December 18, 2019

FILENAME 1: pedsf.5pallcan.txt

LRECL 1: 6,240

SEER Cases diagnosed 1991-2015

PEDSF File created on May 22, 2019

SEER data extracted from November, 2018 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 1991-2017.

PROGRAMMER: Angela Meekins, Information Management Services, Inc.

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

NOTE1: With Linkage 2014 PEDSF no longer includes SEER cases from Los Angeles and San Jose from 1988 thru 1991. LA and SJ now include data from 1992 on.

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.

NOTE3: If you are using old data received prior to Linkage 2010 and link-up the data for **Louisiana** patients using the "oshufcase" it's important to note that some old IDs are repeated as current IDs for different patients. It is suggested that after patients are matched on "oshufcase" they should use the current shufcase so duplicate IDs won't lead to different patients.

NOTE4: The SEER program removed Cervix InSitu patients diagnosed between 1973-1995 from the November 2017 SEER Submission. These patients are no longer included in the PEDSF as well.

NOTE5: With the 2018 Linkage the PEDSF layout has many changes including census tracts, state, county, rural code, tract certainty variables and poverty indicator are now in the SEER repeating section. Entitlement variables have moved and are written out together instead of yearly, this includes Part D data.

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | NOTES |
|-----|----------------------------------|---------------|-----------------|--|
| 1 | Patient ID (patient_id) | 10 | Master File | |
| 1 | SEER Cases - Registry | 2 | SEER | First two digits in the patient ID. 02 = Connecticut (1973+) 20 = Detroit (1973+) 21 = Hawaii (1973+) 22 = Iowa (1973+) 23 = New Mexico (1973+) 25 = Seattle (1974+) 26 = Utah (1973+) 42 = Kentucky (2000+) 43 = Louisiana (2000+) 44 = New Jersey (2000+) 87 = Georgia 88 = California NOTE: Specific California and Georgia registries will be differentiated in the repeated registry codes (starting in column 2390, variables reg1-reg10) |
| 3 | Case Number | 8 | Assigned by IMS | Encrypted SEER Case Number |
| 11 | Five Percent Indicator (fivepct) | 1 | Master File | Always: Y = Included in this file |
| 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) |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|-----|---|---------------|----------------|--|
| 17 | Match Flag 5 (Istflg5) | 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 | Linkage Flag (linkflag) | 1 | Created at IMS | Indicates the linkage the patient is first associated with. 1 = 2005 linkage, earliest claim year is 1991 2 = 2008 linkage, earliest claim year is 1998 3 = 2010 linkage, earliest claim year is 2000 4 = 2012 linkage, earliest claim year is 2002 5 = 2014 linkage, earliest claim year is 2004 6 = 2016 linkage, earliest claim year is 2006 7 = 2018 linkage, earliest claim year is 2008 |
| 23 | Medicare Month of Birth (birthm) | 2 | EDB | MM, zeroes for invalid |
| 27 | Medicare Year of Birth (birthyr) | 4 | EDB | YYYY, zeroes for invalid |
| 31 | 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 12/31/17 |
| 39 | Valid Date of Death (vrfydth) | 1 | EDB | N = No Y = Yes |

| Decem | ber | 18. | 2019 |
|-------|----------------------|-----|------|
| | \sim \sim \sim | | |

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | NOTES |
|-----|------------------------------|---------------|----------------|---|
| 40 | 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 |
| 41 | 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 12/16 were treated as not dead for comparison purposes. 0 = Not dead by 12/16. 1 = Dead, both files agree. 2 = Dead, off by 1-3 months. 3 = Dead, off by 4-6 months. 4 = Dead in MEDICARE only, 5 = Dead but number of months could not be calculated due to the fact that month was missing for either SEER or MEDICARE. |
| 42 | Medicare Sex (m_sex) | 1 | EDB | 1 = Male 2 = Female |
| 43 | Medicare Race (race) | 1 | EDB | 1 = White 2 = Black 3 = Other 4 = Asian 5 = Hispanic 6 = N. Am. Native 0 = Unknown |

<u>Demographic Information</u>
The SEER information was extracted from the last available record for the patient.

SEER Sex (s_sex) 1 SEER 1 = Male44

| COL | <u>FIELD</u> | LENGTH | scember 18, SOURCE | |
|-----|-----------------------------|--------|-----------------------|--|
| | | | | 2 = Female |
| 47 | Race Recode Y (rac_recy) | 1 | SEER | 1 = White 2 = Black 3 = American Indian/Alaska Native 4 = Asian or Pacific Islander 9 = Unknown |
| 48 | Race Recode A (rac_reca) | 1 | SEER | 1 = White 2 = Black 3 = Other(American Indian/AK Native, Asian/Pacific Islander) 9 = Unknown |
| 49 | SEER Race/Ethnicity (srace) | 2 | SEER | 01 = White 02 = Black 03 = American Indian/Alaska Native 04 = Chinese 05 = Japanese 06 = Filipino 07 = Hawaiian 08 = Korean (1988+) 10 = Vietnamese (1988+) 11 = Laotian (1988+) 12 = Hmong (1988+) 13 = Kampuchean (1988+) 14 = Thai (1994+) 15 = Asian Indian or Pakistani, NOS (1988+) 16 = Asian Indian (2010+) 17 = Pakistani (2010+) 20 = Micronesian, NOS (1991+) 21 = Chamorran (1991+) 22 = Guamanian, NOS (1991+) 25 = Polynesian, NOS (1991+) 26 = Tahitian (1991+) 27 = Samoan (1991+) 28 = Tongan (1991+) 30 = Melanesian, NOS (1991+) 31 = Fiji Islander (1991+) 32 = New Guinean (1991+) 96 = Other Asian (1991+) 97 = Pacific Islander, NOS (1991+) 98 = Other 88 = No further race documented |

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | NOTES |
|-----|--|---------------|--------|--|
| | | | | 99 = Unknown |
| 53 | Origin recode NHIA (Hispanic, Non-Hisp) (origrecb) | 1 | SEER | 0 = Non-Spanish-Hispanic-Latino 1 = Spanish-Hispanic-Latino |
| 54 | NHIA Derived Hispanic Origin (nhiade) | 1 | SEER | 0 = Non-Spanish-Hispanic-Latino 1 = Mexican 2 = Puerto Rican 3 = Cuban 4 = South or Central American excluding Brazil 5 = Other specified Spanish/Hispanic origin including Europe 6 = Spanish/Hispanic/Latino, NOS 7 = NHIA surname match only 8 = Dominican Republic |
| 55 | Vital Status Recode (Study cutoff used) (stat_rec) | 1 | SEER | 1 = Alive 4 = Dead |
| 58 | Year of Birth (yr_brth) | 4 | SEER | 1800-2017 |
| 62 | Date of Birth Flag (dbrflag) | 2 | SEER | 12 = A proper value is applicable but not known |
| 64 | SEER Month of Death (ser_dodm) | 2 | SEER | 00 = Alive 01-12 = Valid Month 99 or Missing = Unknown Date complete through 12/31/2016 |
| 66 | SEER Year of Death (ser_dody) | 4 | SEER | 0000 = Alive 1991-2016 = Valid Year 9999 of Missing = Unknown Date complete through 12/31/2016 |
| 70 | Date of Death Flag (deathflag) | 2 | SEER | 0 = Alive 12 = A proper value is applicable but not known |
| 72 | 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 |
| 73 | Cause of Death ICD-8 or 9 (cod89v) | 4 | SEER | 0000 = Alive at last contact 7777 = State Death Certificate not |

| COL | FIELD | <u>LENGTH</u> | SOURCE | |
|-----|--|---------------|-------------------|--|
| | | | | available. 7797 = Death Certificate available, no COD listed. |
| 77 | 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. |
| 81 | 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. |
| 86 | 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. |
| 91 | Number of SEER records (count) | 2 | Created at IMS | Number of eligible SEER records. SEER records prior to 1992 for Rural Georgia, LA, San Jose and records prior to 2000 for Greater California, New Jersey, Kentucky, and Louisiana were not included in PEDSF. |
| 93 | Old Encrypted Casenum (oshufcase) | 8 | IMS | The encrypted case number for Louisiana and Seattle in the 2008 linkage. |
| 101 | Filler | 9 | | |
| 110 | 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 |
| 111 | End Stage Renal | 1 | EDB | N = No ESRD |

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | <u>NOTES</u> |
|-----|---|---------------|--------|---|
| | Disease Ever (esrd_ever) | | | Y = ESRD If an ESRD start date is found on the EDB file then this variable is set to Y. |
| 112 | End Stage Renal Disease start date (ersstm1,ersstd1, erssty1) | 5*8 | EDB | MMDDYYYY Start dates for Renal Disease. |
| | (ersstm5,ersstd5, erssty5) | | | |
| 152 | End Stage Renal Disease end date (erstrm1, erstrd1, erstry1) | 5*8 | EDB | MMDDYYYY End dates for Renal Disease. |
| | (erstrm5,erstrd5, erstry5) | | | |
| 192 | Medicare Status Code (medcd1991- medcd2017) | 27*2 | EDB | Blank = Not Enrolled 10 = Aged 11 = Aged with ESRD 20 = Disabled 21 = Disabled with ESRD 31 = ESRD only 99 = Enrolled but Medicare Status is Unknown Variable is selected based on dates. The last status code for the year is kept. |

Geographic Information

EDB program data are from the EDB record corresponding to the last location for that entitlement year. If patient was not entitled in that year then the variable will be missing.

| 246 | Zip Code (zip5_1991- zip5_2017) | 27*5 | EDB | Encrypted Zip code, 5 digits. * Special permission required for unencrypted zip code. |
|-----|------------------------------------|------|-----|--|
| 381 | Zip Code (zip4_1991- zip4_2017) | 27*4 | EDB | Last four digits of zip code. Blanked out when Encrypted Zip code is given. * Special permission required for unencrypted zip code. |
| 489 | State (state1991- state2017) | 27*2 | EDB | FIPS State Code from EDB |
| 543 | County (cnty1991- | 27*3 | EDB | FIPS County Code from EDB |

| <u>COL</u> | <u>FIELD</u> | LENGTH | SOURCE | |
|------------|---|--------|--------|---|
| | cnty2017) | | | |
| 624 | Rural-Urban code (urban1991- urban2017) | 27*2 | SEER | 1993, 2003 and 2013 Rural-Urban continuum codes were extracted from SEER*stat. 1993 continuum codes were used for years 1991-1998, 2003 continuum codes were used for years 1999-2008 and 2013 continuum codes were used from 2009-2017. 00-09, 99 see documentation for Rural-Urban continuum codes starting on page 31 of this document. |
| | Rural-Urban recode (urbrur1991- urbrur2017) | 27*1 | SEER | Rural-Urban Code (no longer on the file but is created in the input statement using the urban codes) 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) |
| 678 | Registry Code (registry1991- registry2017) | 27*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. |

| Decem | ber 1 | 8. | 2019 |
|-------|-------|----|------|
| | | | |

| <u>COL</u> | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | <u>NOTES</u> |
|------------|---|---------------|--------|---|
| 732 | Registry Code #2 (reg2cd1991- reg2cd2017) | 27*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.

EDB Entitlement Information

EDB program data is used to create monthly indicators for Part A, B and HMO entitlement based on Start and End dates provided in the EDB.

| 790 | Entitlement Indicators | 324*1 | EDB | One indicator for each month in the year. |
|-----|-------------------------------|-------|-----|---|
| | (mon1-mon324) | | | 0 = Not entitled |
| | | | | 1 = Part A only |
| | | | | 2 = Part B only |
| | | | | 3 = Part A and B |
| | | | | Note: $mon1 = 1/1991$, $mon2 = 2/1991$, |
| | | | | mon3 = 3/1991mon324 = 12/2017 |

| <u>COL</u> | FIELD | LENGTH | scember 18, <u>SOURCE</u> | NOTES |
|------------|---|--------|------------------------------|--|
| 1114 | HMO Indicators (gho1-gho324) | 324*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 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: gho1 = 1/1991, gho2 = 2/1991, gho3 = 3/1991 gho324 = 12/2017 |
| 1438 | State Buy-in Monthly Indicators (allflag1-allflag324) | 324*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: allflag1 = 1/1991, allflag2 = 2/1991, allflag3 = 3/1991allflag324 = 12/2017 |
| 1762 | Part A Months (ptacnt1991-ptacnt2017) | 27*2 | EDB | Number of months covered for Part A in each year from 1991-2017. 00 = No coverage that year 01-12 = Number of months of coverage |
| 1816 | Part B Months (ptbcnt1991-ptbcnt2017) | 27*2 | EDB | Number of months covered for Part B in each year from 1991-2017. 00 = No coverage that year 01-12 = Number of months of coverage |
| 1870 | HMO Months (hmocnt1991- hmocnt2017) | 27*2 | EDB | Number of months as an HMO Member in each year from 1991-2017. 00 = No coverage that year 01-12 = Number of months of coverage |
| 1924 | State Buy-in Months (stbuyin1991-stbuyin2017) | 27*2 | EDB | Number of months with state buy-in coverage In each year from 1991-2017. 00 = No coverage that year 01-12 = Number of months of coverage |

December 18, 2019

COL FIELD LENGTH SOURCE NOTES

Repeated Part D Denominator Information

Entitlement information from the Part D Denominator file are provided from 2007 to 2016.

| Blank | values indicate the patient | had no Part | D record fo | or that year. |
|-------|--|-------------|-----------------|--|
| 1978 | Plan-Value Indicators (Jan. – Dec.) (plan07_01-plan07_12) (plan08_01-plan08_12) (plan15_01-plan15_12) (plan16_01-plan16_12) Or read them in all together like other entitlement variables. (planmon193-planmon312) | 120*1 | Part D Denom | 12 monthly 1-byte indicators for each year from 2007-2016. Indicates for each month of the Denominator reference year, the Part D enrollment, which is based on the 1 st digit or the Part D contract number. Recodes only indicating type of plan; no 5-digit plan numbers. Blank = No Part D record 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) |
| 2098 | State Reported Dual Eligible Status Code (Jan. – Dec.) (dual07_01-dual07_12) (dual08_01-dual08_12) (dual15_01-dual15_12) (dual16_01-dual16_12) Or read them in all together like other entitlement variables. (dualmon193-dualmon312) | 120*2 | Part D Denom | 12 monthly 2-byte indicators Indicates for each month of the Denominator reference year (2007-2016), the dual eligibility status, if any, for the beneficiary. Blank = No Part D record 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

| December | 18, | 2019 |
|----------|-----|------|
|----------|-----|------|

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | <u>NOTES</u> |
|------|---|---------------|-----------------|--|
| | | | | 08 = Other Dual Eligibles (Non-QMB, SLMB, QWDI, or QI) w/Medicaid coverage including RX 09 = Other Dual Eligibles but without Medicaid coverage 99 = Unknown |
| 2338 | Plan Coverage Months (ptd07-ptd16) | 10*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, E, or X. |
| 2358 | Medicaid Dual Eligible Months (dualcnt07-dualcnt16) | 10*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 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. Please note that diagnoses are sorted on SEQUENCE NUMBER and not date of diagnosis. Meaning all malignant cancers (sequence # 00-59) are first then nonmalignant cancers (sequence 60+)

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | |
|------|---|---------------|--------|--|
| 2390 | 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+) |
| 2392 | Marital Status at Diagnosis (marst1-marst10) (mar_stat) | 1 | SEER | See Attachment A |
| 2393 | Filler | 3 | | |
| 2396 | Age at Diagnosis (agedx1-agedx10) (age_dx) | 3 | SEER | See Attachment A |
| 2399 | Sequence Number (seq1-seq10) (seq_num) | 2 | SEER | See Attachment A |
| 2401 | Month of Diagnosis (modx1-modx10) (date_mo) | 2 | SEER | See Attachment A |
| 2403 | Year of Diagnosis (yrdx1-yrdx10) (date_yr) | 4 | SEER | See Attachment A |
| 2407 | Primary Site (site1-site10) (site02v) | 4 | SEER | See Attachment A |

| Decemi | er ' | 18 | 2019 |
|--------|---------|----|------|
| | | | |

| <u>COL</u> | <u>FIELD</u> | LENGTH | SOURCE | NOTES |
|------------|--|--------|--------|------------------|
| 2411 | Laterality (lat1-lat10) (lateral) | 1 | SEER | See Attachment A |
| 2412 | Histology ICD-O-2 (1992-2000) (hist2_1-hist2_10) (hist02v) | 4 | SEER | See Attachment A |
| 2416 | Behavior code ICD-O-2 (1992-2000) (beh2_1-beh2_10) (beh02v) | 1 | SEER | See Attachment A |
| 2417 | Histology ICD-O-3 (hist1-hist10) (hist03v) | 4 | SEER | See Attachment A |
| 2421 | Behavior code ICD-O-3 (beh1-beh10) (beh03v) | 1 | SEER | See Attachment A |
| 2422 | Grade (grade1-grade10) (grade) | 1 | SEER | See Attachment A |
| 2423 | Diagnostic Confirmation (dxconf1-dxconf10) (dx_conf) | 1 | SEER | See Attachment A |
| 2424 | Type of Reporting Source (src1-src10) (rept_src) | 1 | SEER | See Attachment A |
| 2425 | EOD 10 - Tumor Size (1988-2003) (e10sz1-e10sz10) (eod10_sz) | 3 | SEER | See Attachment A |
| 2428 | EOD 10 – Tumor Extent (1988-2003) (e10ex1-e10ex10) (eod10_ex) | 2 | SEER | See Attachment A |

| December 18, 2019 |
|-------------------|
|-------------------|

| <u>COL</u> | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------------|---|---------------|--------|------------------|
| 2430 | EOD 10 – Prostate Pathology Extension (1995-2003) (e10pe1-e10pe10) (eod10_pe) | 2 | SEER | See Attachment A |
| 2432 | EOD 10 – Lymph Nodes (1988-2003) (e10nd1-e10nd10) (eod10_nd) | 1 | SEER | See Attachment A |
| 2433 | EOD 10 - # Regional Positive Nodes (1988+) (e10pn1-e10pn10) (eod10_pn) | 2 | SEER | See Attachment A |
| 2435 | EOD 10 - # Regional Nodes Examined (1988+) (e10ne1-e10ne10) (eod10_ne) | 2 | SEER | See Attachment A |
| 2437 | EOD – OLD 13 Digit (1973-1982) (eod13_1-eod13_10) (eod13) | 13 | SEER | See Attachment A |
| 2450 | EOD – OLD 2 Digit (1973-1982) (eod2_1-eod2_10) (eod2) | 2 | SEER | See Attachment A |
| 2452 | EOD 4 -size (1983-1987) (e4siz1-e4siz10) (eod4) | 2 | SEER | See Attachment A |
| 2454 | EOD 4 – extent (1983-1987) (e4ext1-e4ext10) (eod4) | 1 | SEER | See Attachment A |
| 2455 | EOD 4 – nodes (1983-1987) (e4nod1-e4nod10) (eod4) | 1 | SEER | See Attachment A |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------|--|---------------|--------|---|
| 2456 | Coding system – extent of disease (1973-2003) (eod_cd1-eod_cd10) (eodcode) | 1 | SEER | See Attachment A |
| 2457 | Tumor Marker 1 (1990-2003) (tumor1_1-tumor1_10) (tumor_1v) | 1 | SEER | See Attachment A |
| 2458 | Tumor Marker 2 (1990-2003) (tumor2_1-tumor2_10) (tumor_2v) | 1 | SEER | (Prostate data has been removed from this field) See Attachment A |
| 2459 | Tumor Marker 3 (1998-2003) (tumor3_1-tumor3_10) (tumor_3v) | 1 | SEER | See Attachment A |
| 2460 | CS Tumor Size (2004+) (cstum1-cstum10) (cs_size) | 3 | SEER | See Attachment A |
| 2463 | CS Extension (2004+) (csex1-csex10) (cs_ext) | 3 | SEER | See Attachment A |
| 2466 | CS Lymph Nodes (2004+) (cslym1-cslym10) (cs_node) | 3 | SEER | See Attachment A |
| 2469 | CS Mets at Dx (2004+) (csmet1-csmet10) (cs_mets) | 2 | SEER | See Attachment A |
| 2471 | CS Site-Specific Factor 1 (2004+) (cs1st1-cs1st10) (cs_ssf1) | 3 | SEER | See Attachment A |
| 2474 | CS Site-Specific Factor 2 (2004+) (cs2st1-cs2st10) (cs_ssf2) | 3 | SEER | (Prostate data has been removed from this field) See Attachment A |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------|---|---------------|--------|------------------|
| 2477 | CS Site-Specific Factor 3 (2004+) (cs3st1-cs3st10) (cs_ssf3) | 3 | SEER | See Attachment A |
| 2480 | CS Site-Specific Factor 4 (2004+) (cs4st1-cs4st10) (cs_ssf4) | 3 | SEER | See Attachment A |
| 2483 | CS Site-Specific Factor 5 (2004+) (cs5st1-cs5st10) (cs_ssf5) | 3 | SEER | See Attachment A |
| 2486 | CS Site-Specific Factor 6 (2004+) (cs6st1-cs6st10) (cs_ssf6) | 3 | SEER | See Attachment A |
| 2489 | CS Site-Specific Factor 25 (2004+) (cs25st1-cs25st10) (cs_ssf25) | 3 | SEER | See Attachment A |
| 2492 | Derived AJCC T, 6 th edition (2004+) (dajcct1-dajcct10) (d_ajcc_t) | 2 | SEER | See Attachment A |
| 2494 | Derived AJCC N, 6 th edition (2004+) (dajccn1-dajccn10) (d_ajcc_n) | 2 | SEER | See Attachment A |
| 2496 | Derived AJCC M, 6 th edition (2004+) (dajccm1-dajccm10) (d_ajcc_m) | 2 | SEER | See Attachment A |
| 2498 | Derived AJCC Stage Group, 6 th edition (2004+) (dajccstg1- dajccstg10) | 2 | SEER | See Attachment A |
| 2500 | (d_ajcc_s) Derived SS1977 (2004+) (dss77s1-dss77s10) (d_ssg77) | 1 | SEER | See Attachment A |

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | NOTES |
|------|---|---------------|--------|------------------|
| 2501 | Derived SS2000 (2004+) (dss00s1-dss00s10) (d_ssg00) | 1 | SEER | See Attachment A |
| 2502 | Derived AJCC-Flag (dajcflg1-dajcflg1) (d_ajcc_f) | 1 | SEER | See Attachment A |
| 2503 | CS Version Original (2004+) (csvf1-csvf10) (csv_org) | 6 | SEER | See Attachment A |
| 2509 | CS Version Latest (2004+) (csvl1-csvl10) (csv_der) | 6 | SEER | See Attachment A |
| 2515 | CS Version Input Current (2004+) (cscurrent1- cscurrent10) (csv_cur) | 6 | SEER | See Attachment A |
| 2521 | RX Summ – Surgery of Primary Site (1998+) (sxprif1-sxprif10) (surgprim) | 2 | SEER | See Attachment A |
| 2523 | RX –Summ Scope Reg LN Sur (2003+) (sxscof1-sxscof10) (scope) | 1 | SEER | See Attachment A |
| 2524 | RX SummSurg Oth Reg/Dis (2003+) (sxsitf1-sxsitf10) (surgoth) | 1 | SEER | See Attachment A |
| 2525 | Number of Regional Lymph Nodes Examined (1998-2002) (numnd1-numnd10) (surgnode) | 2 | SEER | See Attachment A |
| 2527 | Reason no cancer- directed surgery (nosrg1-nosrg10) (no_surg) | 1 | SEER | See Attachment A |

| <u>COL</u> | FIELD | LENGTH | ecember 18, SOURCE | |
|------------|---|--------|-----------------------|---|
| 2528 | Site Specific Surgery (1983 – 1997) (sssurg1-sssurg10) (ss_surg) | 2 | SEER | See Attachment A |
| 2530 | Scope of regional lymph node surgery (1998-2002) (sxscop1-sxscop10) (scope02) | 1 | SEER | See Attachment A |
| 2531 | Surgery of other reg/dist sites (1998-2002) (sxsite1-sxsite10) (srgoth02) | 1 | SEER | See Attachment A |
| 2532 | Filler | 2 | | |
| 2534 | Type of follow-up expected (typefu1-typefu10) (typefup) | 1 | SEER | See Attachment A |
| 2535 | Age at Diagnosis recode (ager1-ager10) (age_rec) | 2 | SEER | See Attachment A |
| 2537 | 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. |
| 2542 | Recode ICD-O-2 to 9 (icdot09_1- icdot09_10) (icdot09v) | 4 | SEER | See Attachment A |
| 2546 | Recode ICD-O-2 to 10 (icdot10_1- icdot10_10) (icdot10v) | 4 | SEER | See Attachment A |
| 2550 | ICCC site recode ICD-O-3/WHO 2008 (ICCC3WHO1- ICCC3WHO10) (ICCC3WHO) | 3 | SEER | See Attachment A |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------|--|---------------|--------|------------------|
| 2553 | ICCC site recode extended ICD-O-3/ WHO 2008 (ICC3XWHO1- ICC3XWHO10) (ICCC3XWHO) | 3 | SEER | See Attachment A |
| 2556 | Behavior Recode for Analysis (behtrend1- behtrend10) (behanal) | 1 | SEER | See Attachment A |
| 2557 | Histology Recode - Broad Groupings (histrec1-histrec10) (histrec) | 2 | SEER | See Attachment A |
| 2559 | Histology Recode – Brain Groupings (hisrcb1-hisrcb10) (brainrec) | 2 | SEER | See Attachment A |
| 2561 | CS Schema – v 0204 (cs04sch1-cs04sch10) (cs0204schema) | 3 | SEER | See Attachment A |
| 2564 | Filler | 3 | | |
| 2567 | SEER Historic Stage A (hstst1-hstst10) (hst_stga) | 1 | SEER | See Attachment A |
| 2568 | AJCC Stage 3 rd Edition (1988-2003) (ajccstg1-ajccstg10) (ajcc_stg) | 2 | SEER | See Attachment A |
| 2570 | SEER Modified AJCC Stage 3 rd edition (1988-2003) (aj3sr1-aj3sr10) (aj_3seer) | 2 | SEER | See Attachment A |
| 2572 | SEER Summary Stage 1977 (1995-2000) (sss77v1-sss77v10) (ssg77) | 1 | SEER | See Attachment A |
| 2573 | SEER Summary Stage 2000 (2001-2003) (sssm2Z1-sssm2Z10) (ssg2000) | 1 | SEER | See Attachment A |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------|---|---------------|--------|------------------|
| 2574 | First Malignant Primary Indicator (frstprm1-frstprm10) (firstprm) | 1 | SEER | See Attachment A |
| 2575 | FIPS State Code (statecd1-statecd10) (stcounty) | 2 | SEER | See Attachment A |
| 2577 | FIPS County Code (cnty1-cnty10) (stcounty) | 3 | SEER | See Attachment A |
| 2580 | IHS Link (ihscd1-ihscd10) | 1 | SEER | See Attachment A |
| 2581 | Summary Stage 2000 (1998+) (summ2k1- summ2k10) (hist_ssg_2000) | 1 | SEER | See Attachment A |
| 2582 | AYA site recode/WHO 2008 (ayawho1-ayawho10) (aya_recode) | 2 | SEER | See Attachment A |
| 2584 | Lymphoma subtype recode/WHO 2008 (lymwho1-lymwho10) (lymphoma_recode) | 2 | SEER | See Attachment A |
| 2586 | SEER cause-specific death classification (vsrtdx1-vsrtdx10) (dth_class) | 1 | SEER | See Attachment A |
| 2587 | SEER other cause of death classification (odthclass1-odthclass10) (o_dth_class) | 1 | SEER | See Attachment A |
| 2588 | CS Tumor Size/Ext Eval (2004+) (csts1-csts10) (cstseval) | 1 | SEER | See Attachment A |
| 2589 | CS Reg Lymph Nodes Eval (2004+) (csrg1-csrg10) (csrgeval) | 1 | SEER | See Attachment A |

| <u>COL</u> | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------------|---|---------------|--------|------------------|
| 2590 | CS Mets Eval (2004+) (csmt1-csmt10) (csmteval) | 1 | SEER | See Attachment A |
| 2591 | Primary by International Rules (intprim1-intprim10) (intprim) | 1 | SEER | See Attachment A |
| 2592 | ER Status Recode Breast Cancer (1990+) (erstat1-erstat10) (erstatus) | 1 | SEER | See Attachment A |
| 2593 | PR Status Recode Breast Cancer (1990+) (prstat1-prstat10) (prstatus) | 1 | SEER | See Attachment A |
| 2594 | CS Schema AJCC 6 th ed (csschema) | 2 | SEER | See Attachment A |
| 2596 | CS Site-Specific Factor 8 (2004+) (cs8st1-cs8st10) (cs_ssf8) | 3 | SEER | See Attachment A |
| 2599 | CS Site-Specific Factor 10 (2004+) (cs10st1-cs10st10) (cs_ssf10) | 3 | SEER | See Attachment A |
| 2602 | CS Site-Specific Factor 11 (2004+) (cs11st1-cs11st10) (cs_ssf11) | 3 | SEER | See Attachment A |
| 2605 | CS Site-Specific Factor 13 (2004+) (cs13st1-cs13st10) (cs_ssf13) | 3 | SEER | See Attachment A |
| 2608 | CS Site-Specific Factor 15 (2004+) (cs15st1-cs15st10) (cs_ssf15) | 3 | SEER | See Attachment A |
| 2611 | CS Site-Specific Factor 16 (2004+) (cs16st1-cs16st10) (cs_ssf16) | 3 | SEER | See Attachment A |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------|---|---------------|--------|------------------|
| 2614 | Lymph-vascular Invasion (2004+) (vasinv1-vasinv10) (vasinv) | 1 | SEER | See Attachment A |
| 2615 | Survival Months (srvm1-srvm10) (srv_time_mon) | 4 | SEER | See Attachment A |
| 2619 | Survival Months Flag (srvmflag1- srvmflag10) (srv_time_mon_flag) | 1 | SEER | See Attachment A |
| 2620 | Filler | 1 | | |
| 2621 | Insurance Recode (2007+) (insrecpb1- insrecpb10) (insrec_pub) | 1 | SEER | See Attachment A |
| 2622 | Derived AJCC T 7 th ed (dajcc7t1-dajcc7t10) (dajcc7t) | 3 | SEER | See Attachment A |
| 2625 | Derived AJCC N 7 th ed (dajcc7n1-dajcc7n10) (dajcc7n) | 3 | SEER | See Attachment A |
| 2628 | Derived AJCC M 7 th ed (dajcc7m1-dajcc7m10) (dajcc7m) | 3 | SEER | See Attachment A |
| 2631 | Derived AJCC 7 Stage Group (dajcc7_01-dajcc7_10) (dajcc7stg) | 3 | SEER | See Attachment A |
| 2634 | Breast Adjusted AJCC 6 th T (1988+) (adjajc6t1-adjajc6t10) (adjtm_6value) | 2 | SEER | See Attachment A |
| 2636 | Breast Adjusted AJCC 6 th N (1988+) (adjajc6n1-adjajc6n10) (adjnm_6value) | 2 | SEER | See Attachment A |

| COL | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------|---|---------------|--------|------------------|
| 2638 | Breast Adjusted AJCC 6 th M (1988+) (adjajc6m1-adjajc6m10) (adjm_6value) | 2 | SEER | See Attachment A |
| 2640 | Breast Adjusted AJCC 6 th Stage (1988+) (adjajc6_01-adjajc6_10) (adjajccstg) | 2 | SEER | See Attachment A |
| 2642 | CS Site-Specific Factor 7 (cs7st1-cs7st10) (cs7site) | 3 | SEER | See Attachment A |
| 2645 | CS Site-Specific Factor 9 (cs9st1-cs9st10) (cs9site) | 3 | SEER | See Attachment A |
| 2648 | CS Site-Specific Factor 12 (cs12st1-cs12st10) (cs12site) | 3 | SEER | See Attachment A |
| 2651 | Derived HER2 Recode (2010+) (her2rec1-her2rec10) (her2) | 1 | SEER | See Attachment A |
| 2652 | Breast Subtype (2010+) (brstsub1-brstsub10) (brst_sub) | 1 | SEER | See Attachment A |
| 2653 | Lymphoma – Ann Arbor Stage (1983+) (annarbor1- annarbor10) (annarbor) | 1 | SEER | See Attachment A |
| 2654 | CS Mets at DX-Bone (2010+) (csmetsdxb_pub1-csmetsdxb_pub10) (csmetsdxb_pub) | 1 | SEER | See Attachment A |
| 2655 | CS Mets at DX-Brain (2010+) (csmetsdxbr_pub1-csmetsdxbr_pub10) (csmetsdxbr_pub) | 1 | SEER | See Attachment A |

| <u>COL</u> | FIELD | <u>LENGTH</u> | SOURCE | <u>NOTES</u> |
|------------|--|---------------|--------|---|
| 2656 | CS Mets at DX-Liver (2010+) (csmetsdxliv_pub1-csmetsdxliv_pub10) (csmetsdxliv_pub) | 1 | SEER | See Attachment A |
| 2657 | CS Mets at DX-Lung (2010+) (csmetsdxlung_pub1- csmetsdxlung_pub10) (csmetsdxlung_pub) | 1 | SEER | See Attachment A |
| 2658 | T value-based on AJCC 3 rd (1988-2003) (t_value1-t_value10) (t_value) | 2 | SEER | See Attachment A |
| 2660 | N value-based on AJCC 3 rd (1988-2003) (n_value1-n_value10) (<i>n_value</i>) | 2 | SEER | See Attachment A |
| 2662 | M value-based on AJCC 3 rd (1988-2003) (m_value1-m_value10) (m_value) | 2 | SEER | See Attachment A |
| 2664 | Total number of in situ/malignant tumors (maligcnt1-maligcnt10) (maligcount) | 2 | SEER | See Attachment A |
| 2666 | Total number of benign/borderline tumors (benbdcnt1-benbdcnt10) (benbordcount) | 2 | SEER | See Attachment A |
| 2668 | Census Tract (centrt1990_1- centrt1990_10) | 6 | SEER | 1970/80/90 Encrypted census tract * Special permission required for unencrypted census tract. |
| 2674 | Census Tract 2000 (centrt2000_1- centrt2000_10) | 6 | SEER | 2000 encrypted census tract * Special permission required for unencrypted census tract. |
| 2680 | Census Tract 2010 (centrt2010_1- centrt2010_10) | 6 | SEER | 2010 encrypted census tract * Special permission required for unencrypted census tract. |

| <u>COL</u> | FIELD | <u>LENGTH</u> | SOURCE | NOTES |
|------------|--|---------------|--------|--|
| 2686 | Census Track Flag (trtcode1-trtcode10) | 1 | SEER | Use in conjunction with Census Tract (centrt1990_1-centrt1990_10) 0 = No Census Tract 1 = 1970 Census tract definitions 2 = 1980 Census tract definitions 3 = 1990 Census tract definitions |
| 2687 | Census Tract Certainty (trtcert1990_1-trtcert1990_10) | 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 has only one census tract 9 = Unable to assign census tract based on available information |
| 2688 | Census Tract Certainty 2000 (trtcert2000_1-trtcert2000_10) | 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 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 |

| <u>COL</u> | FIELD | LENGTH | ecember 18, <u>SOURCE</u> | |
|------------|---|--------|------------------------------|--|
| 2689 | Census Tract Certainty 2010 (trtcert2010_1-trtcert2010_10) | 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 |
| 2690 | Census Tract Poverty Indicator (cenpvind1-cenpvind10) | 1 | SEER | 1= 0%-<5% poverty 2 = 5% to <10% poverty 3 = 10% to <20% poverty 4 = 20% to 100% poverty 9 = Unknown |
| 2691 | Rural Urban Commuting Area Codes 2000 (RUCA 00) (US only) (ruca00c1- ruca00c10) | 2 | SEER | 1 = Urban commuting area-Secondary code1.0,1.1,2.0,2.1,3.0,4.1,5.1,7.1,8.1, or 10.1 2 = Not an urban commuting area – all other RUCA codes except 99 9 = Primary RUCA code is 99 (valid unknown) 96=Invalid state abbreviation, county code, or census tract 97=Insufficient population data for county to determine RUCA code 98=Valid state/county/tract combination but RUCA code not found 99=Missing or unknown state, county or census tract |
| 2693 | Rural Urban Commuting Area Codes 2010 (RUCA 10) (US Only) (ruca10c1- ruca10c10) | 2 | SEER | 1 = Urban commuting area-Secondary code1.0,1.1,2.0,2.1,3.0,4.1,5.1,7.1,8.1, or 10.1 2 = Not an urban commuting area – all other RUCA codes except 99 9 = Primary RUCA code is 99 (valid unknown) 96=Invalid state abbreviation, county code, or |

| COL | FIELD | LENGTH | SOURCE | |
|------|---|--------|--------|---|
| | | | | census tract 97=Insufficient population data for county to determine RUCA code 98=Valid state/county/tract combination but RUCA code not found 99=Missing or unknown state, county or census tract |
| 2695 | Urban Rural Indicator Codes 2000 (URIC 00) (US Only) (ruic00c1- ruic00c10) | 2 | SEER | 1= All urban – the % of the pop in an urban area is 100% 2 = Mostly urban – the % of the pop in an urban are is between >=50% and <100% 3 = Mostly rural - the % of the pop in an urban area is between >0% and <50% 4 = All rural – the % of the pop in an urban area is 0% 9 = The % of the pop in an urban or rural area is unknown (valid unknown) 96=Invalid state, county code or census tract 97=Insufficient pop data for county to determine census indicator 98=Valid state/county/tract combination but census indicator not found 99=Missing or unknown state, county or census tract |
| 2697 | Urban Rural Indicator Codes 2010 (URIC 10) (US Only) (ruic10c1- ruic10c10) | 2 | SEER | 1= All urban – the % of the pop in an urban area is 100% 2 = Mostly urban – the % of the pop in an urban are is between >=50% and <100% 3 = Mostly rural - the % of the pop in an urban area is between >0% and <50% 4 = All rural – the % of the pop in an urban area is 0% 9 = The % of the pop in an urban or rural area is unknown (valid unknown) 96=Invalid state, county code or census tract 97=Insufficient pop data for county to determine census indicator 98=Valid state/county/tract combination but census indicator not found 99=Missing or unknown state, county or census tract |

| COL | FIELD | <u>LENGTH</u> | SOURCE | |
|------|--|---------------|---------------|--|
| 2699 | Rural-Urban Continuum code 1993 (rucd93ss1-rucd93ss10) | 2 | SEER* Stat | 00=Central counties of metro areas of 1 million population or more 01=Fringe counties of metro areas of 1 million population or more 02=Counties in metro areas of 250,000 to 1 million population 03=Counties in metro areas of fewer than 250,000 population 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 to 19,999, adjacent to a metro area 07=Urban population of 2,500 to 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 88=Unknown-Alaska/Hawaii State/not official USDA Rural-Urban Continuum code 99=Unknown/not official USDA Rural-Urban Continuum code More information can be found here: Rural-Urban |
| 2701 | Rural-Urban Continuum code 2003 (rucd03ss1-rucd03ss10) | 2 | SEER* Stat | 01=Counties of metro areas of 1 million population or more 02=Counties in metro areas of 250,000 to 1 million population 03=Counties in metro areas of fewer than 250,000 population 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 to 19,999, adjacent to a metro area 07=Urban population of 2,500 to 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 88=Unknown-Alaska/Hawaii State/not official |

| COL | <u>FIELD</u> | LENGTH | SOURCE | |
|------|---|--------|---------------|--|
| | | | | USDA Rural-Urban Continuum code 99=Unknown/not official USDA Rural-Urban Continuum code More information can be found here: Rural-Urban |
| 2703 | Rural-Urban Continuum code 2013 (rucd13ss1-rucd13ss10) | 2 | SEER* Stat | 01=Counties of metro areas of 1 million population or more 02=Counties in metro areas of 250,000 to 1 million population 03=Counties in metro areas of fewer than 250,000 population 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 to 19,999, adjacent to a metro area 07=Urban population of 2,500 to 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 88=Unknown-Alaska/Hawaii State/not official USDA Rural-Urban Continuum code 99=Unknown/not official USDA Rural-Urban Continuum code More information can be found here: Rural-Urban |
| 2705 | Health Service Area (hsa_ss1-hsa_ss10) | 4 | SEER* Stat | This variable was derived from the original Health Service Area definitions and is Described in more detail on this website: SEER*Stat |
| 2709 | Health Service Area (NCI Modified) (hsanci_ss1-hsanci_ss10) | 4 | SEER* Stat | This variable was derived from the original Health Service Area definitions, but has been modified and is described in more detail on this website: SEER*Stat |
| 2713 | County at DX geocode 1990 (cnty90dx1- cnty90dx10) | 3 | SEER | County codes can be found here: SPM_AppendixA.pdf |

| <u>COL</u> | FIELD | LENGTH | SOURCE | |
|------------|---|--------|--------|---|
| 2716 | County at DX geocode 2000 (cnty00dx1-cnty00dx10) | 3 | SEER | County codes can be found here: SPM_AppendixA.pdf |
| 2719 | County at DX geocode 2010 (cnty10dx1-cnty10dx10) | 3 | SEER | County codes can be found here: SPM_AppendixA.pdf |
| 2722 | Medicare Status Code (med_stcd1- med_stcd10) | 2 | EDB | 10 = Aged 11 = Aged with ESRD 20 = Disabled 21 = Disabled with ESRD 31 = ESRD only Variable is selected based on dates. The last status code prior to diagnosis is kept. |
| 2724 | 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) |
| 2725 | 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) |
| 2726 | First course of reconstruction (1998-2002) (recstr1-recstr10) (reconst) | 1 | SEER | 0 = No reconstruction/restoration 1 = Reconstruction, NOS (unknown if flap) 2 = Implant; reconstruction WITHOUT flap 3 = Reconstruction WITH flap, NOS 4 = Latissimus dorsi flap 5 = Abdominus recti flap 6 = Flap, NOS plus implant 7 = Latissimus dorsi flap plus implant 8 = Abdominus recti plus implant 9 = Unknown; not stated; death certificate only |
| 2727 | Radiation (rad1-rad10) (radiatn) | 1 | SEER | 0 = None 1 = Beam radiation 2 = Radioactive implants 3 = Radioisotopes 4 = Combination of beams with implants or isotopes 5 = Radiation, NOS method or source not specified 6 = Other radiation (1973-1987 cases only) 7 = Refused |

| COL | <u>FIELD</u> | LENGTH | SOURCE | |
|------|---|--------|--------|--|
| | | | | 8 = Recommended, unknown if administered 9 = Unknown |
| 2728 | Radiation to Brain and /or CNS (radbrn1-radbrn10) (rad_brn) | 1 | SEER | 0 = None 1 = Radiation 7 = Refused 8 = Recommended, unknown if administered 9 = Unknown |
| 2729 | Radiation sequence with surgery (radsurg1-radsurg10) (rad_surg) | 1 | SEER | 0 = No radiation and/or cancer-directed surgery 2 = Radiation prior to surgery 3 = Radiation after surgery 4 = Radiation before and after surgery 5 = Intraoperative radiation 6 = Intraoperative rad with other rad before/after surgery 7 = Surgery both before and after radiation 9 = Sequence unknown, but both were given |
| 2730 | 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 |
| 2731 | 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. |
| 2732 | 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. |
| 2733 | 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. |
| 2734 | 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. |

| COL | FIELD | LENGTH | scember 18, SOURCE | |
|------|---|--------|-----------------------|---|
| 2735 | 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 |
| 2736 | 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. |
| 2737 | 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. |
| 2738 | 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. |
| 2739 | 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. |
| 2740 | 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. |
| 2741 | 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 |

| COL | FIELD | De LENGTH | ecember 18, <u>SOURCE</u> | 2019 NOTES |
|------|--|--------------|------------------------------|---|
| | | | | conflict has been reviewed. |
| 2742 | 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. |
| 2743 | 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 |
| 2745 | 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 |
| 2747 | Month Therapy Started (monrx1-monrx10) (monthrx) | 2 | SEER | Month of Therapy 00 = No Therapy 01-12 = Valid month blank = Unknown |
| 2749 | 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. |
| 2753 | 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 |

| COL | <u>FIELD</u> | <u>LENGTH</u> | SOURCE | NOTES |
|------|--|---------------|--------|---|
| 2754 | 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+) |
| 2755 | 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.

| <u>COL</u> | FIELD | LENGTH | scember 18, SOURCE | |
|------------|---|--------|-----------------------|---|
| 2757 | 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 65 = TRICARE 66 = Military 67 = Veterans Affairs 68 = Indian/Public Health Service 99 = Insurance status unknown |
| 2761 | OncoType DX recurrence score (onco_score1- onco_score10) (oncotype_score) | 3 | SEER | 000-100 *Special permission required to receive this variable |
| 2764 | OncoType DX risk group (onco_rg1-onco_rg10) (oncotype_rg) | 1 | SEER | 1 = Low 2 = Intermediate 3 = High *Special permission required to receive this variable |
| 2765 | OncoType DX reason no score (onco_rns1- onco_rns10) (oncotype_rns) | 1 | SEER | 1 = Cancelled Test – Cancelled by physician or patients before result was returned 2 = Failed Test – Inadequate tissue sample 3 = Outside Analysis Criteria 4 = Non-Unique – Patient had multiple tumors *Special permission required to receive this variable |

| COL | FIELD | <u>LENGTH</u> | SOURCE | <u>NOTES</u> |
|------|---|---------------|--------|--|
| 2766 | OncoType DX delivered year (onco_year1-onco_year10) (oncotype_year) | 4 | SEER | 2004-2015 *Special permission required to receive this variable |
| 2770 | OncoType DX delivered month (onco_month1-onco_month10) (oncotype_month) | 2 | SEER | 01-12 *Special permission required to receive this variable |
| 2772 | OncoType DX months since diagnosis (onco_time1-onco_time10) (oncotype_time) | 3 | SEER | 000-996 = Valid months 997 = Unknown 998 = Test before DX Unknown occurs when month of diagnosis is unknown and year of diagnosis is not greater than test year. *Special permission required to receive this variable |

December 18, 2019

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

December 18, 2019

- 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

December 18, 2019

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