Data Extract Creation for SEER Submission

Last Updated by Claire Daniels, 8/29/2007

1. The final extract will take the form of a flat text file. The file should be named using the following naming convention:
   rr.nov07.txd
   where rr represents the two-character abbreviation for the registry.

2. Each row in the text file should correspond with one tumor. The length of each row should be 1946 characters.

3. The following fields are required in the extract:

Position	Length	NAACCR Item #	NAACCR Item Name	Eureka Item Name	Source
1-1	1	10	Record Type	Record_Type	fixed value [I]
2-9	8	20	Patient ID Number	Patient_ID	P
19-19	1	50	NAACCR Record Version	NAACCR Record Version	fixed value [B]
20-29	10	40	Registry ID	Registry ID	fixed value [0000001541]
30-31	2	60	Tumor Record Number	Central_Tum_No	T
40-47	8			Reg_Pat_No	T
40-49	10	45	*NPI Registry ID	NPIRegistryID	
72-73	2	80	Addr at DX--State	Addr_DX_State	T
83-85	3	90	County at DX	County_DX	T
86-91	6	110	Census Tract 1970/80/90	Census_Tract_90	T
92-92	1	120	Census Cod Sys 1970/80/90	Census Cod Sys 1970/80/90	calculated
93-98	6	130	Census Tract 2000	Census_Tract_00	T
100-100	1	364	Census Tr Cert 1970/80/90	Census_Cert_90	T
101-101	1	365	Census Tr Certainty 2000	Census_Cert_00	T
102-102	1	150	Marital Status at DX	Marital_Status	T
103-104	2	160	Race 1	Race_1	P
105-106	2	161	Race 2	Race_2	P
107-108	2	162	Race 3	Race_3	P
109-110	2	163	Race 4	Race_4	P
111-112	2	164	Race 5	Race_5	P
115-115	1	190	Spanish/Hispanic Origin	Spanish_Origin	P
116-116	1	200	Computed Ethnicity	Computed Ethnicity	calculated
117-117	1	210	Computed Ethnicity Source	Computed Ethnicity Source	calculated
118-118	1	220	Sex	Sex	P
119-121	3	230	Age at Diagnosis	Age_At_Diagnosis	calculated
122-129	8	240	Birth Date	Birth_Date	P
130-132	3	250	Birthplace	Birth_Place	P
231-231	1	191	NHIA Derived Hisp Origin	NHIA_Derived_Hisp_Origin	P
232-232	1	192	IHS Link	IHSLink	P

281-282	2	380	Sequence Number--Central	Seq_No_Central	T
283-290	8	390	Date of Diagnosis	Date_DX	T
291-294	4	400	Primary Site	Site	T
295-295	1	410	Laterality	Laterality	T
296-299	4	420	Histology (92-00) ICD-O-2	Hist_Type_2	T
300-300	1	430	Behavior (92-00) ICD-O-2	Hist_Behavior_2	T
301-304	4	522	Histologic Type ICD-O-3	Hist_Type_3	T
305-305	1	523	Behavior Code ICD-O-3	Hist_Behavior_3	T
306-306	1	440	Grade	Hist_Grade	T
311-311	1	490	Diagnostic Confirmation	DX_Conf	T
312-312	1	500	Type of Reporting Source	Report_Source	T
322-323	*2*	*501*	*Casefinding Source	casefindingSource	
324-324	1	442	Ambiguous Terminology	AmbiguousTerminology	T
325-332	8	443	Date of Conclusive DX	DateOfConclusiveDX	T
333-334	2	444	Multi Tum Rept As One Prim	MultiTumRptAsOnePrim	T
335-342	8	445	Date of Multiple Tumors	DateOfMultipleTumors	T
343-344	2	446	Multiplicity Counter	MultiplicityCounter	T
445-446	2	630	Primary Payer at DX	Pay_Source_1	
528-528	1	759	SEER Summary Stage 2000	Sum_Stage_00	T or calculated
529-529	1	760	SEER Summary Stage 1977	Sum_Stage_77	T or calculated
531-533	3	780	EOD--Tumor Size	Tum_Size	T
534-535	2	790	EOD--Extension	Extension	T
536-537	2	800	EOD--Extension Prost Path	Extension_Path	T
538-538	1	810	EOD--Lymph Node Involv	Nodes_Involved	T
539-540	2	820	Regional Nodes Positive	Nodes_Pos	T
541-542	2	830	Regional Nodes Examined	Nodes_Exam	T
562-562	1	870	Coding System for EOD	Coding System for EOD	calculated
626-626	1	1150	Tumor Marker 1	Tum_Markers_1	T
627-627	1	1160	Tumor Marker 2	Tum_Markers_2	T
628-628	1	1170	Tumor Marker 3	Tum_Markers_3	T
629-631	3	2800	CS Tumor Size	CS_Tum_Size	T
632-633	2	2810	CS Extension	CS_Ext	T

635-636	2	2830	CS Lymph Nodes	CS_LN	T
638-639	2	2850	CS Mets at DX	CS_Mets_DX	T
641-643	3	2880	CS Site-Specific Factor 1	CS_Site_Spec_F1	T
644-646	3	2890	CS Site-Specific Factor 2	CS_Site_Spec_F2	T
647-649	3	2900	CS Site-Specific Factor 3	CS_Site_Spec_F3	T
650-652	3	2910	CS Site-Specific Factor 4	CS_Site_Spec_F4	T
653-655	3	2920	CS Site-Specific Factor 5	CS_Site_Spec_F5	T
656-658	3	2930	CS Site-Specific Factor 6	CS_Site_Spec_F6	T
659-660	2	2940	Derived AJCC T	Derived_AJCC_T	T
662-663	2	2960	Derived AJCC N	Derived_AJCC_N	T
665-666	2	2980	Derived AJCC M	Derived_AJCC_M	T
668-669	2	3000	Derived AJCC Stage Group	Derived_AJCC_Stg_Grp	T
670-670	1	3010	Derived SS1977	Derived_SS1977	T
671-671	1	3020	Derived SS2000	Derived_SS2000	T
672-672	1	3030	Derived AJCC--Flag	Derived_AJCC_Flag	T
673-673	1	3040	Derived SS1977--Flag	Derived_SS1977_Flag	T
674-674	1	3050	Derived SS2000--Flag	Derived_SS2000_Flag	T
705-710	6	2935	CS Version 1st	CS_Version_1st	T
711-716	6	2936	CS Version Latest	CS_Version_Latest	T
835-842	8	1260	Date of Initial RX--SEER	Date_RX	T
859-860	2	1290	RX Summ--Surg Prim Site	Surg_Prim_Sum	T
861-861	1	1292	RX Summ--Scope Reg LN Sur	Scope_LN_Sum	T
862-862	1	1294	RX Summ--Surg Oth Reg/Dis	Surg_Other_Sum	T
863-864	2	1296	RX Summ--Reg LN Examined	Surg_LN_Ex_Sum	T
867-867	1	1330	RX Summ--Reconstruct 1st	Surg_Sum_Recon	T
868-868	1	1340	Reason for No Surgery	Reason_No_Surg	T
873-873	1	1360	RX Summ--Radiation	Rad_Sum	T
874-874	1	1370	RX Summ--Rad to CNS		calculated
875-875	1	1380	RX Summ--Surg/Rad Seq	Rad_Seq	T
876-877	2	3250	RX Summ--Transplnt/Endocr	Transp_Endo_Sum	T
878-879	2	1390	RX Summ--Chemo	Chemo_Sum	T
880-881	2	1400	RX Summ--Hormone	Horm_Sum	T

882-883	2	1410	RX Summ--BRM	Immuno_Sum	T
884-884	1	1420	RX Summ--Other	Other_RX_Sum	T
888-889	2	1460	RX Coding System--Current	RX Coding System--Current	calculated
931-931	1	1639	RX Summ--Systemic Sur Seq	RXSummSystemicSurSeq	T
939-940	2	1646	RX Summ--Surg Site 98-02	Surg_Prim_Sum_98_02	T
941-941	1	1647	RX Summ--Scope Reg 98-02	Scope_LN_Sum_98_02	T
942-942	1	1648	RX Summ--Surg Oth 98-02	Surg_Other_Sum_98_02	T
1124-1124	1	1990	Over-ride Age/Site/Morph	OR_Age_Site	T
1125-1125	1	2000	Over-ride SeqNo/DxConf	OR_Seq_DX_Conf	T
1126-1126	1	2010	Over-ride Site/Lat/SeqNo	OR_Site_Lat_SeqNo	T
1127-1127	1	2020	Over-ride Surg/DxConf	OR_Surg_DX_Conf	T
1128-1128	1	2030	Over-ride Site/Type	OR_Site_Type	T
1129-1129	1	2040	Over-ride Histology	OR_Hist_Behavior	T
1130-1130	1	2050	Over-ride Report Source	OR_DC_Seq	T
1131-1131	1	2060	Over-ride Ill-define Site	OR_Multi_Ildef	T
1132-1132	1	2070	Over-ride Leuk, Lymphoma	OR_Lymph_Leuk	T
1133-1133	1	2071	Over-ride Site/Behavior	OR_Site_Behavior	T
1134-1134	1	2072	Over-ride Site/EOD/DX Dt	OR_Site_Stage	T
1135-1135	1	2073	Over-ride Site/Lat/EOD	OR_Site_Lat_EOD	T
1136-1136	1	2074	Over-ride Site/Lat/Morph	OR_Site_Lat_Hist	T
1147-1147	1	1980	ICD-O-2 Conversion Flag		
1198-1198	1	2120	SEER Coding Sys--Current	SEER Coding Sys--Current	fixed value [7]
1199-1199	1	2130	SEER Coding Sys--Original	SEER Coding Sys--Original	calculated
1214-1214	1	2180	SEER Type of Follow-Up		calculated
1215-1216	2	2190	SEER Record Number		calculated
1243-1243	1	2116	ICD-O-3 Conversion Flag	ICDO3_Conv_Flag	T
1294-1301	8	1750	Date of Last Contact	Date_Lat_Pat_FU	P
1302-1302	1	1760	Vital Status	Vital_Status	P
1388-1391	4	1910	Cause of Death	Cause_Death	P
1392-1392	1	1920	ICD Revision Number	ICD Revision Number	calculated

* Not added to extract yet

P = select from patient

T = select from tumor

The “position” column refers to the position in the text file that each data item should occupy. This corresponds with the NAACCR Layout version 11.

4. Identify which tumors to include using the following selection criteria:
   - Tumor Deleted Flag is NULL.
   - Date of Diagnosis is in the range 1988 to **2005**.
   - The tumor matches one of the following site and histology combinations:
     a) The tumor is invasive (ICD-O-3 histology behavior code is 3) and
        - is not a skin cancer (sites C440-C449) with the following ICD-O-3 histology types: 8000-8005, 8010-8046, 8050-8084, 8090-8110; OR

     b) The tumor is in situ (ICD-O-3 histology behavior code is 2), and
        - is not a skin cancer (sites C440-C449) with the following ICD-O-3 histology types: 8000-8005, 8010-8046, 8050-8084, 8090-8110, and
        - is not a cervical cancer (site C530-C539) diagnosed 1996 or later; OR

     c) ICD-O-3 behavior is 0 (benign) or 1 (borderline) for the intracranial and central nervous system sites (C700 – C709, C710 – C719, C720 – C729, C751 – C753) for diagnosis years 2004 and later; OR
     d) The tumor was diagnosed prior to 2001, the site is ovary (C569), the ICD-O-2 histology type is 8442, 8451, 8462, 8472 or 8473, and the ICD-O-2 histology behavior code is 3.
   - County of Diagnosis is a known California county (i.e. codes 001 to 058).

Note: Reportability is defined by the date of diagnosis and the coding manual that was in effect for that date of diagnosis. The SEER Coding Manual for 2004 defines the reportability for cases diagnosed in 2004 and forward (until a new coding manual is in effect) and it can't be used for cases diagnosed prior to 2004.

5. Perform the following conversions for each item, or use the rules given to generate the item.

Record Type [10]
Set to "I".

Patient ID Number [20]
Right justify and zero fill.

NAACCR Record Version [50]
Set to "B".

Registry ID [40]
Set to 0000009700 (California Cancer Registry) if generating a file containing all records.
Set to 0000001541 (SEER California excluding LA, SJ & SF) if generating a file containing records with a County of Diagnosis in Greater California (Regions 2-6 and 7/10).

Tumor Record Number [60]
Right justify and zero fill.

Regional Patient Number (not a NAACCR data item)
Select the primary regional patient number for the tumor. If there is none, select blank.

Addr at DX—State [80]
Convert to upper case.

County at DX [90]
Convert to a FIPS county code. Right justify and zero fill. Eureka codes can be converted to FIPS codes by multiplying the Eureka code by 2 and subtracting 1.

Census Tract 1970/80/90 [110]
Valid SEER Census Tract 1970/80/90 codes: Blank, 000000, 000100-949999, 999999.
Convert California-specific codes 999993-999998 to SEER code 999999.

Census Cod Sys 1970/80/90 [120]
If Census Tract 1970/80/90 is blank, then set to blank.
If Census Tract 1970/80/90 is 000000, then set to 0.
Otherwise, set to 3.

Census Tract 2000 [130]
Valid SEER Census Tract 2000 codes: Blank, 000000, 000100-999998, 999999.
Convert California-specific codes 999993-999998 to SEER code 999999.

Census Tr Cert 1970/80/90 [364]
No conversion required.

Census Tr Certainty 2000 [365]
If blank and Date of Diagnosis is 2000 or later, then convert to 9.

Marital Status at DX [150]
No conversion required.

Race 1 [160], Race 2 [161], Race 3 [162], Race 4 [163], Race 5 [164]
Convert Eureka code 90 to NAACCR code 96.

Spanish/Hispanic Origin [190]
No conversion required.

Computed Ethnicity [200]

If Date of Diagnosis is before 1994, then set to blank.
Otherwise, use the CCR Volume III logic for SPANISH-SURNAME, namely:

```
If not male (Sex ≠ 1)
        If there is no Maiden_Name
                move 3 to SPANISH-SURNAME
        else if Maiden_Name* is Spanish
                move 7 to SPANISH-SURNAME
        else
                move 1 to SPANISH-SURNAME
else (male)
        move 2 to SPANISH-SURNAME.

If SPANISH-SURNAME = 7 stop.

If Last_Name* is Spanish
        if SPANISH-SURNAME = 3
                move 6 to SPANISH-SURNAME
        else if SPANISH-SURNAME = 1
                move 4 to SPANISH-SURNAME
else (Hispanic male)
        move 5 to SPANISH-SURNAME.
```

* If Name is hyphenated and the hyphenated name is not found on the Spanish Surname table, look for each portion of the name on the Spanish Surname table.

When checking to see whether a Maiden Name or Last Name is on the Spanish Surname table, alias Maiden Names and alias Last Names should also be checked.

The Spanish Surname table is listed in Appendix O of CCR Volume I.

Computed Ethnicity Source [210]
If Date of Diagnosis is before 1994, then set to blank.
Otherwise, set to 6.

Sex [220]
No conversion required.

Age at Diagnosis [230]
Calculate according to CCR Volume III, Appendix 6. The logic is reproduced below:

> If either century of birth or century of diagnosis is unknown (99), then AGE-DX = 999.
>
> CCYY DX - CCYY BIRTH = age.
>
> If month of birth or diagnosis is unknown (99), assume unknown month (and day): MMDD = 0701, otherwise, if day of birth or diagnosis is unknown (99), assume unknown day: DD = 15.
>
> If MMDD DX < MMDD Birth
> then AGE-DX = age - 1
> else AGE-DX = age.
>
> If age > 120 move 120 to AGE-DX.
>
> If age < 0, display an error message (DATE DX < DATE BIRTH).

If age is less than zero, then set it to zero instead of displaying an error message. Also, right justify and zero fill.

Birth Date [240]
Convert yyyymmdd to mmddyyyy. Set dd to 99.

Birthplace [250]
No conversion required.

NHIA Derived Hisp Origin [191]
No conversion required.

IHS Link [192]
No conversion required.

Sequence Number—Central [380]
No conversion required.

Date of Diagnosis [390]
Convert yyyymmdd to mmddyyyy. Set dd to 99.

Primary Site [400]
No conversion required.

Laterality [410]
No conversion required.

Histology (92-00) ICD-O-2 [420]
If Date of Diagnosis is 2001 or later, then convert to blank. (Note: in theory, this should be blank in Eureka.)

Behavior (92-00) ICD-O-2 [430]
If Date of Diagnosis is 2001 or later, then convert to blank. (Note: in theory, this should be blank in Eureka.)

Histologic Type ICD-O-3 [522]
No conversion required.

Behavior Code ICD-O-3 [523]
No conversion required.

Grade [440]
No conversion required.

Diagnostic Confirmation [490]
No conversion required.

Type of Reporting Source [500]
No conversion required.

SEER Summary Stage 2000 [759], SEER Summary Stage 1977 [760]
If Date of Diagnosis is 1993 or earlier, then

- Populate SEER Summary Stage 2000 from Sum_Stage_00
- Populate SEER Summary Stage 1977 from Sum_Stage_77.

If Date of Diagnosis is 1994-2003, then

- Use SEER's Summary Stage program to calculate SEER Summary Stage 2000 and SEER Summary Stage 1977. This program is written in C and the source code can be obtained from SEER.

If Date of Diagnosis is 2004 or later, then

- Set SEER Summary Stage 2000 and SEER Summary Stage 1977 to blank.

EOD—Tumor Size [780], EOD—Extension [790], EOD—Extension Prost Path [800], EOD—Lymph Node Involv [810]
If Date of Diagnosis is 1993 or earlier, then fill with 9s.
If Date of Diagnosis is 2004 or later, then set to blank.

Regional Nodes Positive [820]
If Date of Diagnosis is 1993 or earlier, then fill with 9s.
If Date of Diagnosis is 1994-2003 and Regional Nodes Examined is "95" and Regional Nodes Positive is "95", then set to "97".

Regional Nodes Examined [830]
If Date of Diagnosis is 1993 or earlier, then fill with 9s.

Coding System for EOD [870]
If Date of Diagnosis is 2004 or later, then set to blank.
Otherwise, set to "4".

Tumor Marker 1 [1150], Tumor Marker 2 [1160]
If Date of Diagnosis is before 1990, then generate 9.
If Date of Diagnosis is 1990-2003 and Site is C500-C509 and Type of Reporting Source is 6, then generate 0.
If Date of Diagnosis is 1990-2003 and Site is C500-C509 and Type of Reporting Source is 7, then generate 9.
If Date of Diagnosis is 1990-2003 and Site is C500-C509 and Type of Reporting Source is not 6 or 7, then take the tumor-level value.
If Date of Diagnosis is 1998-2003 and Site is C619 or C620-C629 and Type of Reporting Source is 6, then generate 0.
If Date of Diagnosis is 1998-2003 and Site is C619 or C620-C629 and Type of Reporting Source is 7, then generate 9.
If Date of Diagnosis is 1998-2003 and Site is C619 or C620-C629 and Type of Reporting Source is not 6 or 7, then take the tumor-level value.
If Date of Diagnosis is 1990-1997 and Site is not C500-C509, then generate 9.
If Date of Diagnosis is 1998-2003 and Site is not C500-C509 or C619 or C620-C629, then generate 9.
If Date of Diagnosis is later than 2003, then generate blank.

Tumor Marker 3 [1170]
If Date of Diagnosis is before 1998, then generate 9.
If Date of Diagnosis is 1998-2003 and Site is C620-C629 and Type of Reporting Source is 6, then generate 0.
If Date of Diagnosis is 1998-2003 and Site is C620-C629 and Type of Reporting Source is 7, then generate 9.
If Date of Diagnosis is 1998-2003 and Site is C620-C629 and Type of Reporting Source is not 6 or 7, then take the tumor-level value.
If Date of Diagnosis is 1998-2003 and Site is not C620-C629, then generate 9.
If Date of Diagnosis is later than 2003, then generate blank.

Collaborative Staging Fields

- **CS Tumor Size [2800]**
- **CS Extension [2810]**
- **CS Lymph Nodes [2830]**
- **CS Mets at DX [2850]**
- **CS Site-Specific Factor 1 [2880]**
- **CS Site-Specific Factor 2 [2890]**
- **CS Site-Specific Factor 3 [2900]**
- **CS Site-Specific Factor 4 [2910]**
- **CS Site-Specific Factor 5 [2920]**
- **CS Site-Specific Factor 6 [2930]**
- **Derived AJCC T [2940]**
- **Derived AJCC N [2960]**
- **Derived AJCC M [2980]**
- **Derived AJCC Stage Group [3000]**
- **Derived SS1977 [3010]**
- **Derived SS2000 [3020]**
- **Derived AJCC—Flag [3030]**
- **Derived SS1977—Flag [3040]**
- **Derived SS2000—Flag [3050]**
- **CS Version 1st [2935]**
- **CS Version Latest [2936]**

No conversion required.

Date of Initial RX—SEER [1260]
Use the CCR Volume III logic for Date_RX, namely:

> Computer generate by choosing the earliest date from either Date_Surg, Date_Rad, Date_Chemo, Date_Horm, Date_Immuno , or Date_Other_RX, Date_Def_Surg and Date_Transp_Endo.

Also, use the following rules:

- If any date is blank or 8s, then convert it to 0s.
- If all dates are 0s, then Date RX is 0s.
- If all dates are 0s and 9s, then Date RX is 9s.
- Otherwise, select the earliest date that is not 0s or 9s.

Convert the resulting date to mmddyyyy format and set dd to 99.

RX Summ—Surg Prim Site [1290]
If Date of Diagnosis is before 1998, then set to blank.

RX Summ—Scope Reg LN Sur [1292]
If Date of Diagnosis is before 2003, then set to blank.

RX Summ—Surg Oth Reg/Dis [1294]
If Date of Diagnosis is before 2003, then set to blank.

RX Summ—Reg LN Examined [1296]
If Date of Diagnosis is not 1998-2002, then set to blank.

RX Summ—Reconstruct 1st [1330]
Select from tumor if Date of Diagnosis is 1998-2002 and Primary Site begins with C50, otherwise set to blank.

Reason for No Surgery [1340]
Use the first conversion that applies, if any:

- If Date of Diagnosis is 1998-2002 and (RX Summ—Surg Site 98-02 is 10-90 or RX Summ—Scope Reg 98-02 is 1-6 or RX Summ—Surg Oth 98-02 is 1-8), then set to 0.
- If Date of Diagnosis is 1998-2002 and RX Summ—Surg Site 98-02 is 00 and Type of Reporting Source is not 6 and Reason for No Surgery is 9, then set to 6.
- If Date of Diagnosis is before 2003 and Type of Report Source is 6, then set to 2.
- If Date of Diagnosis is before 2003 and Reason for No Surgery is 5, then set to 2.
- If Date of Diagnosis is 2003 or later and Type of Reporting Source is 6 or 7, then set to 9.
- If Date of Diagnosis is 2003 or later and RX Summ—Surg Prim Site is 00 and Reason for No Sugery is 9, then set to 6.

RX Summ—Radiation [1360]
If Eureka has 0 for "RX Summ—Radiation" and either 7 or 8 for "Reason for No Radiation", then convert to Eureka's "Reason for No Radiation" value.

RX Summ—Rad to CNS [1370]
If the following conditions are true:

- Type of Reporting Source is 6, and
- Date of Diagnosis is before 1998, and
- Primary Site is C340-C349 (lung) or Histology Type 2 is 9800-9941 (leukaemia)

then set to 0.
Otherwise, set to 9.

RX Summ—Surg/Rad Seq [1380]
No conversion required.

RX Summ—Transplnt/Endocr [3250]
No conversion required.

RX Summ—Chemo [1390], RX Summ—Hormone [1400]
If Date of Diagnosis is 2002 or earlier and the Eureka value is 82, 85 or 86, then convert to 00.

RX Summ—BRM [1410]
No conversion required.

RX Summ—Other [1420]
No conversion required.

RX Coding System—Current [1460]
If Date of Diagnosis is before 2003, then set to 05 (treatment data coded according to 1998 SEER Manual).
Otherwise, set to 06 (treatment data coded according to FORDS Manual).

RX Summ—Systemic Sur Seq [1639]
If Date of Diagnosis is 2006 or earlier, then set to blank.

RX Summ—Surg Site 98-02 [1646], RX Summ—Scope Reg 98-02 [1647], RX Summ—Surg Oth 98-02 [1648]
If Date of Diagnosis is not 1998-2002, then set to blank.

Override Flags

Over-ride Age/Site/Morph [1990]
Over-ride SeqNo/DxConf [2000]
Over-ride Site/Lat/SeqNo [2010]
Over-ride Surg/DxConf [2020]
Over-ride Site/Type [2030]
Over-ride Histology [2040]
Over-ride Report Source [2050]
Over-ride Ill-define Site [2060]
Over-ride Leuk, Lymphoma [2070]
Over-ride Site/Behavior [2071]
Over-ride Site/EOD/DX Dt [2072]
Over-ride Site/Lat/EOD [2073]
Over-ride Site/Lat/Morph [2074]

Convert 0 to blank.

ICD-O-2 Conversion Flag [1980]

If Date of Diagnosis is 1991 or earlier, then set to 1 (primary site and morphology converted without review).
If Date of Diagnosis is 1992-2000, then set to 0 (primary site and morphology originally coded in ICD-O-2).
Otherwise, set to blank (not converted).

SEER Coding Sys—Current [2120]

If Date of Diagnosis is 2006 or earlier, then set to 7 (January 2004 SEER Coding Manual).
Otherwise, set to 8 (January 2007 SEER Coding Manual).

SEER Coding Sys—Original [2130]
Calculate based on the Date of Diagnosis:

Date of Diagnosis	Code	Meaning
January 1987-April 1988	1	1987 SEER Coding Manual
May 1988-December 1988 (including unknown month)	2	May 1988 SEER Coding Manual
1989-1991	3	January 1989 SEER Coding Manual
1992-1997	4	January 1992 SEER Coding Manual
1998-2002	5	January 1998 SEER Coding Manual
2003	6	January 2003 SEER Coding Manual
2004+	7	January 2004 SEER Coding Manual

SEER Type of Follow-Up [2180]
If Type of Reporting Source is 6 or 7, then set to 1 ("Autopsy Only" or "Death Certificate Only" case);
else if Primary Site is C530-C539 and either "Behavior (92-00) ICD-O-2" is 2 or "Behavior Code ICD-O-3" is 2, then set to 3 (in situ cancer of the cervix uteri only);
otherwise set to 2 (active follow-up case).

SEER Record Number [2190]
Generate a sequential number for each record submitted per patient, ordered by Sequence Number Central. For example, if the patient has 5 tumors in Eureka, but only 3 are being submitted, with sequence numbers 02, 03 and 05, then the tumor with sequence number 02 should have a SEER Record Number of 01, the tumor with sequence number 03 should have a SEER Record Number of 02, and the tumor with sequence number 05 should have a SEER Record Number of 03. The SEER Record Number should be generated separately for each file that is being submitted.

ICD-O-3 Conversion Flag [2116]
No conversion required.

Date of Last Contact [1750]
Convert yyyymmdd to mmddyyyy. Set dd to 99.

Vital Status [1760]
Convert Eureka code 0 to SEER code 4.

Cause of Death [1910]
No conversion required.

ICD Revision Number [1920]
If Eureka Vital Status is 1 (alive) then set to 0 (patient alive at last followup).
If Eureka Vital Status is 0 (dead) and Date of Last Contact is before 1999, then set to 9 (ICD-9).
If Eureka Vital Status is 0 (dead) and Date of Last Contact is 1999 or later, then set to 1 (ICD-10).

6. The following files should be generated:
    - One file containing all records. For this file, the Registry ID should be set to 0000009700 (California Cancer Registry).
    - One file containing records with a County of Diagnosis in Greater California (Regions 2-6 and 7/10). For this file, the Registry ID should be set to 0000001541 (SEER California excluding LA, SJ & SF).

7. The final submission files should be sorted by:
    - Registry ID (NAACCR Item #40)
    - Patient ID Number (NAACCR Item #20)
    - Sequence Number Central (NAACCR Item #380)

8. The submission files should be run through SEER*Edits and any edit errors should be corrected.

References

SEER Web Site: http://seer.cancer.gov/
SEER Data Submission Requirements
SEER Program Coding and Staging Manual
NAACCR Web Site: http://www.naaccr.org/
NAACCR Standards for Cancer Registries, Volume II (currently using version 11)
California Cancer Reporting System Standards, Volume I, Appendices
California Cancer Reporting System Standards, Volume III