

Surveillance Research Program, NCI, Technical Report #2015-01

Data Nuances – Tips in analyzing Collaborative Stage (CS) data with a focus on CS version 2 for data 2010+ (November 2012 and 2013 SEER data submissions)

Lynn A. G. Ries and Brenda K. Edwards

This document is meant as a technical appendix to the articles published in a supplement to Cancer (<http://www.seer.cancer.gov/publications/collabstaging/>). It focuses on the Collaborative Stage (CS) data fields in SEER*Stat, a software package designed to analyze the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program's data. It gives pointers and discusses pitfalls that an analyst using the CS data items should understand before attempting to analyze this complex data. It focuses on the site-specific factors (SSFs), specifically for 2010, the first complete year of data collection for many of the SSFs. However, this is not meant to minimize the wealth of information available in the other CS fields such as tumor size, extension, lymph nodes, and metastases at diagnosis. SEER has collected extent of disease (EOD) information since it began in 1973. This EOD information can be collapsed into different staging systems. CS collects EOD information and then a computer algorithm derives the stages and components of stage. The collection of CS started with cases diagnosed in 2004 and was designed to be compatible with AJCC Cancer Staging Manual, 6th Edition (AJCC 6th) T, N, M and stage¹ and Summary Stage 1977 (SS77) and Summary Stage 2000 (SS200)² and SEER Historic stage² and will be referred to as CSv1. A major update in CS occurred for cases diagnosed in 2010+ for compatibility with AJCC Cancer Staging Manual, 7th Edition (AJCC 7th) T, N, M, and stage³ and was backward compatible with AJCC 6th and SS77 and SS2000 and added many clinically relevant SSFs. This major update is referred to as CSv2. [Figure 1](#) gives a timeline of the EOD collection in SEER and the different staging systems that were in effect; the dark green arrow bars show the years that the staging types (AJCC 6th, AJCC 7th, Historic, SS77 and SS2000) are available for most sites in SEER*Stat. Besides the two major versions, there were more minor revisions – see section on Version Control below. This document focuses on CS v0204. For ease of reference, the SEER*Stat data dictionary field names are used for the different CS fields in this document.

Note: <https://cancerstaging.org/cstage/> is the official web site for CS. From here one can navigate to different versions of CS including site/histology groupings called schemas, documentation, coding rules, etc. Navigating to hyperlinks listed in the remaining document for specific CS v0204 topics may yield a 'credential' error. Please just close the error window (top right marked by an X) and then the

correct web page should appear. In addition, the variable names given in this document correspond to the names of schemas in CS.

Schema Selection:

In the CS system, the primary site and histology define which schema was used to collect all of the CS information under that version. For a few schemas, additional information may be needed to define the schema and that information is stored in SSF 25. The URL (<https://cancerstaging.org/cstage/schema/Pages/version0204.aspx>) [if it asks you to sign in, click the 'X' on the right side of the top border] lists all of the schemas for CS v0204. To make it easier to find which schema includes a particular primary site and histology, there is an application (app) which will give you the schema name: <http://www.seer.cancer.gov/seertools/cstest/> by just filling in site, histology, and behavior (2 = in situ and 3 = invasive) and it will return the schema name. [Table 1](#) gives the primary sites and histologies that define each schema.

The most notable change in schemas between the AJCC 6th and 7th editions was the addition of melanoma for head and neck cancers. Another notable change was the split from Colon into Colon, NETColon, Appendix, GISTColon, and CarcinoidAppendix. SSF 25 was defined when additional information beyond the International Classification of Diseases for Oncology, Third Edition (ICD-O-3)⁴ site and histology was needed to define a schema. [Table 1](#) details the schemas that need additional information beyond site and histology and the SSF 25 codes that distinguish one schema from another.

When one analyzes T, N, M and AJCC stage data, it is important to use the correct CS schema recode in SEER*Stat. If one is analyzing trends in T, N, M or stage based on AJCC 6th, one should use the CS Schema - AJCC 6th edition variable in SEER*Stat. If one is analyzing SSFs, especially 2010+, one should use CS Schema v0204. CS Schema v0204 should also be used for trends in T, N, M or stage based on AJCC 7th.

Schemas versus SEER Site Recode

Similarly, there is a [SEER site recode](#) which is based on site and histology. SEER site recodes are a list of site/histology groupings which are traditionally used in publications such as the [SEER Cancer Statistics Review](#) and the [Annual Report to the Nation](#). Some schemas, however, are defined slightly differently from the traditional SEER site recodes. Just like the SEER site recode, the CS schemas use both the topography (site code) and morphology of the ICD-O-3⁴. For example, each case with an ICD-O-3

topography for lung of C34.0-C34.3 or C34.8-C34.9 would fall into either the Lymphoma schema if the histology is lymphoma or the Lung schema if the histology isn't a lymphoma. For the same topography codes, the SEER site recode would also separate lung and the lung lymphomas which would be in the lymphomas. [Table 2](#) shows the SEER site recode versus the CS Schema v0204 recode. For example, if one chooses Stomach in the SEER site recode, one would end up analyzing several different CS v0204 schemas: Stomach, NETStomach, GISTStomach, and EsophagusGEJunction. Part of what is generally considered 'colon' in SEER publications would be in multiple schemas: Colon, NETColon, GISTColon, Appendix, GISTAppendix, CarcinoidAppendix, and DigestiveOther in CS v0204. Since some of the SSFs are different for these schemas, the analyses could include data that shouldn't be combined since the SSFs may be different. For example, if one chose colon under site recode and chose SSF 11 (Mitotic count): for the colon schema this field would be 988 (not coded) but for NETColon schema this field would have data on mitotic count. Whenever data on SSFs is analyzed, one should use the schema variable and not SEER site recode. [Table 3](#) shows the CS Schema v0204 by the SEER site recode. If one chose MelanomaSkin in the CS Schema v0204 variable, one would get melanoma as defined in the SEER site recode but also included would be skin of male and female genital sites which is grouped with the female or male genital site in the SEER Site recode instead of melanoma.

Schema versus AJCC 6th and 7th

AJCC 6th and 7th recommend that T, N, M, and stage be applied to a more limited subset of histologies than those listed in the schema definition ([Table 1](#)); for many sites this excludes sarcomas from the analyses. These histologies are designated as not applicable (NA) in the T, N, M and AJCC stage so that they can be easily excluded from analyses.

Orientation to a CS schema CS v0204

From the official CS web site (<https://cancerstaging.org/cstage/>) one can navigate to a particular schema for documentation on the CS codes and algorithm. As one navigates to a specific schema for a particular CS version, the first 'page' is the schema index page which contains a list of the fields and tables which are defined for that schema (Figure 2). The schema index page displays each defined field for that schema. If the field is undefined, it shows a 988 on the schema index page if it is a three-digit field and 98 if it is a two-digit field. The field names are hyperlinked to go to the individual data fields and their codes and definitions. Note there are variables defined for all schemas such as Lymph vascular invasion

(required only for Penis and Testis) and the CS Mets at DX fields for bone, brain, lung, and liver which are considered CS fields but don't appear on the schema index page.

For the SSFs, the name of the SSF is given on the schema index page to make it easy to hyperlink to the SSF that one wants. Many SSFs were defined but they weren't required by any standardsetter (NCI, Centers for Diseases Control and Prevention (CDC), the Commission on Cancer (CoC), and the Canadian Council of Cancer Registries (CCCR)). Each standardsetter decided which SSFs they would require for data collection under a specific CS version number – see Version Control below. There is an application that will generate a table of the required SSFs for each standardsetter in all of the CSv2 versions, <http://www.seer.cancer.gov/csregstatus/>. Just select 'Get Started' on the right and then select the '+' on the next page and select from the dropdown boxes.

Not all of the SSFs are released on the SEER research file. Some SSFs were never required and other SSFs are still being evaluated for completeness before a decision will be made on their release. Others aren't being released because after review, they were found to be incomplete or inaccurate. In the November 2012 submission, some SSFs were released under a custom data request. As of the April 2014 release those SSFs are included in the research file. The list of SSFs available on the November 2013 research file (released April 2014) is available on-line: <http://seer.cancer.gov/seerstat/databases/ssf/>

On the schema index page after SSF 25, there is a histology inclusion table for AJCC 7th which shows the histologies for which AJCC 7th defines T, N, M and stage. The next group is a histology exclusion table for AJCC 6th; this table defines the histologies for which T, N, M and stage are not defined. The histologies for which AJCC 6th and 7th are defined form a subset of the cases for which the schema is used. The AJCC TNM 7 Stage, AJCC TNM 6 Stage and Summary Stage tables documents which variables and values define AJCC 7th, 6th, and Summary Stages, respectively. If any other tables are shown, they are additional documentation needed to show how the values from various fields are combined to define T, N, M, stage, or intermediary values.

The CS Schema on-line pages also show the mapping of the different fields to the derived fields such as T, N, M and various stages. For example, CS Mets at DX would have four columns to the right of the description that contain the mapping to the M code for AJCC 7th and AJCC 6th and to localized, regional or distant stage for SEER SS77 and SS2000, respectively. The extra tables are provided for

documentation when the mapping is based on more than one data field, e.g., when T is based on both tumor size and extension.

While it appears that there are unique site-specific tables for each schema, many schemas share some definitions for fields. For example, there are only a couple of unique tables for Regional Nodes Positive. From the schema index page, the 'Revision Info' tab at the bottom in CS v0204 shows the internal ID for each table and whether it is used by other schemas (i.e., common tables). Click on 'common' to see which schemas share a particular table. Items such as extension, regional nodes, and mets at DX tend to be schema specific and therefore schemas cannot be combined for this type of analysis. CS 'revision info' is not listed in CSv0205.

Version Control

Most SEER data items were collected for a particular year of diagnosis. If one were interested in how a particular field was coded for a case diagnosed in 2011, one could go the SEER coding manual used for that year of diagnosis and look up the codes. For CS, it is a bit more complicated. For CS, there are three flags (CS version input original [NAACCR item #2935]; CS version input current [NAACCR item #2937]; CS version derived (2004+) [NAACCR item #2936]) that point to the CS version the case was originally coded, was last updated and had the derivations performed, respectively. Each of these flags contains a version number to track under which set of codes and rules the case was originally coded, updated or the derivation fields calculated. So for 2004 cases, they should be coded in CSv1 but if they were originally processed after the registry started using CSv2, they would have to follow the requirements and codes for CSv2. All cases diagnosed 2010+, had to be processed in CSv2. However, a registry could decide when they wanted to operationalize CSv2. In order to not have to maintain multiple CS versions, once CSv2 was in operation all new data processed including cases 2004-2009 had to follow the codes, coding instructions, and requirements for CSv2. Nearly every year there were updates to CS and changes in requirements. These three administrative flags for version control have the first two digits to designate the major revision (01 for CSv1 and 02 for CSv2) and the next 4 digits the specific version within the major revision. For analysts, the CS version control numbers can point to the correct documentation needed for analyses. The item CS Version Input Original [NAACCR item #2935] reflects the CS version that was used to originally code the case including the codes and coding instructions for that version. The coding of the 2010 cases could have been coded under the rules and requirements for any version of CSv2 depending on when each registry implemented the latest versions and when they processed the case. In some instances, the derivation algorithm evaluates CS version

input original to determine which set of criteria is used to calculate T, N, M, or any of the various stages. CS version Input current [NAACCR item #2937] reflects the CS version last used to update and/or convert the case. Note that the CS version input current is not specific to a particular CS field; that is, not all CS items had to be reviewed under the latest version when one field was updated. The bottom line is that the CS version input current is based on when the case was originally coded but can be updated based on the version used to update or convert one or more CS data items. The CS version derived (2004+) [NAACCR item #2936] variable shows the CS version used to derive the CS Derived item fields including T, N, M, AJCC stage for both 6th and 7th (2010+ only) and SEER SS77 and SS2000. There is more information in the CS Implementation Guide for CS v0204 (<https://cancerstaging.org/cstage/coding/Documents/ImplementationGuide02.04.pdf>) [if it asks you to sign in, click the 'X' on the right side of the top border] regarding the 3 versions flags, the version numbers, and changes between versions.

SEER Data Submission

The data presented in the supplement to Cancer (<http://onlinelibrary.wiley.com/doi/10.1002/cncr.v120.S23/issuetoc>) are from the November 2012 submission of data to NCI. The data through diagnosis year 2010 (November 2012 submission) were released on the SEER research file in April 2013. Similarly data through 2011 (November 2013 submission) were released in April 2014 for analyses. The CS derived items should be derived by computer algorithm under CS v0204 for all cases 2004+ for both submissions – see CS version derived (2004+) – [NAACCR item #2936] above under version control. The documentation for CS v0204 can be used for analyses of any of the 2004-2011 CS data with a couple of caveats. CS v0204 documents all of the valid codes for CS v0204 ([Valid, Invalid, Obsolete Tables](#)) and in addition, documents all codes which were valid under all previous versions including obsoleted, reviewed, and changed codes. For obsolete and modified codes, it also documents the version under which the change occurred. CS v0204 does NOT document under which version a new code was added. Earlier versions of the documentation should be used to determine this or by comparing the list of valid codes across versions. Especially if there is a large increase in one code, it should be checked to make sure that it is a real increase instead of a new code which was added later.

Data Dictionary

Explanations for CS data items vary in specificity. A few are listed here. A very simplistic non-schema specific explanation is given below in the sections under CS Information collected. A data dictionary for variables released under SEER*Stat is given on the SEER web site:

<http://www.seer.cancer.gov/data/seerstat/nov2012/seerstat-variable-dictionary-nov2012.pdf>

The data descriptions are very short and are not meant as complete documentation for the file but rather pointers to where additional information can be found and warnings about pitfalls using some of the data items. The following URL gives a little more detail about the variables in SEER*Stat and is adapted from NAACCR Volume II (www.naaccr.org):

(<http://www.seer.cancer.gov/data/seerstat/nov2012/TextData.FileDescription.pdf>).

Documentation for the individual codes for a particular data item for a particular site/histology group is on-line and is listed under the schema name – see

<https://cancerstaging.org/cstage/schema/Pages/version0204.aspx> [if it asks you to sign in, click the 'X' on the right side of the top border]. To understand what a schema is please See Schema selection above.

The Coding Instructions for CS v0204 can be found at [if it asks you to sign in, click the 'X' on the right side of the top border]: <https://cancerstaging.org/cstage/coding/Pages/version0204.aspx>

Display Codes for Derived Items

The derived AJCC 6th and 7th T, N, M, stage, and descriptors in SEER*Stat show their display code to make selection easier. The display code is the description by which a human understands what the code means such as a derived AJCC 6th code of T1a. They are, however, stored and transmitted via a numeric code (or blanks or null) for ease of computer handling. [Tables 4-8](#) give the display codes and storage/transmission values. For example, a T1a is stored as a 12 for AJCC 6th and 120 for AJCC 7th. If you only analyze data under SEER*Stat, one doesn't have to understand the storage codes. However, if you are analyzing raw SEER data, the storage codes would have to be used.

Notes on Selected Schemas not Covered in the Supplement

Oral cavity and pharynx: Oral cavity and pharynx is broken into several different schemas in AJCC 7th. This includes the individual sub-sites such as BuccalMucosa, FloorMouth, GumLower, etc. plus melanomas for each of the sub-sites such as MelanomaBuccalMucosa, MelanomaEpiglottisAnterior, MelanomaFloorMouth, MelanomaGumLower, etc. There are very few melanomas of the oral cavity and

pharynx and for the analyst it may have been easier to add just one schema for all of the melanomas of the oral cavity and pharynx, but this would have caused incompatibility issues when the analyst wanted to look at all histologies for a particular sub-site of the oral cavity or extremely long lists of site-specific regional nodes would have to have been used. Therefore, individual schemas were designed based on the individual sub-site in the oral cavity and pharynx and its associated regional lymph nodes. In general, the lip and oral cavity sub-sites can all be grouped together and the melanomas of the oral cavity, pharynx, and larynx can be grouped together to analyze AJCC 7th T, N, M and stage for 2010+. For AJCC 6th, T, N, M, and stage, use the CS Schema Recode – AJCC 6th and then combine the sub-sites of the oral cavity to look at trends in these derived values. Analyses of individual CS extension, lymph nodes, mets fields, and SSFs need to be done on the specific schema for a particular sub-site and schemas should not be grouped for analyses unless the variable under analyses has the same codes and coding structures across all of the combined schemas. In some instances the same code number for a particular field such as CS extension means different things depending on the schema.

Testis: Since testis is not presented as a separate chapter in the Supplement, a couple of pointers here may help the analyst to understand why there are so many SSFs and which ones should be used. In AJCC 6th and 7th the Serum Tumor Marker designated by AJCC as S is needed for staging and is based on a combination of AFP, hCG, and LDH. In 2004 SSFs were introduced to collect this information, SSF1, SSF 2, and SSF 3, respectively. When AJCC 7th was first published, it listed that these 3 values should be from tests prior to orchiectomy: SSF 7, 9, 10, respectively, and limit SSF1, SSF2, and SSF3 values for cases that have been diagnosed and coded prior to CSv2. After the publication of AJC 7th, there was a correction that stated that post-orchiectomy test values should be used (SSF13, 15, and 16, respectively). For AJCC 6th, the mapping for the S value is based on SSF1, SSF2, and SSF3 if the case was originally coded under CSv1 and SSF13, SSF15, and SSF16 if it is a CSv2 case. For AJCC 7th, the stage is based on SSF13, SSF15, and SSF16.

Useful analytic tools

- SEER*Stat (<http://www.seer.cancer.gov/seerstat/>)
 - A SEER developed software package to analyze SEER incidence, survival, multiple primaries, etc.
 - Contains recodes to make analyses easier
 - For more information on CS variables, see <http://seer.cancer.gov/seerstat/variables/seer/ajcc-stage>
- Test-o-matic and SEER web selection tool (CS*Test - <http://www.seer.cancer.gov/seertools/cstest/>)

- Shows the schema name for a given site and histology (plus SSF25 where needed)
- Allows one to look at various CS input variables; the tool will derive various stage values (including T, N, and M) and yield errors if the input combinations are invalid.
- Valid and invalid codes - www.cancerstaging.org/cstage
 - Available for each version in CSv2
 - Valid codes: easy way to compare across versions to see when new codes were added
 - Valid codes are an easy way to limit the analyses to just the codes that are valid for a particular schema instead of assuming all values 000-999 are valid. Hint: the range of valid codes can be copied directly from the valid spreadsheet to the SEER*Stat 'Selected:' 'Unlabeled Values' when creating a variable.

CS information collected

- **CS tumor size (2004+)**: Information on tumor size. Available for 2004+. Earlier cases may be converted and new codes added which weren't available for use prior to the current version of CS. For more information, see <http://seer.cancer.gov/seerstat/variables/seer/ajcc-stage>.
- **CS extension (2004+)**: Information on extension of the tumor. Available for 2004+. Note: this item was originally a 2 digit field and was expanded to 3 digits during conversion to CSv2. Generally, a zero was added to the right of the existing 2 digit field except for 99 which became 999.
- **CS Tumor Size/Ext Eval (2004+)**: not required before 2007: Available for 2004+, but not required until 2007. Will be blank when it wasn't collected. Sometimes used in determining stage. Documents the basis of information used to code CS tumor size and CS extension.
- **CS lymph nodes (2004+)**: Information on involvement of regional lymph nodes.
- **CS Reg Node Eval (2004+)**: not required before 2007: Available for 2004+, but not required until 2007. Will be blank when it wasn't collected. Sometimes used in determining stage. Documents the basis of information used to code CS lymph nodes.
- **Regional nodes positive (1988+)**: Contains the number of positive regional lymph nodes based on biopsy or dissection. Coded under CS for 2004+. Cases coded 1988-2003 used slightly different definitions - see SEER coding manual for those years.
- **Regional nodes examined (1988+)**: Contains the number of regional lymph nodes examined based on biopsy or dissection. Coded under CS for 2004+. Cases coded 1988-2003 used slightly different definitions - see SEER coding manual for those years.
- **CS mets at dx (2004+)**: Information on involvement of distant sites and/or distant lymph nodes.
- **CS Mets Eval (2004+)**: not required before 2007: Available for 2004+, but not required until 2007. Will be blank when it wasn't collected. Sometimes used in determining M. Documents the basis of information used to code CS mets at dx.
- **CS mets at dx – bone, brain, lung, liver**: collected for CS v0202+ but made available for 2010+ since it is first full year for the information.
- **Grade path value and grade path system**: collected for 2010+ but information is missing for most cases. Information not on research file.
- **CS site-specific factor 25**: this is coded 988 unless it is needed to determine the schema. For example, it is needed to determine if tumors which started in the stomach should be grouped with GEJ instead of the Stomach schema.
- **CS site-specific factor 1- 24 (2004+)**: See site-specific schemas for SSFs and their codes and definitions. Some may only be available for certain years, e.g., 2010+.
- **Lymph-vascular Invasion (2004+)**: Required for cases originally coded under CSv2 or diagnosed 2010+ for the schemas for penis and testis only. On the research file LVI is shown for testis because it is needed for AJCC 6th staging. For penis, LVI is needed for AJCC 7th staging and so it is only available on the research file for 2010+. See ASCII text file description: <http://www.seer.cancer.gov/data/seerstat/nov2011/TextData.FileDescription.pdf>

Administrative CS fields

- **CS Schema v0204**: Based on site and histology, selection of the site group allows for the alignment of the CS codes with the schema that they were collected under. Use of SEER site recode may cause additional or fewer cases and should not be used when analyzing CS data unless one knows the differences. This recode should be used to analyze CS 2010 SSFs and AJCC 7th T, N, M and stage.
- **CS Schema – AJCC 6th Edition**: This recode should be used when analyzing AJCC 6th T, N, M, or stage in order to obtain the sites and histologies that were applicable.
- **CS version input original (2004+) [NAACCR item #2935]**: Data item shows what version was in effect the first time that CS was coded for this case. This data item along with CS version input current gives information on what document to use for the CS codes.
- **CS version input current (2004+) [NAACCR item #2937]**: Data item shows what version was in effect the last time that any CS field was updated for this case. Special codes for converted cases. This data item along with CS version input original gives info on what document to use for the CS codes.
- **CS version derived (2004+) [NAACCR item #2936]**: Data item shows what CS version was used to derive the CS Derived item fields including T, N, M, AJCC stage for both 6th and 7th (2010+ only) and SEER SS77 and SS2000.

Derived AJCC T, N, M, and stage AJCC 6th

- **Derived AJCC T, 6th ed (2004+):** The T category for AJCC 6th is derived from CS data elements for 2004+ cases. For some schemas, it is based on only CS extension, while for others it is based on size and extension, or a combination of extension and other variables. Not all schemas have a derived T.
- **Derived AJCC N, 6th ed (2004+):** The N category for AJCC 6th is derived from CS data elements for 2004+ cases. For some schemas, it is based on only CS lymph nodes while for others it is based on number of positive (SSFs usually) and CS lymph nodes, or a combination of CS lymph nodes and other variables. Not all schemas have a derived N.
- **Derived AJCC M, 6th ed (2004+):** The M category for AJCC 6th is derived from CS data elements for 2004+ cases. For some schemas, it is based on only CS mets at dx while for others it is based on a combination of CS mets at dx and other variables. Not all schemas have a derived M.
- **Derived AJCC Stage Group, 6th ed (2004+):** The stage category for AJCC 6th is derived from CS data elements for 2004+ cases. It is usually based on the derived T, N, and M for AJCC 6th but may also include other data items such as grade.
- **NA (not applicable):** is not an AJCC T, N, M, or stage category. It is used as a placeholder for histologies that AJCC does not define these groups. For many solid tumors, sarcomas are not staged. Excluding AJCC 6th stage = NA is the easiest way to exclude the histologies that AJCC does not provide AJCC 6th stage. Similarly for AJCC 6th stage. The histologies that aren't defined may be different for AJCC 6th compared to AJCC 7th.

AJCC 7th

- **Derived AJCC T, 7th ed (2010+):** The T category for AJCC 7th is derived from CS data elements for 2010+ cases. For some schemas, it is based on only CS extension, while for others it is based on size and extension, or a combination of extension and other variables. Not all schemas have a derived T.
- **Derived AJCC N, 7th ed (2010+):** The N category for AJCC 7th is derived from CS data elements for 2010+ cases. For some schemas, it is based on only CS lymph nodes while for others it is based on number of positive (SSFs usually) and CS lymph nodes, or a combination of CS lymph nodes and other variables. Not all schemas have a derived N.
- **Derived AJCC M, 7th ed (2010+):** The M category for AJCC 7th is derived from CS data elements for 2010+ cases. For some schemas, it is based on only CS mets at dx while for others it is based on a combination of CS mets at dx and other variables. Not all schemas have a derived M.
- **Derived AJCC Stage Group, 7th ed (2010+):** The stage category for AJCC 7th is derived from CS data elements for 2010+ cases. It is usually based on the derived T, N, and M for AJCC 7th but may also include other data items such as grade.
- **NA (not applicable):** is not an AJCC T, N, M, or stage category. It is used as a placeholder for histologies that AJCC does not define for these groups. In the schemas for many solid tumors, sarcomas are not staged. Excluding AJCC 7th stage = NA is the easiest way to exclude the histologies that AJCC does not provide AJCC 7th stage. Similarly for AJCC 6th stage. The histologies that aren't defined may be different for AJCC 7th compared to AJCC 6th.

Other derived stages: information on the availability of these derived stages and the years and SEER site recodes that information is available is provided at: <http://seer.cancer.gov/analysis/stage.html>

- **Derived SS1977 (2004+):** Derived from CS data items. It is a simplified version of stage: in situ, localized, regional, distant, & unknown. See SEER Summary Stage 2000 for definitions [<http://www.seer.cancer.gov/tools/ssm/>].
- **Derived SS2000 (2004+):** Derived from CS data items. It is a simplified version of stage: in situ, localized, regional, distant, & unknown. See SEER Summary Stage 2000 for definitions [<http://www.seer.cancer.gov/tools/ssm/>].
- **SEER historic stage A:** Available for many sites back to 1973. SEER Historic Stage A is Derived from CS data items for 2004+ and Extent of Disease (EOD) from 1973-2003. It is a simplified version of stage: in situ, localized, regional, distant, & unknown. See SEER Summary Stage 2000 for definitions [<http://www.seer.cancer.gov/tools/ssm/>].

Other recoded/derived information available in SEER*Stat:

- Schema recodes: recoded variables were defined in SEER*Stat to aid in distinguishing which schemas were defined under which version of CS – see Version Control above.
- **ER Status Recode Breast Cancer (1990+):** Created by combining information from Tumor marker 1 (1990-2003) [NAACCR item #1150], with information from CS site-specific factor 1 (2004+) [NAACCR item #2880]. This field is blank for non-breast cases and cases diagnosed before 1990.
- **PR Status Recode Breast Cancer (1990+):** Created by combining information from Tumor marker 2 (1990-2003) [NAACCR item #1150], with information from CS site-specific factor 2 (2004+) [NAACCR item #2880]. This field is blank for non-breast cases and cases diagnosed before 1990.
- **Derived HER2 Recode (2010+)** – Derives a summary HER2 for breast cases diagnosed 2010+. Created with combined information from several CS site-specific factors. For more information, see <http://seer.cancer.gov/seerstat/databases/ssf/her2-derived.html>.
- **Breast Subtype (2010+):** Created with combined information from ER Status Recode Breast Cancer (1990+), PR Status Recode Breast Cancer (1990+), and Derived HER2 Recode (2010+).
- **Lymphoma - Ann Arbor stage under Stage AJCC:** creates a derived Ann Arbor stage across different EOD/CS years of diagnosis for years 1983+.
- **Other recoded information in SEER*Stat:** <http://www.seer.cancer.gov/analysis/>

Summary

AJCC 6th T, N, M and stage are available for cases diagnosed in 2004+. AJCC 7th T, N, M, and stage are available for 2010+ cases. Data before 2004 had a more simplified version of extent of disease than CS. An on-going SEER project is to apply AJCC 6th criteria to the earlier data in order to create longer term trends such as AJCC 6th stage for cases diagnosed in 1988 and later. An overview of the stage data and data submissions in SEER*Stat can be found at <http://seer.cancer.gov/seerstat/variables/seer/ajcc-stage/>.

Some SSFs are available since 2004 but most are only available for 2010+. While the SSFs were required for data collection by version number, the data collection for the SSFs introduced in CSv2 would only be complete for 2010 forward and are therefore, only released for analyses for 2010+. Exception: when the data are needed to derive stage or any component of stage, they are shown for any earlier year that they were collected.

Schema is very important and the correct schema variable needs to be used depending on what is being analyzed. For example, AJCC 6th trends need the AJCC 6 schema variable to correctly select the cases. A dummy value of 'NA' was added to T, N, M and stage to indicate that AJCC does not recommend staging for these particular cases and they should be excluded from analyses of stage.

There is documentation for CS codes, coding structures, instructions, coding rules, how the codes are mapped to staging components, etc. on the www.cancerstaging.org/cstage web site. Within the schema there are links to most of the data items for that schema and in addition to the codes and descriptions, there are the mappings for the derived fields and some of the instructions for the field are included in the notes. More detailed coding instructions can be found in Part 1 (Sections 1 and 2) of the CS manual [<https://cancerstaging.org/cstage/coding/Pages/Version-02.05.aspx>]. In particular, the links to the coding instructions and the site-specific schemas may be invaluable in analyzing this complex information.

1. Greene FL, Page DL, Fleming ID, et al. AJCC Cancer Staging Manual. 6th Ed. Chicago: Springer-Verlag; 2002.
2. Young JL Jr, Roffers SD, Ries LAG, Friit AG, Hurlbut AA (eds), SEER Summary Staging Manual – 2000: Codes and Coding Instructions, National Cancer Institute, NIH Pub. No 01-4969, Bethesda, MD, 2001. Available on-line:
3. Edge SB, Byrd DR, Compton CC, et al. AJCC Cancer Staging Manual. 7th Ed. Chicago: Springer-Verlag; 2010.
4. Fritz A, Jack A, Parkin DM, et al. International Classification of Diseases for Oncology. 3rd ed. Geneva: World Health Organization; 2000.

Figure 1: Timeline of AJCC Stage, SEER EOD, and Derived Stage over Time

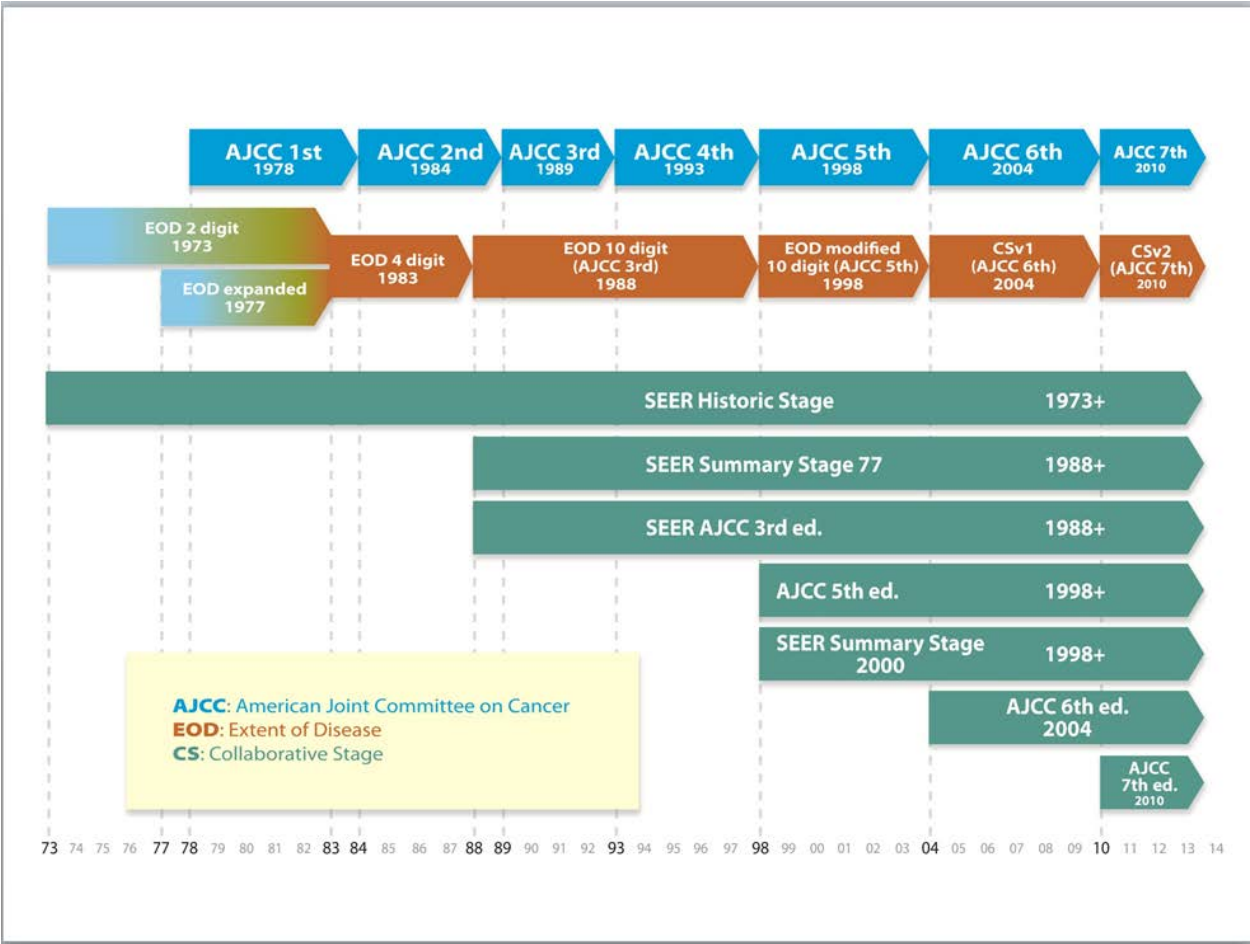


Table 1: List of CS Schema Names Based on Site and Histology and Sometimes Site Specific Factor (SSF) 25 for CS version 2.

AJCC 7th edition schemas names	ICD-O-3 Primary site: C _ _ _ . _ _	ICD-O-3 Histology: M- _ _ _ _ _	SSF 25 values when needed
LipUpper	000,003	8000-8713, 8800-9136,9141-9582,9700-9701	
LipLower	001, 004, 006	8000-8713, 8800-9136,9141-9582,9700-9701	
LipOther	002,005,008-009	8000-8713, 8800-9136,9141-9582,9700-9701	
TongueAnterior	020-023, 028-029	8000-8713, 8800-9136,9141-9582,9700-9701	
GumUpper	030	8000-8713, 8800-9136,9141-9582,9700-9701	
GumLower	031, 062	8000-8713, 8800-9136,9141-9582,9700-9701	
GumOther	039	8000-8713, 8800-9136,9141-9582,9700-9701	
FloorMouth	040-041,048-049	8000-8713, 8800-9136,9141-9582,9700-9701	
PalateHard	050	8000-8713, 8800-9136,9141-9582,9700-9701	
MouthOther	058-059, 068-069	8000-8713, 8800-9136,9141-9582,9700-9701	
BuccalMucosa	060-061	8000-8713, 8800-9136,9141-9582,9700-9701	
TongueBase	019, 024	8000-8713, 8800-9136,9141-9582,9700-9701	
PalateSoft	051-052	8000-8713, 8800-9136,9141-9582,9700-9701	
Oropharynx	090-091,098-099,100,102-104,108-109	8000-8713, 8800-9136,9141-9582,9700-9701	
Nasopharynx	110, 112,113,118-119	8000-8713, 8800-9136,9141-9582,9700-9701	
Nasopharynx (continued)	111	8000-8713, 8800-9136,9141-9582,9700-9701	010, 100, 981
PharyngealTonsil	111	8000-8713, 8800-9136,9141-9582,9700-9701	020
Hypopharynx	129,130-132,138-139	8000-8713, 8800-9136,9141-9582,9700-9701	
PharynxOther	140, 142, 148	8000-8713, 8800-9136,9141-9582,9700-9701	
EpiglottisAnterior	101	8000-8713, 8800-9136,9141-9582,9700-9701	
LarynxGlottic	320	8000-8713, 8800-9136,9141-9582,9700-9701	
LarynxSupraglottic	321	8000-8713, 8800-9136,9141-9582,9700-9701	
LarynxSubglottic	322	8000-8713, 8800-9136,9141-9582,9700-9701	
LarynxOther	323,328-329	8000-8713, 8800-9136,9141-9582,9700-9701	
NasalCavity	300	8000-8713, 8800-9136,9141-9582,9700-9701	
SinusMaxillary	310	8000-8713, 8800-9136,9141-9582,9700-9701	
SinusEthmoid	311	8000-8713, 8800-9136,9141-9582,9700-9701	

SinusOther	312-313,318-319	8000-8713, 8800-9136,9141-9582,9700-9701	
ParotidGland	079	8000-9136,9141-9582,9700-9701	
SubmandibularGland	080	8000-9136,9141-9582,9700-9701	
SalivaryGlandOther	081,088-089	8000-9136,9141-9582,9700-9701	
Thyroid	739	8000-9136,9141-9582,9700-9701	
MelanomaLipUpper	000,003	8720-8790	
MelanomaLipLower	001, 004, 006	8720-8790	
MelanomaLipOther	002,005,008-009	8720-8790	
MelanomaTongueAnterior	020-023, 028-029	8720-8790	
MelanomaGumUpper	030	8720-8790	
MelanomaGumLower	031, 062	8720-8790	
MelanomaGumOther	039	8720-8790	
MelanomaFloorMouth	040-041,048-049	8720-8790	
MelanomaPalateHard	050	8720-8790	
MelanomaMouthOther	058-059, 068-069	8720-8790	
MelanomaBuccalMucosa	060-061	8720-8790	
MelanomaTongueBase	019, 024	8720-8790	
MelanomaPalateSoft	051-052	8720-8790	
MelanomaOropharynx	090-091,098-099,100,102-104,108-109	8720-8790	
MelanomaNasopharynx	110-113,118-119	8720-8790	
MelanomaHypopharynx	129,130-132,138-139	8720-8790	
MelanomaPharynxOther	140, 142, 148	8720-8790	
MelanomaEpiglottisAnterior	101	8720-8790	
MelanomaLarynxGlottic	320	8720-8790	
MelanomaLarynxSupraglottic	321	8720-8790	
MelanomaLarynxSubglottic	322	8720-8790	
MelanomaLarynxOther	323,328-329	8720-8790	
MelanomaNasalCavity	300	8720-8790	
MelanomaSinusMaxillary	310	8720-8790	
MelanomaSinusEthmoid	311	8720-8790	
MelanomaSinusOther	312-313,318-319	8720-8790	
Esophagus	150-155,158-159	8000-8934,8940-9136,9141-9582,9700-9701	
EsophagusGEJunction	160	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	982
EsophagusGEJunction (continued)	161-162	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	020, 040, 060,
Stomach	161-162	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	000, 030, 100, 999
Stomach (continued)	163-166,168-169	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	981
SmallIntestine	170-173,178-179	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	
Appendix	181	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	
CarcinoidAppendix	181	8153,8240-8242,8246,8249	
Colon	180,182-189	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	

Rectum	199, 209	8000-8152,8154-8231,8243-8245,8247,8248,8250-8934,8940-9136,9141-9582,9700-9701	
Anus	210-212,218	8000-9136,9141-9582,9700-9701	
GISTEsophagus	150-155,158-159	8935-8936	
GISTStomach	160-166,168-169	8935-8936	
GISTSmallIntestine	170-173,178-179	8935-8936	
GISTAppendix	181	8935-8936	
GISTColon	180,182-189	8935-8936	
GISTRectum	199, 209	8935-8936	
GISTPeritoneum	480-482,488	8935-8936	
NETStomach	160-166,168-169	8153,8240-8242,8246,8249	
NETSmallIntestine	170-173,178-179	8153,8240-8242,8246,8249	
NETColon	180,182-189	8153,8240-8242,8246,8249	
NETRectum	199, 209	8153,8240-8242,8246,8249	
NETAmpulla	241	8153,8240-8242,8246,8249	
Liver	220	8000-8157,8162-8175,8190-9136,9141-9582,9700-9701	
Liver (continued)	221	8170-8175	
BileDuctsIntraHepat	221	8000-8162, 8180-9136,9141-9582,9700-9701	
BileDuctsIntraHepat (continued)	220	8160,8161,8180	
Gallbladder	239	8000-9136,9141-9582,9700-9701	
CysticDuct	240	8000-9136,9141-9582,9700-9701	030
BileDuctsPerihilar	240	8000-9136,9141-9582,9700-9701	010, 020, 050, 060, 100, 999
BileDuctsDistal	240	8000-9136,9141-9582,9700-9701	040, 070
BiliaryOther	248-249	8000-9136,9141-9582,9700-9701	
AmpullaVater	241	8000-8152,8154-8231,8243-8245,8247,8248,8250-9136,9141-9582,9700-9701	
PancreasHead	250	8000-9136,9141-9582,9700-9701	
PancreasBodyTail	251-252	8000-9136,9141-9582,9700-9701	
PancreasOther	253-254,257-259	8000-9136,9141-9582,9700-9701	
Lung	340-343,348-349	8000-9136,9141-9582,9700-9701	
Pleura	384	8000-9136,9141-9582,9700-9701	
Bone	400-403,408-409,410-414,418-419	8000-9136,9141-9582,9700-9701	
HeartMediastinum	380-383,388	8000-9136,9141-9582,9700-9701	
SoftTissue	470-476,478-479,490-496,498-499	8000-9136,9141-9582,9700-9701	
Peritoneum	481-482,488 & male only	8000-8934,8940-9136,9141-9582,9700-9701	001, 003, 004, 009, 100, 981
Peritoneum (continued)	481-482,488 & female only	8580-8589,8680-8921,9120-9136,9141-9582,9700-9701	981
Retroperitoneum	480	8000-8934,8940-9136,9141-9582,9700-9701	
Skin	440,442-449	8000-8246,8248-8713,8800-9136,9141-9582	
Scrotum	632	8000-8246,8248-8713,8800-9136,9141-9582	
MerkelCellSkin	440,442-449	8247	
MerkelCellVulva	510-512,518-519	8247	
MerkelCellPenis	600-602,608-609	8247	
MerkelCellScrotum	632	8247	
MelanomaSkin	440-449,510-512,518-519,600-602,608-609,632	8720-8790	
Breast	500-506,508-509	8000-9136,9141-9582,9700-9701	
Vulva	510-512,518-519	8000-8246,8248-8713,8800-9136,9141-9582	

Vagina	529	8000-9136,9141-9582,9700-9701	
Cervix	530-531,538-539	8000-9136,9141-9582,9700-9701	
CorpusCarcinoma	540-543,548-549,559	8000-8790,8950,8951,8980-8981,9700-9701	
CorpusSarcoma	540-543,548-549,559	8800-8932,8934-8941,8959-8974,8982-9136,9141-9582	
CorpusAdenosarcoma	540-543,548-549,559	8933	
Ovary	569	8000-9136,9141-9582,9700-9701	
PeritoneumFemaleGen	481-482,488 & female only	8000-8576, 8590-8671, 8930-8934, 8940-9110	002
FallopianTube	570	8000-9136,9141-9582,9700-9701	
Placenta	589	8000-9136,9141-9582,9700-9701	
Penis	600-602,608-609	8000-8246,8248-8713,8800-9136,9141-9582	
Prostate	619	8000-9136,9141-9582,9700-9701	
Testis	620-621,629	8000-9136,9141-9582,9700-9701	
KidneyParenchyma	649	8000-9136,9141-9582,9700-9701	
KidneyRenalPelvis	659, 669	8000-9136,9141-9582,9700-9701	
Bladder	670-679	8000-9136,9141-9582,9700-9701	
Urethra	680	8000-9136,9141-9582,9700-9701	
AdrenalGland	740-741,749	8000-9136,9141-9582,9700-9701	
SkinEyelid	441	8000-8713,8800-9136,9141-9508,9510-9514, 9520-9582	
Conjunctiva	690	8000-8713,8800-9136,9141-9508,9520-9582,9700-9701	
MelanomaConjunctiva	690	8720-8790	
MelanomaChoroid	693	8720-8790	
MelanomaCiliaryBody	694	8720-8790	010, 100
Melanomalris	694	8720-8790	020
Retinoblastoma	690-696,698-699	9510-9514	
LacrimalGland	695	8000-8713,8800-9136,9141-9508,9520-9582,9700-9701	015, 100
LacrimalSac	695	8000-8713,8800-9136,9141-9508,9520-9582,9700-9701	025
Orbit	696	8000-9136,9141-9508,9520-9582,9700-9701	
MelanomaEyeOther	691,692,695,698-699	8720-8790	
EyeOther	691-694,698-699	8000-8713,8800-9136,9141-9508,9520-9582,9700-9701	
LymphomaOcularAdnexa	441, 690, 695-696	9590-9699,9702-9738, 9811-9818,9820-9837	
Brain	700,710-719	8000-9136,9141-9582,9700-9701	
CNSOther	701,709,720-725,728-729	8000-9136,9141-9582,9700-9701	
IntracranialGland	751,752,753	8000-9136,9141-9582,9700-9701	
EndocrineOther	379,750,754-755,758-759	8000-9136,9141-9582,9700-9701	
Lymphoma	000-440,442-689,691-694, 698-809	9590-9699, 9702-9729,9735,9737,9738	
Lymphoma (continued)	000-419,422-423,425-440,442-689,691-694,698-809	9811-9818,9823, 9827,9837	
MycosisFungoides	440-449,510-512,518-519,600-602,608-609,632	9700-9701	
MyelomaPlasmaCellDisorder	000-440, 442-689, 691-694, 698-809	9731,9732,9734	
DigestiveOther	260,268-269	8000-9136,9141-9582,9700-9701	
MiddleEar	301	8000-9136,9141-9582,9700-9701	
Trachea	339	8000-9136,9141-9582,9700-9701	
RespiratoryOther	390,398-399	8000-9136,9141-9582,9700-9701	
AdnexaUterineOther	571-574	8000-9136,9141-9582,9700-9701	
GenitalFemaleOther	577-579	8000-9136,9141-9582,9700-9701	

GenitalMaleOther	630-631, 637-639	8000-9136,9141-9582,9700-9701	
UrinaryOther	681,688-689	8000-9136,9141-9582,9700-9701	
KaposiSarcoma	000-809	9140	
HemeRetic	000-809	9740-9809,9840-9992	
HemeRetic (continued)	420, 421, 424	9811-9818,9823, 9827,9837	
HemeRetic (continued)	000-440, 442-689, 691-694, 698-809	9733,9820,9826,9831-9836	
IIIDefinedOther	420-424,760-765,767-768,770-775,778-779,809	8000-9136,9141-9582,9700-9701	

Table 2: Number of Cases for SEER Site Recode by CS Schema Name (CSv0204), 2010 diagnoses excluding DCO & Autopsy cases, SEER Nov 2012 submission, SEER 18 Areas

SEER Site Recode	Schema Name or Abbreviation (see below)	Number of cases
Oral Cavity and Pharynx	Lip & oral cav	3,743
Oral Cavity and Pharynx	Pharynx	5,169
Oral Cavity and Pharynx	Melanoma Head & Neck	22
Oral Cavity and Pharynx	Major salivary	1,030
Oral Cavity and Pharynx	Larynx	21
Esophagus	Esophagus	3,818
Esophagus	GISTEsophagus	<5
Stomach	Stomach	3,798
Stomach	NETStomach	405
Stomach	GISTStomach	406
Stomach	EsophagusGEJunction	1,856
Small Intestine	SmallIntestine	711
Small Intestine	GISTSmallIntestine	221
Small Intestine	NETSmallIntestine	1,033
Colon excluding Rectum	Colon	25,606
Colon excluding Rectum	NETColon	298
Colon excluding Rectum	GISTColon	15
Colon excluding Rectum	Appendix	588
Colon excluding Rectum	GISTAppendix	<5
Colon excluding Rectum	CarcinoidAppendix	114
Colon excluding Rectum	DigestiveOther	21
Rectum and Rectosigmoid Junction	Rectum	10,173
Rectum and Rectosigmoid Junction	NETRectum	1,090
Rectum and Rectosigmoid Junction	GISTRectum	17
Anus, Anal Canal and Anorectum	Anus	2,441
Liver and Intrahepatic Bile Duct	BileDuctsIntraHepat	862
Liver and Intrahepatic Bile Duct	Liver	6,167
Gallbladder	Gallbladder	1,026
Other Biliary	BileDuctsPerihilar	510
Other Biliary	BileDuctsDistal	338
Other Biliary	BiliaryOther	214
Other Biliary	AmpullaVater	506
Other Biliary	NETAmpulla	24
Other Biliary	CysticDuct	17
Pancreas	Pancreas – total	10,415
Retroperitoneum	GISTPeritoneum	6
Retroperitoneum	Retroperitoneum	322
Peritoneum, Omentum and Mesentery	Peritoneum	40
Peritoneum, Omentum and Mesentery	GISTPeritoneum	11
Peritoneum, Omentum and Mesentery	PeritoneumFemaleGen	429
Other Digestive Organs	DigestiveOther	503
Other Digestive Organs	Peritoneum	<5
Other Digestive Organs	PeritoneumFemaleGen	<5
Nose, Nasal Cavity and Middle Ear	Melanoma Head & Neck	55
Nose, Nasal Cavity and Middle Ear	Nasal Cav & Sinus	511
Nose, Nasal Cavity and Middle Ear	MiddleEar	14
Larynx	Larynx	3,184
Lung and Bronchus	Lung	49,020
Pleura	Pleura	30
Trachea, Mediastinum and Other Respiratory Organs	Trachea	56
Trachea, Mediastinum and Other Respiratory Organs	RespiratoryOther	<5
Trachea, Mediastinum and Other Respiratory Organs	HeartMediastinum	116
Bones and Joints	Bone	758
Soft Tissue including Heart	SoftTissue	2,844
Soft Tissue including Heart	HeartMediastinum	26
Melanoma of the Skin	MelanomaSkin	32,913
Other Non-Epithelial Skin	Skin	1,021

Other Non-Epithelial Skin	MerkelCellSkin	473
Other Non-Epithelial Skin	SkinEyelid	85
Breast	Breast	72,622
Cervix Uteri	Cervix	3,325
Corpus and Uterus, NOS	CorpusCarcinoma	11,787
Corpus and Uterus, NOS	CorpusSarcoma	477
Corpus and Uterus, NOS	CorpusAdenosarcoma	60
Ovary	Ovary	5,704
Vagina	Vagina	702
Vulva	MelanomaSkin	81
Vulva	Vulva	2,967
Other Female Genital Organs	FallopianTube	348
Other Female Genital Organs	Placenta	32
Other Female Genital Organs	AdnexaUterineOther	9
Other Female Genital Organs	GenitalFemaleOther	135
Prostate	Prostate	56,519
Testis	Testis	2,380
Penis	MelanomaSkin	7
Penis	Penis	508
Other Male Genital Organs	MelanomaSkin	<5
Other Male Genital Organs	Scrotum	98
Other Male Genital Organs	GenitalMaleOther	34
Urinary Bladder	Bladder	17,385
Kidney and Renal Pelvis	KidneyParenchyma	12,423
Kidney and Renal Pelvis	KidneyRenalPelvis	1,012
Ureter	KidneyRenalPelvis	673
Other Urinary Organs	Urethra	206
Other Urinary Organs	UrinaryOther	149
Eye and Orbit	MelanomaChoroid	354
Eye and Orbit	MelanomaCiliaryBody	33
Eye and Orbit	MelanomaConjunctiva	55
Eye and Orbit	Melanomalris	15
Eye and Orbit	MelanomaEyeOther	34
Eye and Orbit	Conjunctiva	141
Eye and Orbit	LacrimalGland	15
Eye and Orbit	LacrimalSac	12
Eye and Orbit	Orbit	49
Eye and Orbit	Retinoblastoma	78
Eye and Orbit	EyeOther	42
Brain and Other Nervous System	Brain	11,547
Brain and Other Nervous System	CNSOther	2,961
Thyroid	Thyroid	11,571
Other Endocrine including Thymus	AdrenalGland	255
Other Endocrine including Thymus	EndocrineOther	305
Other Endocrine including Thymus	IntracranialGland	3,300
Hodgkin Lymphoma	Lymphoma	2,285
Non-Hodgkin Lymphoma	LymphomaOcularAdnexa	199
Non-Hodgkin Lymphoma	MycosisFungoides	425
Non-Hodgkin Lymphoma	Lymphoma	16,325
Non-Hodgkin Lymphoma	IIIDefinedOther	<5
Myeloma	LymphomaOcularAdnexa	<5
Myeloma	MyelomaPlasmaCellDisorder	5,254
Leukemia	HemeRetic	11,002
Mesothelioma	Any schema with a mesothelioma	<5
Kaposi Sarcoma	KaposiSarcoma	434
Miscellaneous	HemeRetic	6,646
Miscellaneous	IIIDefinedOther	6,912

Abbreviations used in table:

SEER site recode of Oral Cavity and Pharynx includes Lip, Oral Cavity, and Pharynx

CS Lip & oral cav = BuccalMucosa, FloorMouth, GumLower, GumOther, GumUpper, LipLower, LipOther, LipUpper,

MouthOther, PalateHard, TongueAnterior

Pharynx = Hypopharynx, Nasopharynx, Oropharynx, PalateSoft, PharyngealTonsil, PharynxOther, TongueBase

Melanoma Head & Neck = MelanomaBuccalMucosa,
MelanomaEpiglottisAnterior, MelanomaFloorMouth,
MelanomaGumLower, MelanomaGumOther,
MelanomaGumUpper, MelanomaHypopharynx,
MelanomaLarynxGlottic, MelanomaLarynxSubglottic,
MelanomaLarynxSupraglottic, MelanomaLarynxOther,
MelanomaLipLower, MelanomaLipOther,
MelanomaLipUpper, MelanomaMouthOther,
MelanomaNasalCavity, MelanomaNasopharynx,
MelanomaLarynxOther, MelanomaLarynxSubglottic,
MelanomaLipLower, MelanomaLipOther,
MelanomaLipUpper, MelanomaMouthOther,
MelanomaNasalCavity,
MelanomaNasopharynx, MelanomaOropharynx,
MelanomaPalateHard, MelanomaPalateSoft,
MelanomaPharynxOther, MelanomaSinusEthmoid,
MelanomaSinusMaxillary, MelanomaSinusOther,
MelanomaTongueAnterior, MelanomaTongueBase

Major salivary = ParotidGland, SalivaryGlandOther, SubmandibularGland

Pancreas - total = PancreasBodyTail, PancreasHead, PancreasOther

Nasal Cav & Sinus = NasalCavity, SinusEthmoid, SinusMaxillary, SinusOther

Larynx = EpiglottisAnterior, LarynxGlottic, LarynxOther, LarynxSubglottic, LarynxSupraglottic

Table 3: Number of Cases for CS Schema by SEER Site Recode, SEER Nov 2012 submission, 2010 diagnosis excluding DCO & autopsy cases, SEER 18 Areas

CS Schema V0204 Name	SEER site recode	Number
Lip & oral cav	Oral Cavity and Pharynx	3,743
Pharynx	Oral Cavity and Pharynx	5,169
Melanoma Head & Neck	Oral Cavity and Pharynx	22
Melanoma Head & Neck	Nose, Nasal Cavity and Middle Ear	55
Major salivary	Oral Cavity and Pharynx	1,030
Stomach	Stomach	3,798
Stomach	Mesothelioma	<5
NETStomach	Stomach	405
GISTStomach	Stomach	406
Esophagus	Esophagus	3,818
GISTEsophagus	Esophagus	<5
EsophagusGEJunction	Stomach	1,856
SmallIntestine	Small Intestine	711
GISTSmallIntestine	Small Intestine	221
NETSmallIntestine	Small Intestine	1,033
Colon	Colon excluding Rectum	25,606
NETColon	Colon excluding Rectum	298
GISTColon	Colon excluding Rectum	15
Appendix	Colon excluding Rectum	588
GISTAppendix	Colon excluding Rectum	<5
CarcinoidAppendix	Colon excluding Rectum	114
Rectum	Rectum and Rectosigmoid Junction	10,173
NETRectum	Rectum and Rectosigmoid Junction	1,090
GISTRectum	Rectum and Rectosigmoid Junction	17
Anus	Anus, Anal Canal and Anorectum	2,441
Gallbladder	Gallbladder	1,026
BileDuctsIntraHepat	Liver and Intrahepatic Bile Duct	862
BileDuctsPerihilar	Other Biliary	510
BileDuctsDistal	Other Biliary	338
Liver	Liver and Intrahepatic Bile Duct	6,167
Pancreas - total	Pancreas	10,415
BiliaryOther	Other Biliary	214
AmpullaVater	Other Biliary	506
NETAmpulla	Other Biliary	24
DigestiveOther	Colon excluding Rectum	21
DigestiveOther	Other Digestive Organs	503
Nasal Cav & Sinus	Nose, Nasal Cavity and Middle Ear	511
Lung	Lung and Bronchus	49,020
Lung	Mesothelioma	28
Larynx	Oral Cavity and Pharynx	21
Larynx	Larynx	3,184
Trachea	Trachea, Mediastinum and Other Respiratory Organs	56
Pleura	Pleura	30
Pleura	Mesothelioma	679
MiddleEar	Nose, Nasal Cavity and Middle Ear	14
RespiratoryOther	Trachea, Mediastinum and Other Respiratory Organs	<5
MelanomaSkin	Melanoma of the Skin	32,913
MelanomaSkin	Vulva	81
MelanomaSkin	Penis	7
MelanomaSkin	Other Male Genital Organs	<5
Skin	Other Non-Epithelial Skin	1,021
MerkelCellSkin	Other Non-Epithelial Skin	473
Bone	Bones and Joints	758
SoftTissue	Soft Tissue including Heart	2,844
SoftTissue	Mesothelioma	<5
HeartMediastinum	Trachea, Mediastinum and Other Respiratory Organs	116
HeartMediastinum	Soft Tissue including Heart	26

HeartMediastinum	Mesothelioma	5
Peritoneum	Peritoneum, Omentum and Mesentery	40
Peritoneum	Other Digestive Organs	<5
Peritoneum	Mesothelioma	34
GISTPeritoneum	Retroperitoneum	6
GISTPeritoneum	Peritoneum, Omentum and Mesentery	11
Retroperitoneum	Retroperitoneum	322
Retroperitoneum	Mesothelioma	3
PeritoneumFemaleGen	Peritoneum, Omentum and Mesentery	429
PeritoneumFemaleGen	Other Digestive Organs	<5
PeritoneumFemaleGen	Mesothelioma	28
Breast	Breast	72,622
CorpusCarcinoma	Corpus and Uterus, NOS	11,787
CorpusSarcoma	Corpus and Uterus, NOS	477
CorpusAdenosarcoma	Corpus and Uterus, NOS	60
Ovary	Ovary	5,704
Cervix	Cervix Uteri	3,325
Vulva	Vulva	2,967
Vagina	Vagina	702
FallopianTube	Other Female Genital Organs	348
Placenta	Other Female Genital Organs	32
AdnexaUterineOther	Other Female Genital Organs	9
AdnexaUterineOther	Mesothelioma	<5
GenitalFemaleOther	Other Female Genital Organs	135
GenitalFemaleOther	Mesothelioma	<5
Prostate	Prostate	56,519
Testis	Testis	2,380
Scrotum	Other Male Genital Organs	98
Penis	Penis	508
GenitalMaleOther	Other Male Genital Organs	34
GenitalMaleOther	Mesothelioma	<5
Bladder	Urinary Bladder	17,385
KidneyParenchyma	Kidney and Renal Pelvis	12,423
KidneyRenalPelvis	Kidney and Renal Pelvis	1,012
KidneyRenalPelvis	Ureter	673
AdrenalGland	Other Endocrine including Thymus	255
Urethra	Other Urinary Organs	206
UrinaryOther	Other Urinary Organs	149
Thyroid	Thyroid	11,571
EndocrineOther	Other Endocrine including Thymus	305
Brain	Brain and Other Nervous System	11,547
CNSOther	Brain and Other Nervous System	2,961
IntracranialGland	Other Endocrine including Thymus	3,300
KaposiSarcoma	Kaposi Sarcoma	434
MelanomaChoroid	Eye and Orbit	354
MelanomaCiliaryBody	Eye and Orbit	33
MelanomaConjunctiva	Eye and Orbit	55
MelanomaIris	Eye and Orbit	15
MelanomaEyeOther	Eye and Orbit	34
Conjunctiva	Eye and Orbit	141
CysticDuct	Other Biliary	17
LacrimalGland	Eye and Orbit	15
SkinEyelid	Other Non-Epithelial Skin	85
LacrimalSac	Eye and Orbit	12
Orbit	Eye and Orbit	49
Retinoblastoma	Eye and Orbit	78
EyeOther	Eye and Orbit	42
LymphomaOcularAdnexa	Non-Hodgkin Lymphoma	199
LymphomaOcularAdnexa	Myeloma	<5
MycosisFungoides	Non-Hodgkin Lymphoma	425
Lymphoma	Hodgkin Lymphoma	2,285
Lymphoma	Non-Hodgkin Lymphoma	16,325

MyelomaPlasmaCellDisorder	Myeloma	5,254
HemeRetic	Leukemia	11,002
HemeRetic	Miscellaneous	6,646
IIIDefinedOther	Non-Hodgkin Lymphoma	<5
IIIDefinedOther	Mesothelioma	7
IIIDefinedOther	Miscellaneous	6,912

Same footnotes as Table 2

Table 4: Derived AJCC T: Display and Storage/Transmission Codes

This table shows the allowable values for this Collaborative Stage data item. The Storage Code is the value to be stored in the field of a NAACCR record. The Storage Codes are designed for analysis. The Display String is the corresponding label that should be displayed on the screen or in a report. The meaning of these strings will be clear to the registrar or physician user.

Note: A blank may indicate either that the calculation was not performed or that the calculation resulted in an error.

Display String	AJCC6 Storage Code	AJCC7 Storage Code	Comments
TX	99	999	TX
T0	00	000	T0
Ta	01	010	Ta
Tis	05	050	Tis
Tispu	06	060	Tispu (Urethra only)
Tispd	07	070	Tispd (Urethra only)
T1	10	100	T1
T1mi	11	110	T1mi
T1NOS	19	199	T1 NOS
T1NOS (s)		191	T1 NOS(s)
T1NOS (m)		192	T1 NOS(m)
T1a	12	120	T1a
T1a(s)		121	T1a(s)
T1a(m)		122	T1a(m)
T1a1	13	130	T1a1
T1a2	14	140	T1a2
T1b	15	150	T1b
T1b(s)		151	T1b(s)
T1b(m)		152	T1b(m)
T1b1	16	160	T1b1
T1b2	17	170	T1b2
T1c	18	180	T1c
T1d		181	T1d
T2	20	200	T2
T2(s)		201	T2(s)
T2(m)		202	T2(m)
T2NOS	29	299	T2 NOS
T2a	21	210	T2a
T2a1		211	T2a1
T2a2		212	T2a2
T2aNOS		213	T2a NOS
T2b	22	220	T2b
T2c	23	230	T2c
T2d		240	T2d
T3	30	300	T3
T3(s)		301	T3(s)
T3(m)		302	T3(m)
T3NOS	39	399	T3 NOS
T3a	31	310	T3a
T3b	32	320	T3b
T3c	33	330	T3c

T3d		340	T3d
T4	40	400	T4
T4NOS	49	499	T4 NOS
T4NOS (s)		491	T4 NOS(s)
T4NOS (m)		492	T4 NOS(m)
T4a	41	410	T4a
T4a(s)		411	T4a(s)
T4a(m)		412	T4a(m)
T4b	42	420	T4b
T4b(s)		421	T4b(s)
T4b(m)		422	T4b(m)
T4c	43	430	T4c
T4d	44	440	T4d
T4e		450	T4e
T1aNOS	80	800	T1a NOS
T1bNOS	81	810	T1b NOS
NA	88	888	Not applicable
			For errors: Display string is a blank of length 1. AJCC6 storage code is a blank of length 2. AJCC7 storage code is a blank of length 3.

Table 5: Derived AJCC N: Display and Storage/Transmission Codes

This table shows the allowable values for this Collaborative Stage data item. The Storage Code is the value to be stored in the field of a NAACCR record. The Storage Codes are designed for analysis. The Display String is the corresponding label that should be displayed on the screen or in a report. The meaning of these strings will be clear to the registrar or physician user.

Note: A blank may indicate either that the calculation was not performed or that the calculation resulted in an error.

Code N0NOS was defined in CSv1 but never used. It has been removed in CSv2.

Display String	AJCC6 Storage Code	AJCC7 Storage Code	Comments
NX	99	999	NX
N0	00	000	N0
N0(i-)	01	010	N0(i-)
N0(i+)	02	020	N0(i+)
N0(mol-)	03	030	N0(mol-)
N0(mol+)	04	040	N0(mol+)
N1	10	100	N1
N1NOS	19	199	N1 NOS
N1a	11	110	N1a
N1b	12	120	N1b
N1c	13	130	N1c
N1mi	18	180	N1mi
N2	20	200	N2
N2NOS	29	299	N2 NOS
N2a	21	210	N2a
N2b	22	220	N2b
N2c	23	230	N2c
N3	30	300	N3
N3NOS	39	399	N3 NOS
N3a	31	310	N3a
N3b	32	320	N3b
N3c	33	330	N3c
N4		400	N4
NA	88	888	Not applicable
			For errors: Display string is a blank of length 1. AJCC6 storage code is a blank of length 2. AJCC7 storage code is a blank of length 3.

Table 6: Derived AJCC M: Display and Storage/Transmission Codes

This table shows the allowable values for this Collaborative Stage data item. The Storage Code is the value to be stored in the field of a NAACCR record. The Storage Codes are designed for analysis. The Display String is the corresponding label that should be displayed on the screen or in a report. The meaning of these strings will be clear to the registrar or physician user.

Note: A blank may indicate either that the calculation was not performed or that the calculation resulted in an error.

Display String	AJCC6 Storage Code	AJCC7 Storage Code	Comments
MX	99	999	MX
M0	00	000	M0
M0(i+)		010	M0(i+)
M1	10	100	M1
M1a	11	110	M1a
M1b	12	120	M1b
M1c	13	130	M1c
M1d		140	M1d
M1e		150	M1e
M1NOS	19	199	M1 NOS
NA	88	888	Not applicable
			For errors: Display string is a blank of length 1. AJCC6 storage code is a blank of length 2. AJCC7 storage code is a blank of length 3.

Table 7: Derived AJCC Stage: Display Strings and Storage/Transmission Codes

This table shows the allowable values for this Collaborative Stage data item. The Storage Code is the value to be stored in the field of a NAACCR record. The Storage Codes are designed for analysis. The Display String is the corresponding label that should be displayed on the screen or in a report. The meaning of these strings will be clear to the registrar or physician user.

Note: A blank may indicate either that the calculation was not performed or that the calculation resulted in an error.

Display String	AJCC6 Storage Code	AJCC7 Storage Code	Comments
0	00	000	Stage 0
0a	01	010	Stage 0a
0is	02	020	Stage 0is
I	10	100	Stage I
INOS	11	110	Stage I NOS
IA	12	120	Stage IA
IA1	13	130	Stage IA1
IA2	14	140	Stage IA2
IANOS		121	Stage IA NOS
IB	15	150	Stage IB
IB1	16	160	Stage IB1
IB2	17	170	Stage IB2
IBNOS		151	Stage IB NOS
IC	18	180	Stage IC
IS	19	190	Stage IS
ISA	23	230	Stage ISA (lymphoma only)
ISB	24	240	Stage ISB (lymphoma only)
IEA	20	200	Stage IEA (lymphoma only)
IEB	21	210	Stage IEB (lymphoma only)
IE	22	220	Stage IE (lymphoma only)
II	30	300	Stage II
IINOS	31	310	Stage II NOS
IIA	32	320	Stage IIA
IIANOS		321	Stage IIA NOS
IIA1		322	Stage IIA1
IIA2		323	Stage IIA2
IIB	33	330	Stage IIB
IIC	34	340	Stage IIC
IIEA	35	350	Stage IIEA (lymphoma only)
IIEB	36	360	Stage IIEB (lymphoma only)
IIE	37	370	Stage IIE (lymphoma only)
IISA	38	380	Stage IISA (lymphoma only)
IISB	39	390	Stage IISB (lymphoma only)
IIS	40	400	Stage IIS (lymphoma only)
IIESA	41	410	Stage IIESA (lymphoma only)
IIESB	42	420	Stage IIESB (lymphoma only)
IIES	43	430	Stage IIES (lymphoma only)
III	50	500	Stage III
IIINOS	51	510	Stage III NOS
IIIA	52	520	Stage IIIA

IIIB	53	530	Stage IIIB
IIIC	54	540	Stage IIIC
IIIC1		541	Stage IIIC1
IIIC2		542	Stage IIIC2
IIIEA	55	550	Stage IIIEA (lymphoma only)
IIIEB	56	560	Stage IIIEB (lymphoma only)
IIIE	57	570	Stage IIIE (lymphoma only)
IIISA	58	580	Stage IIISA (lymphoma only)
IIISB	59	590	Stage IIISB (lymphoma only)
IIIS	60	600	Stage IIIS (lymphoma only)
IIIESA	61	610	Stage IIIESA (lymphoma only)
IIIESB	62	620	Stage IIIESB (lymphoma only)
IIIES	63	630	Stage IIIES (lymphoma only)
IV	70	700	Stage IV
IVNOS	71	710	Stage IV NOS
IVA	72	720	Stage IVA
IVA1		721	Stage IVA1
IVA2		722	Stage IVA2
IVB	73	730	Stage IVB
IVC	74	740	Stage IVC
NA	88	888	Not applicable
OCCULT	90	900	Stage Occult
UNK	99	999	Stage Unknown
			For errors: Display string is a blank of length 1. AJCC6 storage code is a blank of length 2. AJCC7 storage code is a blank of length 3.

Table 8: Derived AJCC Descriptor: Display and Storage/Transmission Codes

This table shows the allowable values for the derived Collaborative Stage data items: Derived AJCC T Descriptor, Derived AJCC N Descriptor, and Derived AJCC M Descriptor. The Storage Code is the value to be stored in the field of a NAACCR record. The Storage Codes are designed for analysis. The Display String is the corresponding label that should be displayed on the screen or in a report.

Note 1: A blank of length 1 may indicate either that the calculation was not performed or that the calculation resulted in an error.

Note 2: These descriptors are considered to be modifiers of the derived T, N, and M codes. If the T, N, or M code is absent, there will be nothing to modify and the descriptor should not be displayed or stored.

Note 3: In CS Version 2, display strings can have different lengths. For example, the display string for 'y' is 'yp'. The display string for 'c' is 'c'. The display string for 'N' is "", a blank of length 0.

Display String	Storage Code	Comments
c	c	
p	p	
a	a	
yp	y	
	N	Display String is blank of length 0.
		For errors: Display string is blank of length 1. Storage code is blank of length 1.