMER: Angela Meekins, Information Management Services, Inc.

CONTACT: SEER-Medicare Technical Support, Information Management Services, Inc.
Email: SEER-Medicare@imsweb.com

NOTE1: This PEDSF file is not linkable to non-CAHPS files as the protected IDs are different.

NOTE2: For all non-custom requests, patients have been selected if the cancer(s) requested
appeared in the site array (siterwho1-10) and the associated year of diagnosis (yrdx1-10) fell in
the year of diagnosis window requested.

COL	FIELD	LENGTH	SOURCE	NOTES
1	Protected ID **(phic)**	11	IMS created	Protected ID for patients
12	Match Type **(mat_type)**	1	Match File	CMSs matching algorithm is available at IMS. A = Alphabetic Match B = SSN Match
13	Match Flag 1 **(yobflg1)**	1	Match File	Year of Birth flag for SSN and Alpha Match. Y = Agreement between SEER and Medicare N = Did not agree 0 (Zero) = Cases matched on SSN U = Unknown, Missing data
14	Match Flag 2 **(dobflg2)**	1	Match File	Day of Birth Match for SSN and Alpha Match (See Match Flag 1 for code definition)
15	Match Flag 3 **(yodflg3)**	1	Match File	Month and Year of Death for Alpha Match. (See Match Flag 1 for code definition)
17	Match Flag 5 **(lstflg5)**	1	Match File	Last Name Flag for SSN Match. (See Match Flag 1 for code definition)
18	Match Flag 6 **(fstflg6)**	1	Match File	First Name Flag for SSN Match. (See Match Flag 1 for code definition)
19	Match Flag 7 **(mobflg7)**	1	Match File	Month of Birth Flag for SSN Match. (See Match Flag 1 for code definition)
20	Match Flag 8 **(midflg8)**	1	Match File	Middle Initial Flag for Alpha Match. (See Match Flag 1 for code definition)
21	Number of digits in SSN that matched **(numdigit)**	1	Match File	Alpha search only. 0-8 9 = SSN match U = Unknown; data missing A = digits transposed and matched
22	Date of Death Flag **(dod_flg)**	1	Created at IMS	Shows the level of agreement between SEER and MEDICARE on the patient's month of death. Dates of death after 12/14 were treated as not dead for comparison purposes. 0 = Not dead by 12/14. 1 = Dead, both files agree.

COL	FIELD	LENGTH	SOURCE	NOTES
				2 = Dead, off by 1-3 months.
				3 = Dead, off by 4-6 months.
				4 = Dead in MEDICARE only,
				5 = Dead in SEER only.
				6 = Dead but number of months could not be calculated due to the fact that month was missing for either SEER or MEDICARE.
23	Date of Birth Flag **(dob_flg)**	1	Created at IMS	Shows the level of agreement between SEER and MEDICARE on the patient's month of birth. 0 = Both files agree on birth date. 1 = Birth date off by 1-3 months. 2 = Birth date off by 4-6 months. 3 = Birth date off by 7-11 months. 4 = Birth date off by one year 5 = Birth date off by 13-23 months 6 = Birth date off by 2 years 7 = Birth date off by 25-35 months 8 = Birth date off by 3 years 9 = Birth date off by 37+ months Blank = Birth date is missing
25	Medicare Date of Death **(med_dodm, med_dodd, med_dody)**	8	EDB	MMDDYYYY or blanks for not dead. Reported to CMS by SSA – not date of last contact. Date complete through 12/31/16
33	Medicare Month of Birth **(birthm)**	2	EDB	MM, zeroes for invalid
35	Medicare Year of Birth **(birthyr)**	4	EDB	YYYY, zeroes for invalid
39	Sex **(m_sex)**	1	EDB	1 = Male 2 = Female
40	Race **(race)**	1	EDB	1 = White 2 = Black 3 = Other 4 = Asian 5 = Hispanic 6 = N. Am. Native 0 = Unknown

COL	FIELD	LENGTH	SOURCE	NOTES
41	Original Reason for Entitlement **(rsncd1)**	1	EDB	0 = OASI (age) 1 = DIB (disability) 2 = ESRD 3 = DIB/ESRD This is the first entitlement reason located on EDB
42	Current Reason for Entitlement **(cur_ent)**	1	EDB	Blank = Not Enrolled 0 = OASI (age) 1 = DIB 2 = ESRD 3 = DIB/ESRD Variable is selected by the first diagnosis date at age 65 or older, or the last diagnosis if all are prior to turning age 65. If information not available at those times then the first available data is used.
43	Current Reason for Entitlement Year **(cur_yr)**	4	EDB	YYYY
47	Chronic Renal Disease **(chr_esrd)**	1	EDB	0 (zero) = No ESRD Y = Source of ESRD Variable is set to zero. If there is a start date before the diagnosis date at age 65 or older, or the last diagnosis if all are prior to turning age 65 then it is set to Y.
48	Chronic Renal Disease Year **(chr_esrd_yr)**	4	EDB	YYYY
52	First Chronic Renal Disease Year **(first_esrd_yr)**	4	EDB	First occurrence of Chronic Renal Disease regardless of age.
56	Medicare Status Code **(med_stcd)**	2	EDB	Blank = Not Enrolled 10 = Aged 11 = Aged with ESRD 20 = Disabled 21 = Disabled with ESRD 31 = ESRD only Variable is selected by the first diagnosis date at age 65 or older, or the last diagnosis if all are prior to turning age 65. If information not available at those times

COL	FIELD	LENGTH	SOURCE	NOTES
				then the first available data is used.
58	Medicare Status Code Year **(medst_yr)**	4	EDB	YYYY

Geographic Information

SEER program data are from the SEER record corresponding to the first diagnosis of cancer at age 65 or older. Zip codes are from Medicare enrollment file in the year of first diagnosis at age 65 or older or the last diagnosis if never 65.

COL	FIELD	LENGTH	SOURCE	NOTES
62	State **(state)**	2	SEER	FIPS State Code
64	County **(county)**	3	SEER	FIPS County Code
67	Census Track Flag **(code_sys)**	1	SEER	Use in conjunction with Census Tract (tract) 0 = Not tracted 1 = 1970 Census tract definitions 2 = 1980 Census tract definitions 3 = 1990 Census tract definitions
68	Census Tract **(enctrt90)**	6	SEER	1970/80/90 Encrypted census tract *** Special permission required for unencrypted census tract.**
74	Census Tract 2000 **(enctrt00)**	6	SEER	2000 encrypted census tract *** Special permission required for unencrypted census tract.**
80	Census Tract 2010 **(enctrt10)**	6	SEER	2010 encrypted census tract *** Special permission required for unencrypted census tract.**
86	HSA **(hsa)**	3	ARF	Health Service Area. Taken from the 2004 Area Resource File (ARF).
89	Urban/Rural recode **(urbrur)**	1	ARF	Urban/Rural Code 1 = Big Metro (Urban = 00 or 01) 2 = Metro (Urban = 02 or 03) 3 = Urban (Urban = 04 or 05) 4 = Less Urban (Urban = 06 or 07) 5 = Rural (Urban = 08 or 09) 9 = Unknown (Urban = 99)
90	Urban/Rural code **(urban)**	2	ARF	01-09, 99 see table at end of this document

Demographic Information

The following information comes from the SEER record corresponding to the first diagnosis

COL	FIELD	LENGTH	SOURCE	NOTES
	at age 65 or over, or the latest diagnosis if all are prior to turning age 65.			
92	Sex **(s_sex)**	1	SEER	See Attachment A
95	Race Recode Y **(rac_recy)**	1	SEER	See Attachment A
96	Race Recode A **(rac_reca)**	1	SEER	See Attachment A
98	ICD Code – Cause of Death **(icd_code)**	1	SEER	0 = Patient is alive at last follow-up 1 = Tenth ICD revision 8 = Eighth ICD Revision 9 = Ninth ICD revision
99	Cause of Death ICD -8 or 9 **(cod89v)**	4	SEER	0000 = Alive at last contact 7777 = State Death Certificate not available. 7797 = Death Certificate available, no COD listed.
103	Cause of Death ICD-10 **(cod10v)**	4	SEER	A020-Y891 0000 = Alive at last contact 7777 = State Death Certificate not available. 7797 = Death Certificate available, no COD listed.
107	Cause of Death to site recode KM **(codkm)**	5	SEER	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. Study cutoff date has been applied, i.e. coded as alive if death occurred after study cutoff. Go to the end of this document for a listing of codes and their definition.
112	Cause of Death to site recode **(codpub)**	5	SEER	This recode was introduced to account for several newly valid ICD-10 codes and includes both cancer and non-cancer causes of death. Go to the end of this document for a listing of codes and their definition.
117	NHIA Derived Hispanic Origin **(nhiade)**	1	SEER	See Attachment A
118	SEER Month of Death	2	SEER	00 = Alive

COL	FIELD	LENGTH	SOURCE	NOTES
	(ser_dodm)			blank = Unknown month Date complete through 12/31/14
120	SEER Year of Death (ser_dody)	4	SEER	0000 = Alive 2053 = Unknown year Date complete through 12/31/14
124	SEER Race/Ethnicity (srace)	2	SEER	See Attachment A
127	Origin recode NHIA (Hispanic, Non-Hisp) (origrecb)	1	SEER	See Attachment A
128	Vital Status Recode (stat_rec)	1	SEER	See Attachment A
129	Census Tract Certainty (ctcert90)	1	SEER	Associated with 1970/80/90 Census data 1 = Census tract based on complete and valid street address of residence 2 = Census tract based on residence ZIP+4 3 = Census tract based on residence ZIP+2 4 = Census tract based on residence ZIP code only 5 = Census tract based on ZIP code of post office box 6 = Census tract/BNA based on residence city where city has only one census tract, or based on residence ZIP code where ZIP code has only one census tract 9 = Unable to assign census tract based on available information
130	Census Tract Certainty 2000 (ctcert00)	1	SEER	Associated with 2000 Census data 1 = Census tract based on complete and valid street address of residence 2 = Census tract based on residence ZIP+4 3 = Census tract based on residence ZIP+2 4 = Census tract based on residence ZIP code only 5 = Census tract based on ZIP code of post office box

COL	FIELD	LENGTH	SOURCE	NOTES
				6 = Census tract/BNA based on residence city where city has only one census tract, or based on residence ZIP code where ZIP code has only one census tract
				9 = Unable to assign census tract based on available information
131	Census Tract Certainty 2010 **(ctcert10)**	1	SEER	Associated with 2010 Census data
				1 = Census tract based on complete and valid street address of residence
				2 = Census tract based on residence ZIP+4
				3 = Census tract based on residence ZIP+2
				4 = Census tract based on residence ZIP code only
				5 = Census tract based on ZIP code of post office box
				6 = Census tract based on residence city where city has only one census tract, or based on residence ZIP code where ZIP code has only one census tract
				9 = Unable to assign census tract based on available information
132	Census Tract Poverty Indicator **(census_pov_ ind)**	1	SEER	1= 0%-<5% poverty
				2 = 5% to <10% poverty
				3 = 10% to <20% poverty
				4 = 20% to 100% poverty
				9 = Unknown
133	SEER Year of Birth **(yr_brth)**	4	SEER	See Attachment A
137	Date of Birth Flag **(dbrflag)**	2	SEER	12 = A proper value is applicable but not known.
139	Number of SEER records **(count)**	2	Created at IMS	Number of eligible SEER records. SEER records prior to 1992 for Rural Georgia and records prior to 2000 for Greater California, New Jersey, Kentucky, and Louisiana were not included in PEDSF.
141	Diagnosis indicator **(resnrec)**	1	Created at IMS	0 = Last Dx; Patient always less than 65.
				1 = First Dx at age 65 or later.

COL	FIELD	LENGTH	SOURCE	NOTES
142	Last survey year **(year_last_svy)**	4	Created at IMS	Year of last survey patient participated in. 1997-2005, 2007-2015

Repeated Part D Denominator Information
These variables (columns 146-185) are repeated once for every year from 2007 to 2015. (YY = the year of the file, ex. 07)

COL	FIELD	LENGTH	SOURCE	NOTES
146	Plan-Value Indicators (Jan. – Dec.) **(planYY_01- planYY_12)**	12*1	Part D Denom	12 monthly 1-byte indicators. Indicates for each month of the Denominator reference year, the Part D enrollment, which is based on the 1st digit or the Part D contract number.

Recodes only indicating type of plan; no 5-digit plan numbers.
0 = Not Medicare enrolled for the month
* = Enrolled in Medicare A and/or B, but no MIIR record for the month
H = Managed Care Organizations other than Regional PPO
R = Regional PPO
S = PDP
N = Not Part D Enrolled
E = Employer-sponsored (starting in Jan. 2007)
X = Limited Income Newly Eligible Transition plan (LINET, starting July 2009)

COL	FIELD	LENGTH	SOURCE	NOTES
158	State Reported Dual Eligible Status Code (Jan. – Dec.) **(dualYY_01- dualYY_12)**	12*2	Part D Denom	12 monthly 2-byte indicators. Indicates for each month of the Denominator reference year, the dual eligibility status, if any, for the beneficiary.

00 = Not Medicare enrolled for the month
** = Enrolled in Medicare A and/or B, but no MIIR record for the month
NA = Non-Medicaid
01 = QMB only
02 = QMB and Medicaid coverage including RX
03 = SLMB only
04 = SLMB and Medicaid coverage including RX
05 = QDWI
06 = Qualifying Individuals
08 = Other Dual Eligibles (Non-QMB, SLMB,

COL	FIELD	LENGTH	SOURCE	NOTES
				QWDI, or QI) w/Medicaid coverage including RX 09 = Other Dual Eligibles but without Medicaid coverage 99 = Unknown
182	Plan Coverage Months **(ptdYY)**	2	Part D Denom	Contains the total number of months of Part D plan coverage for the beneficiary. The value in this field will be within the valid range of values 00 through 12, inclusive, dependent on the number of occurrences when the Plan Indicators = H, R, S, or E.
184	Medicaid Dual Eligible Months **(dualcntYY)**	2	Part D Denom	Contains the total numbers of months of dual eligibility for the beneficiary. The value in this field will be within the valid range of values 00 through 12, inclusive, dependent on the number of occurrences when the Medicaid Dual Eligible Indicators not equal to 00, **, NA, or 99.

Repeated Entitlement Information

These variables (columns 506-549) are repeated once for every year from 1996 to 2015.

COL	FIELD	LENGTH	SOURCE	NOTES
506	Entitlement Indicators **(ent1-ent240)**	12*1	EDB	One indicator for each month in the year. 0 = Not entitled 1 = Part A only 2 = Part B only 3 = Part A and B Note: ent1 = 1/1996, ent2 = 2/1996, ent3 = 3/1996 ….ent240 = 12/2015

COL	FIELD	LENGTH	SOURCE	NOTES
518	HMO Indicators **(hmo1-hmo240)**	12*1	EDB	One indicator for each month in the year. 0 = Not a member of HMO 1 = Non-Lock-in, CMS to process provider claims 2 = Non-Lock-in, GHO to process in-plan Part A & in-area Part B claims 4 = Chronic care disease management organizations-FFS plan 5 = ? A = Lock-in, CMS to process provider claims B = Lock-in, GHO to process in-plan Part A & in-area Part B claims C = Lock-in, GHO to process all Part A and Part B claims Note: hmo1 = 1/1996, hmo2 = 2/1996, hmo3 = 3/1996 hmo240 = 12/2015
530	State Buy-in Indicator **(stbuy1-stbuy240)**	12*1	EDB	One indicator for each month in the year. 0 = Not entitled 1 = Part A only 2 = Part B only 3 = Part A and B Note: stbuy1 = 1/1996, stbuy2 = 2/1996, stbuy3 = 3/1996 ….stbuy240 = 12/2015
542	Part A Months **(ptacnt1996-ptacnt2015)**	2	EDB	Number of months covered for part A in each year covering 1996-2015. 00 = No coverage that year 01-12 = Number of months of coverage
544	Part B Months **(ptbcnt1996-ptbcnt2015)**	2	EDB	Number of months covered for part B in each year covering 1996-2015. 00 = No coverage that year 01-12 = Number of months of coverage
546	HMO Months **(hmocnt1996-hmocnt2015)**	2	EDB	Number of months as an HMO Member in each year covering 1996-2015. 00 = No coverage that year 01-12 = Number of months of coverage
548	State Buy-in Months **(stbuycnt1996-stbuycnt2015)**	2	EDB	Number of months with state buy-in coverage In each year covering 1996-2015. 00 = No coverage that year 01-12 = Number of months of coverage

COL	FIELD	LENGTH	SOURCE	NOTES
Repeated Zip Code and Registry Information				
Zip Code and Registry (based on state and county) information has been retained from the EDB file for each year from 1996 to 2015.				
1386	State **(state1996-state2015)**	2	EDB	State code from EDB file.
1388	County **(cnty1996-cnty2015)**	3	EDB	County code from EDB file.
1391	Encrypted Zip Codes (5) **(enzip5_1996-enzip5_2015)**	5	EDB	Encrypted Zip code (5 digits) is assigned as the last zip code the patient resided in that year. ***Special permission required for unencrypted zip codes.**
1396	Zip Codes (4) **(enzip4_1996-enzip4_2015)**	4	EDB	Last 4 digits of zip code. Blanked out when the zip code is encrypted. ***Special permission required for unencrypted zip codes.**
1400	Urban/Rural Code **(urban1996-urban2015)**	2	ARF	01-09, 99 see attachment at end of this document
1402	Urban/Rural Recode **(urbrur1996-urbrur2015)**	1	IMS created	1 = Big Metro (Urban = 00 or 01) 2 = Metro (Urban = 02 or 03) 3 = Urban (Urban = 04 or 05) 4 = Less Urban (Urban = 06 or 07) 5 = Rural (Urban = 08 or 09) 9 = Unknown (Urban = 99)

COL	FIELD	LENGTH	SOURCE	NOTES
1403	Registry Code **(registry1996-registry2015)**	2	EDB	01 = San Francisco 02 = Connecticut 20 = Detroit 21 = Hawaii 22 = Iowa 23 = New Mexico 25 = Seattle 26 = Utah 27 = Atlanta 31 = San Jose 35 = Los Angeles 37 = Rural Georgia 41 = Greater California 42 = Kentucky 43 = Louisiana 44 = New Jersey 47 = Greater Georgia NA = Has not lived in a registry yet **NOTE:** **First registry found is retained until patient moved into a new registry.**
1405	Registry Code #2 **(reg2cd1996-reg2cd2015)**	2	EDB	01 = San Francisco 02 = Connecticut 20 = Detroit 21 = Hawaii 22 = Iowa 23 = New Mexico 25 = Seattle 26 = Utah 27 = Atlanta 31 = San Jose 35 = Los Angeles 37 = Rural Georgia 41 = Greater California 42 = Kentucky 43 = Louisiana 44 = New Jersey 47 = Greater Georgia NA = Other **NOTE:** **Registry is based off the state and county, which is the last residence the patient lived at for that year.**

COL	FIELD	LENGTH	SOURCE	NOTES

Repeated SEER Cancer Information

A patient can have up to 10 diagnoses in SEER, so the below variables are repeated for each cancer diagnosis the patient has in SEER. If the patient has more than 10 only the first 10 are retained for this file.
In parenthesis you will find the variable names. In italics you will find the SEER variable name.

COL	FIELD	LENGTH	SOURCE	NOTES
1806	SEER Registry Code at Diagnosis **(reg1-reg10)** *(reg)*	2	SEER	01 = San Francisco (1973+) 02 = Connecticut (1973+) 20 = Detroit (1973+) 21 = Hawaii (1973+) 22 = Iowa (1973+) 23 = New Mexico (1973+) 25 = Seattle (1974+) 26 = Utah (1973+) 27 = Atlanta (1975+) 31 = San Jose (1992+) 35 = Los Angeles (1992+) 37 = Rural Georgia (1992+) 41 = Greater California (2000+) 42 = Kentucky (2000+) 43 = Louisiana (2000+) 44 = New Jersey (2000+) 47 = Greater Georgia (2000+)
1808	Marital Status at Diagnosis **(marst1-marst10)** *(mar_stat)*	1	SEER	See Attachment A
1809	Age at Diagnosis **(agedx1-agedx10)** *(age_dx)*	3	SEER	See Attachment A
1812	Sequence Number **(seq1-seq10)** *(seq_num)*	2	SEER	See Attachment A
1814	Month of Diagnosis **(modx1-modx10)** *(date_mo)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1816	Year of Diagnosis **(yrdx1-yrdx10)** *(date_yr)*	4	SEER	See Attachment A
1820	Primary Site **(site1-site10)** *(site02v)*	3	SEER	See Attachment A Note: This variable is length 4 in the SEER file. The first character (which is always 'C') was not written to the PEDSF.
1823	Laterality **(lat1-lat10)** *(lateral)*	1	SEER	See Attachment A
1824	Histology ICD-O-2 (1992-2000) **(hist2_1-hist2_10)** *(hist02v)*	4	SEER	See Attachment A
1828	Behavior code ICD-O-2 (1992-2000) **(beh2_1-beh2_10)** *(beh02v)*	1	SEER	See Attachment A
1829	Histology ICD-O-3 **(hist1-hist10)** *(hist03v)*	4	SEER	See Attachment A
1833	Behavior code ICD-O-3 **(beh1-beh10)** *(beh03v)*	1	SEER	See Attachment A
1834	Grade **(grade1-grade10)** *(grade)*	1	SEER	See Attachment A
1835	Diagnostic Confirmation **(dxconf1-dxconf10)** *(dx_conf)*	1	SEER	See Attachment A
1836	Type of Reporting Source **(src1-src10)** *(rept_src)*	1	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1837	EOD 10 - Tumor Size (1988-2003) **(e10sz1-e10sz10)** *(eod10_sz)*	3	SEER	See Attachment A
1840	EOD 10 – Tumor Extent (1988-2003) **(e10ex1-e10ex10)** *(eod10_ex)*	2	SEER	See Attachment A
1842	EOD 10 – Prostate pathology ext (1995-2003) **(e10pe1-e10pe10)** *(eod10_pe)*	2	SEER	See Attachment A
1844	EOD 10 - Nodes (1988-2003) **(e10nd1-e10nd10)** *(eod10_nd)*	1	SEER	See Attachment A
1845	EOD 10 - # positive Nodes (1988+) **(e10pn1-e10pn10)** *(eod10_pn)*	2	SEER	See Attachment A
1847	EOD 10 - # Nodes Examined (1988+) **(e10ne1-e10ne10)** *(eod10_ne)*	2	SEER	See Attachment A
1849	EOD – OLD 13 Digit (1973-1982) **(eod13_1-eod13_10)** *(eod13)*	13	SEER	See Attachment A
1862	EOD – OLD 2 Digit (1973-1982) **(eod2_1-eod2_10)** *(eod2)*	2	SEER	See Attachment A
1864	EOD 4 –size (1983-1987) **(e4siz1-e4siz10)** *(eod4)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1866	EOD 4 – extent (1983-1987) **(e4ext1-e4ext10)** *(eod4)*	1	SEER	See Attachment A
1867	EOD 4 – nodes (1983-1987) **(e4nod1-e4nod10)** *(eod4)*	1	SEER	See Attachment A
1868	Coding system – extent of disease (1973-2003) **(eod_cd1-eod_cd10)** *(eodcode)*	1	SEER	See Attachment A
1869	Tumor Marker 1 (1990-2003) **(tumor1_1-tumor1_10)** *(tumor_1v)*	1	SEER	See Attachment A
1870	Tumor Marker 2 (1990-2003) **(tumor2_1-tumor2_10)** *(tumor_2v)*	1	SEER	(Prostate data has been removed from this field) See Attachment A
1871	Tumor Marker 3 (1998-2003) **(tumor3_1-tumor3_10)** *(tumor_3v)*	1	SEER	See Attachment A
1872	CS Tumor Size (2004+) **(cstum1-cstum10)** *(cs_size)*	3	SEER	See Attachment A
1875	CS Extension (2004+) **(csex1-csex10)** *(cs_ext)*	3	SEER	See Attachment A
1878	CS Lymph Nodes (2004+) **(cslym1-cslym10)** *(cs_node)*	3	SEER	See Attachment A
1881	CS Mets at Dx (2004+) **(csmet1-csmet10)** *(cs_mets)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1883	CS Site-Specific Factor 1 (2004+) **(cs1st1-cs1st10)** *(cs_ssf1)*	3	SEER	See Attachment A
1886	CS Site-Specific Factor 2 (2004+) **(cs2st1-cs2st10)** *(cs_ssf2)*	3	SEER	(Prostate data has been removed from this field) See Attachment A
1889	CS Site-Specific Factor 3 (2004+) **(cs3st1-cs3st10)** *(cs_ssf3)*	3	SEER	See Attachment A
1892	CS Site-Specific Factor 4 (2004+) **(cs4st1-cs4st10)** *(cs_ssf4)*	3	SEER	See Attachment A
1895	CS Site-Specific Factor 5 (2004+) **(cs5st1-cs5st10)** *(cs_ssf5)*	3	SEER	See Attachment A
1898	CS Site-Specific Factor 6 (2004+) **(cs6st1-cs6st10)** *(cs_ssf6)*	3	SEER	See Attachment A
1901	CS Site-Specific Factor 25 (2004+) **(cs25st1-cs25st10)** *(cs_ssf25)*	3	SEER	See Attachment A
1904	Derived AJCC T, 6th edition (2004+) **(dajcct1-dajcct10)** *(d_ajcc_t)*	2	SEER	See Attachment A
1906	Derived AJCC N, 6th edition (2004+) **(dajccn1-dajccn10)** *(d_ajcc_n)*	2	SEER	See Attachment A
1908	Derived AJCC M, 6th edition (2004+) **(dajccm1-dajccm10)** *(d_ajcc_m)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1910	Derived AJCC Stage Group, 6th edition (2004+) **(dajccstg1-dajccstg10)** *(d_ajcc_s)*	2	SEER	See Attachment A
1912	Derived SS1977 (2004+) **(dss77s1-dss77s10)** *(d_ssg77)*	1	SEER	See Attachment A
1913	Derived SS2000 (2004+) **(dss00s1-dss00s10)** *(d_ssg00)*	1	SEER	See Attachment A
1914	Derived AJCC-Flag **(dajcflg1-dajcflg10)** *(d_ajcc_f)*	1	SEER	See Attachment A
1915	Derived SS1977 - Flag **(dss77f1-dss77f10)** *(d_ssg77f)*	1	SEER	1 = SS1977 derived from Collaborative Staging Manual and Coding Instructions, Version 1.0 2 = SS1977 derived from EOD (prior to 2004)
1916	Derived SS2000 – Flag **(dss00f1-dss00f10)** *(d_ssg00f)*	1	SEER	1 = SS2000 derived from Collaborative Staging Manual and Coding Instructions, Version 1.0 2 = SS2000 derived from EOD (prior to 2004)
1917	CS Version 1st (2004+) **(csvf1-csvf10)** *(csv_org)*	6	SEER	See Attachment A
1923	CS Version Latest (2004+) **(csvl1-csvl10)** *(csv_der)*	6	SEER	See Attachment A
1929	CS version input current (2004+) **(cscurrent1-cscurrent10)** *(csv_cur)*	6	SEER	See Attachment A
1935	RX Summ – Surgery of Primary Site (1998+) **(sxprif1-sxprif10)** *(surgprim)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1937	RX –Summ Scope Reg LN Sur (2003+) **(sxscof1-sxscof10)** *(scope)*	1	SEER	See Attachment A
1938	RX Summ--Surg Oth Reg/Dis (2003+) **(sxsitf1-sxsitf10)** *(surgoth)*	1	SEER	See Attachment A
1939	Number of regional lymph nd exam (1998-2002) **(numnd1-numnd10)** *(surgnode)*	2	SEER	See Attachment A
1941	Reason no cancer-directed surgery **(nosrg1-nosrg10)** *(no_surg)*	1	SEER	See Attachment A
1942	Radiation **(rad1-rad10)** *(radiatn)*	1	SEER	See Attachment A
1943	Radiation to Brain and /or CNS **(radbrn1-radbrn10)** *(rad_brn)*	1	SEER	See Attachment A
1944	Radiation sequence with surgery **(radsurg1-radsurg10)** *(rad_surg)*	1	SEER	See Attachment A
1945	Site Specific Surgery (1983 – 1997) **(sssurg1-sssurg10)** *(ss_surg)*	2	SEER	See Attachment A
1947	Scope of regional lymph node surgery (1998-2002) **(sxscop1-sxscop10)** *(scope02)*	1	SEER	See Attachment A
1948	Surgery of other reg/dist sites (1998-2002) **(sxsite1-sxsite10)** *(srgoth02)*	1	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1951	Age-site edit override **(ositage1-ositage10)** *(o_sitage)*	1	SEER	1 = Reviewed and confirmed that age/site/histology combination is correct as coded 2 = Reviewed and confirmed that case was diagnosed in utero 3 = Reviewed and confirmed that conditions 1 and 2 apply
1952	Sequence number-dx conf override **(oseqcon1-oseqcon10)** *(o_seqcon)*	1	SEER	1 = Reviewed: Multiple primaries of special sites in which at least one diagnosis has not been microscopically confirmed have been reviewed.
1953	Site-type-lat-seq override **(oseqlat1-oseqlat10)** *(o_seqlat)*	1	SEER	1 = Reviewed: Multiple primaries of the same histology (3 digit) in the same primary site group have been reviewed.
1954	Surgery-diagnostic conf override **(osurcon1-osurcon10)** *(o_surcon)*	1	SEER	1 = Reviewed: A patient who had (cancer directed) surgery, but the tissue removed was not sufficient for microscopic confirmation.
1955	Site-type edit override **(osittyp1-osittyp10)** *(o_sittyp)*	1	SEER	1 = Reviewed: The coding of an unusual combination of primary site and histologic type has been reviewed.
1956	Histology edit override **(hbenign1-hbenign10)** *(h_benign)*	1	SEER	1 = Unusual Histology behavior combination reviewed. 2 = Unusual Diagnostic Confirmation/Behavior combination reviewed. 3 = Both scenarios above reviewed
1957	Report source sequence override **(orptsrc1-orptsrc10)** *(o_rptsrc)*	1	SEER	1 = Reviewed: A second or subsequent primary with a reporting source of death certificate only has been reviewed and is indeed an independent primary.
1958	Seq-ill-defined site override **(odfsite1-odfsite10)** *(o_dfsite)*	1	SEER	1 = Reviewed: A second or subsequent primary reported with an ill-defined primary site (C76.0 C76.8, C80.9) has been reviewed and is an independent primary.
1959	Leuk-Lymph dx confirmation override **(oleukdx1-oleukdx10)** *(o_leukdx)*	1	SEER	1 = Reviewed: A patient was diagnosed with leukemia or lymphoma and the diagnosis was not microscopically confirmed.

COL	FIELD	LENGTH	SOURCE	NOTES
1960	Site-behavior override **(ositbeh1-ositbeh10)** *(o_sitbeh)*	1	SEER	1 = Reviewed: A patient has an in situ cancer of a nonspecific site and no further information about the primary site is available. Note: 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. This over-ride indicates that the conflict has been reviewed.
1961	Site-EOD-dx date override **(oeoddt1-oeoddt10)** *(o_eoddt)*	1	SEER	1 = Reviewed: A patient had "localized" disease with a non-specific site and no further information about the primary site is available. Note: 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.
1962	Site-laterality-EOD override **(ositeod1-ositeod10)** *(o_siteod)*	1	SEER	1 = Reviewed: A patient had laterality coded non-specifically and EOD coded specifically. Note: 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.
1963	Site-laterality-morph override **(ositmor1-ositmor10)** *(o_sitmor)*	1	SEER	1 = Reviewed: A patient had behavior code of "in situ" and laterality is not stated as right: origin of primary; left: origin of primary; or only one side involved, right or left origin not specified. Note: 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.
1964	Type of follow-up expected **(typefu1-typefu10)** *(typefup)*	1	SEER	See Attachment A
1965	Age at Diagnosis recode **(ager1-ager10)** *(age_rec)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1967	Site Recode ICD-O-3/ WHO 2008 **(siterwho1-siterwho10)** *(siterwho)*	5	SEER	See Attachment A ***This variable is used to select patients for data requests.**
1972	Recode ICD-O-2 to 9 **(icdot09_1-icdot09_10)** *(icdot09v)*	4	SEER	See Attachment A
1976	Recode ICD-O-2 to 10 **(icdot10_1-icdot10_10)** *(icdot10v)*	4	SEER	See Attachment A
1980	ICCC site recode ICD-O-3/WHO 2008 **(ICCC3WHO1-ICCC3WHO10)** *(ICCC3WHO)*	3	SEER	See Attachment A
1983	ICCC site recode extended ICD-O-3/ WHO 2008 **(ICC3XWHO1-ICC3XWHO10)** *(ICCC3XWHO)*	3	SEER	See Attachment A
1986	Behavior recode for analysis **(behtrend1-behtrend10)** *(behanal)*	1	SEER	See Attachment A
1987	Histology recode - broad groupings **(histrec1-histrec10)** *(histrec)*	2	SEER	See Attachment A
1989	Histology recode – Brain groupings **(hisrcb1-hisrcb10)** *(brainrec)*	2	SEER	See Attachment A
1991	CS Schema – v 0204 **(cs04sch1-cs04sch10)** *(cs0204schema)*	3	SEER	See Attachment A
1994	SEER Historic Stage A **(hstst1-hstst10)** *(hst_stga)*	1	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
1995	AJCC Stage 3rd Edition (1988-2003) **(ajccstg1-ajccstg10)** *(ajcc_stg)*	2	SEER	See Attachment A
1997	SEER Modified AJCC Stage 3rd edition (1988-2003) **(aj3sr1-aj3sr10)** *(aj_3seer)*	2	SEER	See Attachment A
1999	SEER Summary Stage 1977 (1995-2000) **(sss77v1-sss77v10)** *(ssg77)*	1	SEER	See Attachment A
2000	SEER summary stage 2000 (2001-2003) **(sssm2Z1-sssm2Z10)** *(ssg2000)*	1	SEER	See Attachment A
2001	First Malignant Primary Indicator **(frstprm1-frstprm10)** *(firstprm)*	1	SEER	See Attachment A
2002	FIPS State Code **(statecd1-statecd10)** *(stcounty)*	2	SEER	See Attachment A
2004	FIPS County Code **(cnty1-cnty10)** *(stcounty)*	3	SEER	See Attachment A
2007	IHS Code **(ihscd1-ihscd10)**	1	SEER	See Attachment A
2008	Summary Stage 2000 (1998+) **(summ2k1-summ2k10)** *(hist_ssg_2000)*	1	SEER	See Attachment A
2009	AYA site recode/WHO 2008 **(ayawho1-ayawho10)** *(aya_recode)*	2	SEER	See Attachment A
2011	Lymphoma subtype recode/WHO 2008 **(lymwho1-lymwho10)** *(lymphoma_recode)*	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
2013	SEER cause-specific death classification **(vsrtdx1-vsrtdx10)** *(dth_class)*	1	SEER	See Attachment A
2014	SEER other cause of death classification **(odthclass1-odthclass10)** *(o_dth_class)*	1	SEER	See Attachment A
2015	CS Tumor Size/Ext Eval (2004+) **(csts1-csts10)** *(exteval)*	1	SEER	See Attachment A
2016	CS Reg Nodes Eval (2004+) **(csrg1-csrg10)** *(nodeeval)*	1	SEER	See Attachment A
2017	CS Mets Eval (2004+) **(csmt1-csmt10)** *(metseval)*	1	SEER	See Attachment A
2018	Primary by International Rules **(intprim1-intprim10)** *(intprim)*	1	SEER	See Attachment A
2019	ER Status Recode Breast Cancer (1990+) **(erstat1-erstat10)** *(erstatus)*	1	SEER	See Attachment A
2020	PR Status Recode Breast Cancer (1990+) **(prstat1-prstat10)** *(prstatus)*	1	SEER	See Attachment A
2021	CS Schema **(cssch1-cssch10)** *(csschema)*	2	SEER	See Attachment A
2023	CS Site-Specific Factor 8 (2004+) **(cs8st1-cs8st10)** *(cs_ssf8)*	3	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
2026	CS Site-Specific Factor 10 (2004+) **(cs10st1-cs10st10)** *(cs_ssf10)*	3	SEER	See Attachment A
2029	CS Site-Specific Factor 11 (2004+) **(cs11st1-cs11st10)** *(cs_ssf11)*	3	SEER	See Attachment A
2032	CS Site-Specific Factor 13 (2004+) **(cs13st1-cs13st10)** *(cs_ssf13)*	3	SEER	See Attachment A
2035	CS Site-Specific Factor 15 (2004+) **(cs15st1-cs15st10)** *(cs_ssf15)*	3	SEER	See Attachment A
2038	CS Site-Specific Factor 16 (2004+) **(cs16st1-cs16st10)** *(cs_ssf16)*	3	SEER	See Attachment A
2041	Lymph-vascular Invasion (2004+) **(vasinv1-vasinv10)** *(vasinv)*	1	SEER	See Attachment A
2042	Survival Months **(srvm1-srvm10)** *(srv_time_mon)*	4	SEER	See Attachment A
2046	Survival Months Flag **(srvmflag1-srvmflag10)** *(srv_time_mon_flag)*	1	SEER	See Attachment A
2047	Filler	5		
2052	Insurance Recode (2007+) **(insrecpb1-insrecpb10)** *(insrec_pub)*	1	SEER	See Attachment A
2053	Derived AJCC T 7th ed **(dajcc7t1-dajcc7t10)** *(dajcc7t)*	3	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
2056	Derived AJCC N 7th ed **(dajcc7n1-dajcc7n10)** (*dajcc7n*)	3	SEER	See Attachment A
2059	Derived AJCC M 7th ed **(dajcc7m1-dajcc7m10)** (*dajcc7m*)	3	SEER	See Attachment A
2062	Derived AJCC 7 Stage Group **(dajcc7_01-dajcc7_10)** (*dajcc7stg*)	3	SEER	See Attachment A
2065	Breast Adjusted AJCC 6th T (1988+) **(adjajc6t1-adjajc6t10)** (*adjtm_6value*)	2	SEER	See Attachment A
2067	Breast Adjusted AJCC 6th N (1988+) **(adjajc6n1-adjajc6n10)** (*adjnm_6value*)	2	SEER	See Attachment A
2069	Breast Adjusted AJCC 6th M (1988+) **(adjajc6m1-adjajc6m10)** (*adjm_6value*)	2	SEER	See Attachment A
2071	Breast Adjusted AJCC 6th Stage (1988+) **(adjajc6_01-adjajc6_10)** (*adjajccstg*)	2	SEER	See Attachment A
2073	CS Site-Specific Factor 7 **(cs7st1-cs7st10)** (*cs7site*)	3	SEER	See Attachment A
2076	CS Site-Specific Factor 9 **(cs9st1-cs9st10)** (*cs9site*)	3	SEER	See Attachment A
2079	CS Site-Specific Factor 12 **(cs12st1-cs12st10)** (*cs12site*)	3	SEER	See Attachment A
2082	Derived HER2 Recode (2010+) **(her2rec1-her2rec10)** (*her2*)	1	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
2083	Breast Subtype (2010+) **(brstsub1-brstsub10)** (*brst_sub*)	1	SEER	See Attachment A
2084	Lymphoma – Ann Arbor Stage (1983+) **(annarbor1-annarbor10)** (*annarbor*)	1	SEER	See Attachment A
2085	CS Mets at DX-Bone (2010+) **(csmetsdxb_pub1-csmetsdxb_pub10)** (*csmetsdxb_pub*)	1	SEER	See Attachment A
2086	CS Mets at DX-Brain (2010+) **(csmetsdxbr_pub1-csmetsdxbr_pub10)** (*csmetsdxbr_pub*)	1	SEER	See Attachment A
2087	CS Mets at DX-Liver (2010+) **(csmetsdxliv_pub1-csmetsdxliv_pub10)** (*csmetsdxliv_pub*)	1	SEER	See Attachment A
2088	CS Mets at DX-Lung (2010+) **(csmetsdxlung_pub1-csmetsdxlung_pub10)** (*csmetsdxlung_pub*)	1	SEER	See Attachment A
2089	T value-based on AJCC 3rd (1988-2003) **(t_value1-t_value10)** (*t_value*)	2	SEER	See Attachment A
2091	N value-based on AJCC 3rd (1988-2003) **(n_value1-n_value10)** (*n_value*)	2	SEER	See Attachment A
2093	M value-based on AJCC 3rd (1988-2003) **(m_value1-m_value10)** (*m_value*)	2	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
2095	Diagnosis Date Flag **(ddxflag1-ddxflag10)**	2	SEER	12 = Date is unknown Missing = A valid date is provided, or the date was not expected to have been transmitted
2097	Therapy date flag **(dthflag1-dthflag10)**	2	SEER	10 = Unknown if therapy was administered 11 = Therapy was not administered 12 = Therapy was administered, but the date is unknown Missing = A valid date is provided, or was not expected to have been transmitted
2099	Month Therapy Started **(monrx1-monrx10)** *(monthrx)*	2	SEER	Month of Therapy 00 = No Therapy 01-12 = Valid month blank = Unknown
2101	Year Therapy Started **(yearrx1-yearrx10)** *(year_rx)*	4	SEER	Year of Therapy. If year is blank, refer to Therapy Date Flag (dthflag) for more information.
2105	Other Therapy **(other_tx1-other_tx10)** *(othr_rx)*	1	SEER	0 = No other cancer-directed therapy 1 = Other cancer-directed therapy 2 = Other experimental cancer-directed therapy 3 = Double blind study, code not yet broken 6 = Unproven therapy (including Laetrile, Krebiozen, etc) 7 = Refused therapy 1-3 above 8 = Recommended, unknown if administered 9 = Unknown
2106	ICD-O coding scheme **(icdo1-icdo10)** *(icdover)*	1	SEER	2 = Originally coded in ICD-O-2, (1973-2000) 3 = Originally coded in ICD-O-3, (2001+)

COL	FIELD	LENGTH	SOURCE	NOTES
2107	NAPIIA Derived API Race **(napiia1-napiia10)** *(napiia)*	2	SEER	The NAACCR Asian Pacific Islander Identification Algorithm version 1 (NAPIIA v1) uses a combination of NAACCR variables to classify cases directly or indirectly as Asian Pacific Islander for analytic purposes. This version of the algorithm is focused on coding cases with a race code of Asian NOS (race code 96) to a more specific Asian race category, using the birthplace and name fields (first, last, and maiden names). Birthplace can be used to indirectly assign a specific race to one of eight Asian race groups (Chinese, Japanese, Vietnamese, Korean, Asian Indian, Filipino, Thai, and Cambodian). Names can be used to indirectly assign a specific race to one of seven Asian groups (Chinese, Japanese, Vietnamese, Korean, Asian Indian, Filipino, and Hmong). Future versions of NAPIIA will incorporate Pacific Islanders and will potentially incorporate name lists for Thai, Cambodian, and Laotians. Codes are the same as for Race/Ethnicity. Blanks in field mean the algorithm was not run.
2109	Primary Payer at DX **(payer_dx1-payer_dx10)** *(payerdx)*	2	SEER	Primary Payer at Diagnosis identifies the patients primary insurance carrier or method of payment at the time of initial diagnosis and/or treatment. 01 = Not insured 02 = Not insured, self pay 10 = Insurance, NOS 20 = Private Insurance: Managed care, HMO, or PPO 21 = Private Insurance: Fee-for-Service 31 = Medicaid 35 = Medicaid - Administered through a Managed Care plan 60 = Medicare/Medicare, NOS 61 = Medicare with supplement, NOS 62 = Medicare - Administered through a Managed Care plan 63 = Medicare with private supplement 64 = Medicare with Medicaid eligibility

COL	FIELD	LENGTH	SOURCE	NOTES
				65 = TRICARE
				66 = Military
				67 = Veterans Affairs
				68 = Indian/Public Health Service
				99 = Insurance status unknown
4856	Filler	1		

Codes for the variables:
-Cause of Death to site recode KM (codkm)
-Cause of Death to site recode (codpub)

```
00000 = Alive
20010 = Lip
20020 = Tongue
20030 = Salivary Gland
20040 = Floor of Mouth
20050 = Gum and Other Mouth
20060 = Nasopharynx
20070 = Tonsil
20080 = Oropharynx
20090 = Hypopharynx
20100 = Other Oral Cavity and Pharynx
21010 = Esophagus
21020 = Stomach
21030 = Small Intestine
21040 = Colon excluding Rectum
21050 = Rectum and Rectosigmoid Junction
21060 = Anus, Anal Canal and Anorectum
21071 = Liver
21072 = Intrahepatic Bile Duct
21080 = Gallbladder
21090 = Other Biliary
21100 = Pancreas
21110 = Retroperitoneum
21120 = Peritoneum, Omentum and Mesentery
21130 = Other Digestive Organs
22010 = Nose, Nasal Cavity and Middle Ear
22020 = Larynx
22030 = Lung and Bronchus
22050 = Pleura
22060 = Trachea,Mediastinum and Other Respiratory Organs
23000 = Bones and Joints
24000 = Soft Tissue including Heart
25010 = Melanoma of the Skin
25020 = Other Non-Epithelial Skin
26000 = Breast
27010 = Cervix Uteri
27020 = Corpus Uteri
27030 = Uterus, NOS
27040 = Ovary
27050 = Vagina
27060 = Vulva
27070 = Other Female Genital Organs
```

28010 = Prostate
28020 = Testis
28030 = Penis
28040 = Other Male Genital Organs
29010 = Urinary Bladder
29020 = Kidney and Renal Pelvis
29030 = Ureter
29040 = Other Urinary Organs
30000 = Eye and Orbit
31010 = Brain and Other Nervous System
32010 = Thyroid
32020 = Other Endocrine including Thymus
33010 = Hodgkin Lymphoma
33040 = Non-Hodgkin Lymphoma
34000 = Myeloma
35011 = Acute Lymphocytic Leukemia
35012 = Chronic Lymphocytic Leukemia
35013 = Other Lymphocytic Leukemia
35021 = Acute Myeloid Leukemia
35031 = Acute Monocytic Leukemia
35022 = Chronic Myeloid Leukemia
35023 = Other Myeloid/Monocytic Leukemia
35041 = Other Acute Leukemia
35043 = Aleukemic, Subleukemic and NOS
36010 = Mesothelioma (ICD-10 only) (not in CODPUB)
36020 = Kaposi Sarcoma (ICD-10 only) (not in CODPUB)
37000 = Miscellaneous Malignant Cancer
38000 = In situ, benign or unknown behavior neoplasm
50000 = Tuberculosis
50010 = Syphilis
50030 = Septicemia
50040 = Other Infectious and Parasitic Diseases including HIV
50050 = Diabetes Mellitus
50051 = Alzheimers (ICD-9 and 10 only)
50060 = Diseases of Heart
50070 = Hypertension without Heart Disease
50080 = Cerebrovascular Diseases
50090 = Atherosclerosis
50100 = Aortic Aneurysm and Dissection
50110 = Other Diseases of Arteries, Arterioles, Capillaries
50120 = Pneumonia and Influenza
50130 = Chronic Obstructive Pulmonary Disease and Allied Cond
50140 = Stomach and Duodenal Ulcers
50150 = Chronic Liver Disease and Cirrhosis
50160 = Nephritis, Nephrotic Syndrome and Nephrosis
50170 = Complications of Pregnancy, Childbirth, Puerperium

50180 = Congenital Anomalies
50190 = Certain Conditions Originating in Perinatal Period
50200 = Symptoms, Signs and Ill-Defined Conditions
50210 = Accidents and Adverse Effects
50220 = Suicide and Self-Inflicted Injury
50230 = Homicide and Legal Intervention
50300 = Other Cause of Death
41000 = State DC not available or state DC available but no COD
99999 = Unknown/missing/invalid COD

Note:

Rural/Urban Continuum as Defined in the 2004 ARF file
(urban/rural code)

The **2003 Rural/Urban Continuum Codes** are from Economic Research Service (ERS), Department of Agriculture. The codes form a classification scheme that distinguishes metropolitan (metro) counties by the population size of their metro area and nonmetropolitan (nonmetro) counties by degree of urbanization and adjacency to a metro area or nonmetro areas. All U.S. counties and county equivalents are grouped according to the official metro status announced by the Office of Management and Budget (OMB) in June 2003, when the population and worker commuting criteria used to identify metro counties were applied to results of the 2000 Census.

Metro counties are distinguished by population size of the Metropolitan Statistical Area of which they are part. Nonmetro counties are classified according to the aggregate size of their urban population. Within the three urban size categories, nonmetro counties are further identified by whether or not they have some functional adjacency to a metro area or areas. A nonmetro county is defined as adjacent if it physically adjoins one or more metro areas, and has at least 2 percent of its employed labor force commuting to central metro counties. Nonmetro counties that do not meet these criteria are classed as nonadjacent.

In concept, the 2003 version of the Rural-Urban Continuum Codes is comparable with that of earlier decades. However, OMB made major changes in its metro area delineation procedures for the 2000 Census, and the Census Bureau changed the way in which rural and urban are measured. Therefore, the new Rural-Urban Continuum Codes are not fully comparable with those of earlier years. OMBs changes added some additional metro areas by no longer requiring that a metro area must have at least 100,000 population if its urbanized area has no place of at least 50,000 people. More importantly, simplifying the worker commuting criteria that determine outlying metro counties had the effect of both adding numerous new outlying counties to metro status while deleting a smaller number that were previously metro.

The Census Bureau made a radical shift in determining rural-urban boundaries by changing and liberalizing the procedures for delineating urbanized areas of 50,000 or more people, and abandoning place boundaries in measuring urban or rural population. The procedures used in defining Urbanized Areas were extended down to clusters of 2,500 or more people, based solely on population density per square mile.

In earlier versions of the Rural-Urban Continuum Codes, metro areas with 1 million population or more were subdivided between central counties (Code 0) and fringe counties (Code 1). The Code 1 group has become much less meaningful in the last two censuses as more and more counties of large metro areas have been rated as central counties by OMB procedures. In 2000, only 1.6 percent of the population of large metro areas was in fringe counties. Therefore, this distinction has been dropped. Codes 0 and 1 have been combined, and the new code 1 represents all counties in metro areas of 1 million or more population.
The 2003 Rural/Urban Continuum Codes are defined as follows:

CODE **METROPOLITAN COUNTIES (1-3)**
01   Counties of metro areas of 1 million population or more
02   Counties in metro areas of 250,000 - 1,000,000 population
03   Counties in metro areas of fewer than 250,000 population

CODE **NON METROPOLITAN COUNTIES (4-9)**
04   Urban population of 20,000 or more, adjacent to a metro area
05   Urban population of 20,000 or more, not adjacent to a metro area
06   Urban population of 2,500-19,999, adjacent to a metro area
07   Urban population of 2,500-19,999, not adjacent to a metro area
08   Completely rural or less than 2,500 urban population, adjacent to a metro area
09   Completely rural or less than 2,500 urban population, not adjacent to a metro area
99   Missing Value