

**Collaborative Stage Data Collection System  
Version 2: 0204**

**Implementation Guide for  
Registries and Vendors**

**Elaine N. Collins  
Minnesota Cancer Surveillance System**

**December 5, 2011**

## Implementation Issues Team

KIMBERLEY BOYUK, MA  
Statistics Canada  
Ottawa, ON

CONNIE BURA  
American Joint Committee on Cancer  
Chicago, IL

SUSAN CAPRON  
Chair, NAACCR Edits Work Group  
Chicago, IL

ELAINE COLLINS, MA, RHIA, CTR  
Minnesota Cancer Surveillance System  
St. Paul, MN

APRIL FRITZ, RHIT, CTR  
A.Fritz and Associates  
Reno, NV

BARRY GORDON, PhD  
C/NET Solutions  
Berkeley, CA

DON GREEN  
Information Management Services, Inc  
Silver spring, MD

DONNA GRESS, RHIT, CTR  
American Joint Committee on Cancer  
Chicago, IL

PETER KIM  
Centers for Disease Control and Prevention  
Atlanta, GA

GARY M. LEVIN, CTR  
Florida Cancer Data System  
Miami, FL

KAREN POLLITT  
American Joint Committee on Cancer  
Chicago, IL

MARTIN MADERA  
American Joint Committee on Cancer  
Chicago, IL

CHUCK MAY  
Information Management Services, Inc  
Silver Spring, MD

MARY MROSZCZYK, CTR  
Massachusetts Cancer Registry  
Boston, MA

TOM RAWSON  
Centers for Disease Control and Prevention  
Atlanta, GA

LYNN RIES, MS  
National Cancer Institute  
Bethesda, MD

JOSEPH ROGERS  
Centers for Disease Control and Prevention  
Atlanta, GA

DAVID RONEY  
Information Management Services, Inc.  
Silver Spring, MD

JENNIFER RUHL, RHIT, CCS, CTR  
National Cancer Institute  
Bethesda, MD

JENNIFER SEIFFERT, MLIS, CTR  
Northrop Grumman Health Solutions  
Warsaw, IN

REDA WILSON, CTR  
Centers for Disease Control and Prevention  
Atlanta, GA

## Table of Contents

1	INTRODUCTION .....	5
1.1	V0204.....	5
1.2	Use of This Implementation Guide .....	6
2	V0204 CHANGES.....	6
2.1	Changes in Content.....	6
2.2	Changes in Process .....	7
2.3	Notable Schema-Specific Changes in V0204 .....	7
2.4	OBSOLETE Codes.....	7
2.5	CS Versioning .....	8
3	REVIEW SPECIFICATIONS .....	12
3.1	OBSOLETE DATA REVIEWED AND CHANGED V0204 .....	12
3.2	OBSOLETE DATA RETAINED AND REVIEWED V0204 and OBSOLETE DATA RETAINED V0204.....	12
3.3	Special Review Situations .....	12
4	SOFTWARE IMPLEMENTATION .....	13
4.1	Source Files .....	13
4.2	Application Interface (API).....	13
4.3	Derivation of Stage Values .....	14
4.4	Version Stamping .....	14
5	DATA CONVERSION FROM V0203 to V0204.....	14
5.1	Conversion Program .....	14
5.2	Conversion Specifications .....	15
5.3	Old Cases, New Version .....	15
5.4	New Cases, Old Version .....	15
6	NPCR Requirements for SEER Summary Stage 2000 .....	15
7	DOCUMENTATION .....	16
7.1	NAACCR Data Dictionary.....	16
7.2	CSv2 User Documentation.....	16
7.3	CSv2 Software .....	17
7.4	CAnswerForum .....	17
8	EDITS .....	17
8.1	NAACCR Version 12.2 Edits Metafile .....	17
8.2	Edits Processing.....	18
8.4	Edits at Conversion .....	18
9	CENTRAL REGISTRY CONSIDERATIONS .....	18
9.1	Data Preparation for Existing Records.....	19
9.2	Data Conversion .....	20
9.3	Data Review After Conversion.....	20
9.4	Submission Testing .....	20
9.5	Edits.....	21
9.6	Quality Monitoring .....	21

10	VENDOR/SOFTWARE DEVELOPER CONSIDERATIONS .....	21
10.1	Timing of CS Version Stamping.....	22
10.2	Timing of Data Entry for 2010 Cases.....	22
10.3	Edits.....	22
10.4	Submission Format Testing.....	22
10.5	Data Conversion Instructions and Review .....	23
10.6	Vendor Information.....	23
11	FACILITY REGISTRY CONSIDERATIONS .....	24
11.1	Release Notes.....	24
11.2	Timing of Data Abstracting and Reporting .....	24
11.3	Data Preparation .....	24
11.4	Data Review.....	24
11.5	Data Submission .....	25
11.6	Questions.....	25

# 1 INTRODUCTION

## 1.1 V0204

Collaborative Stage Data Collection System Version 2, production version 020440, is required for use with cancers diagnosed 1/1/2012 and later and may be used for earlier cases. Data are assumed to be in the North American Association of Central Cancer Registries, Inc. (NAACCR) layout 12.2. The NAACCR 12.2 Edits Metafile should be used after upgrade to the NAACCR 12.2 layout and CS data conversion. Release items include:

- Release notes for v02.04
- CS Manual, Parts I and Part II
  - PDF format
  - Online Help format
- Collaborative Stage Tables
  - XML format
  - HTML format
- Collaborative Stage Library (DLL)
  - API Documentation
  - Source code
  - API sample programs
  - XML & HTML tables used in the build
  - Systems documentation
- Implementation Information
  - Implementation Guide for v02.04
  - Data Conversion Specifications for v02.04, overview and rules (MS Word document)
  - Data Conversion Spreadsheet (MS Excel document)
  - Conversion Utility (NorthCon)
- EDITS metafile
- General Information and Testing Files
  - Test-O-Matic
  - Valid, Invalid, and Obsolete Codes tables
  - Default and Unknown Values for CS items for DCOs
  - Test cases with valid NAACCR records
  - Test cases with invalid NAACCR records

All files and documents for this release are available through the CS Web page at

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

The V0204 files completely replace all earlier CS files. Only data converted into V0204 and collected in V0204 should be maintained after implementation and testing.

Earlier releases of Collaborative Stage Version 2 include:

- 02.00.01, January 2010, for use with 2010 diagnoses
- 02.01.00, February 2010, for use with 2010 diagnoses

- 02.02.00, April 2010, for use with 2010 diagnoses
- 02.03.02, December 2010, for use with 2011 diagnoses

## **1.2 Use of This Implementation Guide**

This implementation guide and the release notes are supplements to the other documents noted above, the software documentation, the conversion specifications, and the user manuals. This document contains some duplication of these other materials as necessary to provide a guide for implementation of V0204. In all cases, the other documentation should be considered the primary authority.

For registry personnel, this guide provides a background for the changes in V0204 and information about case preparation and review. For software developers, in-house or outside vendors, this guide is designed to point out particular steps and considerations in upgrading software to V0204 and assisting registries with the review and upgrade process.

## **2 V0204 CHANGES**

### **2.1 Changes in Content**

CSv2 represented a large increase from CSv1 in the number of data items to be defined and collected in the site-specific factor fields to meet AJCC 7 staging and other data requirements, and many code adjustments in the fields carried over from CSv1 to satisfy backward compatibility for AJCC 6 stage and Summary Stage 1977 and Summary Stage 2000 values. Further extensive changes were made to the CS system in V0203 as a result of a data validation review after the release of V0202. Given the complexity of the system, the time constraints of release schedules, support resources available, and the time delays between release and widespread use in the registry community, ongoing identification of issues and/or problems was expected. The release of V0204 was authorized to correct typographical errors, to resolve problems affecting stage output, and to address other issues as time permitted.

The code changes in V0204 are much smaller in number and review requirements are fewer than were the changes for V0203. New codes are listed in the conversion documentation spreadsheets: 111 codes in 22 schemas. Obsolete codes are also listed in the conversion spreadsheets: 25 codes in 15 schemas. The number of extra tables has increased; these tables, combine codes from two or more data fields to produce a staging component value. Regional Nodes Positive has become a required input data item to derive AJCC stage values for the Merkel Cell schemas. Corrections have been made to derived stage values returned by the CS algorithm.

There are no major changes in the design and function of the CS algorithm in CSv0204. The structure of the tables remains the same. There are no new schemas or newly defined site-specific factors.

Note that 20 CS Edit Over-ride fields are available in the NAACCR record and may be used in edits on CS fields starting with the 12.2 Edits metafile. Over-ride CS 20 has been designated as a

flag for directly coded Summary Stage 2000 [NAACCR data item 759] to support National Program of Cancer Registry (NPCR) requirements; see section 6.

## 2.2 Changes in Process

A database system has been developed to document issues, track response to issues by the Mapping Team, document the status of issues, collect release notes, and generate various types of reports for the Project Management Team. Issues in the database not resolved in V0204 have been marked for later consideration.

The CAnswerForum, maintained by the American Joint Committee on Cancer (AJCC) and the American College of Surgeons Commission on Cancer, (ACoS-COC) has been designated as the primary channel of communication between the registry community and the CS teams regarding questions about CS coding and potential problems with CS codes and output values. Questions in CAnswerForum are addressed by Cancer Technical Advisory Panel (CTAP) volunteers, who identify and forward issues to the database system as required.

## 2.3 Notable Schema-Specific Changes in V0204

- Lung: Subheadings of extra tables reformatted to display more clearly on schema index page
- MelanomaSkin, MerkelCellPenis, MerkelCellScrotum, MerkelCellSkin, MerkelCellVulva: Collection of information in SSF3, Clinical Status of Lymph Nodes, and AJCC staging tables modified to distinguish between and report both cN and pN values. Secondary nodal basins defined as regional nodes.
- Breast: Codes expanded in CS SSF 19, Assessment of Positive Axillary Lymph Nodes, to include more combinations of findings at biopsy and resection.
- CorpusCarcinoma: Histology codes 8950 and 8951 reassigned to this schema from CorpusSarcoma
- Prostate: Upper limit of specified number of biopsy needle cores examined and positive extended to 100.
- Thyroid: (s) and (m) descriptors added to T values for AJCC 7 staging

## 2.4 OBSOLETE Codes

The new data tags used in this release are:

- OBSOLETE DATA CONVERTED V0204 – Code converted according to V0204 conversion specifications. Use of this code generates an ERROR for the AJCC 6 and 7 and SS77 and SS 2000 mappings after conversion, and the code should not be available for use after conversion to V0204.

- CONVERTED AND CODE REUSED V0204 - Code converted according to V0204 conversion specifications. All cases with the code are converted to a new value. The code is given a new definition and remains available for use with the new definition after conversion to V0204. Care must be taken on data entry to assign the code based on the V0204 definition rather than the prior definition.
- OBSOLETE DATA REVIEWED AND CHANGED V0204 – Code MUST be reviewed and changed according to V0204 review specifications. Use of this code generates a CS ERROR for the AJCC 6 and 7 and SS77 and SS2000 mappings, and the code should not be available for use in V0204.
- OBSOLETE DATA RETAINED AND REVIEWED V0204 – Code SHOULD be reviewed and changed according to V0204 review specifications. The review is optional but strongly recommended for cases originally coded in CSv2 (CS Version Original 020001, 020100, 020200, 020302) OR cases with year of diagnosis 2010+. This code will map according to CS table specifications, but fails an edit on OBSOLETE codes for new cases (CS Version Input Original of 020440) if edit included in metafile used by the registry. The code should not be used for new cases.
- OBSOLETE DATA RETAINED V0204 – Code MAY be reviewed and changed according to V0204 review specifications, according to registry requirements for data use and analysis. If undertaken, review could be limited to cases originally coded in CSv2 (CS Version Original 020001, 020100, 020200, 020302) OR cases with year of diagnosis 2010+. This code will map according to CS table specifications, but fails an edit on OBSOLETE codes for new cases (CS Version Input Original of 020440) if edit included in metafile used by the registry. The code should not be used for new cases.

OBSOLETE tables, codes, and code descriptions should always be available to the user for inspection and updating in cases coded in previous CS versions, and the coded information available for data queries. It is recommended that the OBSOLETE tables and codes not be presented for routine data entry into cases abstracted in CSv2. The CS Dynamic Link Library (DLL) and the EDITS program contain strictures on the use of OBSOLETE codes in records based on diagnosis date and CS Version codes.

## 2.5 CS Versioning

CS Versioning is designed to track the history of CS coding for a particular record, and is used by data analysts in evaluating the quality and meaning of reported codes:

- CS Version Input Original: CS production version in use when the record was first abstracted
- CS Version Input Current: CS version in use when the most recent updates were made to the record. Different numbers are assigned to distinguish among records with different CS



versions at original data entry, converted records without later updates, converted records with later updates, and records originally abstracted in the current version.

- CS Version Derived: CS version in use when the most recent stage calculation was applied.

The CS Version Input Original value for the current installed version of CSv2 should be entered on all new cases abstracted in CSv2. The same value should be entered into the CS Version Input Current value at initial data entry; the two values will remain the same until coding is updated to a later version of CSv2 and CS Version Input Current is also updated. A specific CS Version Input Current value is assigned to every converted record at data conversion, based on the existing CS Version Input Original and CS Version Input Current values in the record; the CS Version Input Current value assigned at conversion is not a value that is used in the CS Version Input Original or CS Version Derived fields. (For details about these version numbers, see the table at the end of section 2.5.) The CS Version Input Current value must be updated after conversion if the converted record is reviewed and coding is updated to the current installed version on any of the CS data entry fields.

The versioning system used in CS has limitations. The versioning codes are applied to a whole record, rather than to each code in the record, and abstractors may update some but not all CS codes in a record, so that the record may represent a hybrid of codes from different versions. Depending on vendor implementation, an abstractor may only review a record without updating any codes, and cause the CS Version Input Current value to be updated. If the CS Version Input Current codes have not been correctly assigned at any conversion point, codes to be assigned at later conversions will not accurately reflect the conversion and review history of individual records and codes.

The conversion specifications for V0204 provide detailed instructions for assigning CS Version Input Current to converted records. These instructions should be carefully followed. CS Version Input Current should NOT be automatically assigned to the same value as CS Version Derived when the CS calculation is performed. CS Version Input Current and CS Version Derived carry different meanings.

The CS Version number for the production release of V0204 is 020440. This version number should be used in:

- CS Version Input Original: all cases originally coded in V0204.
- CS Version Input Current: all cases with any codes manually updated after conversion to V0204, using V0204 codes.
- CS Version Derived: all cases with stage derived using the V0204 algorithm.

CS Version Input Current numbers 020410, 020412, 020413, 020420, 020423, and 020430 are entered into cases converted from V0203 and only assigned at conversion:

- CS Version Input Current 020410:  
Cases converted from CSv1 to CSv2 and not manually updated (CS Version Input Current 020000 after first conversion), subsequently converted from V0202 to V0203 and not manually updated (CS Version Input Current 020300 after second conversion), and now converted to V0204 without manual update.
- CS Version Input Current 020412:  
Cases converted from CSv1 to CSv2 and manually updated (CS Version Input Current 020001, 020100, or 020200 after first conversion), subsequently converted from V0202 to V0203 and not manually updated (CS Version Input Current 020301 after second conversion), and now converted to V0204 without manual update.
- CS Version Input Current 020413:  
Cases converted from CSv1 to CSv2 and not updated ((CS Version Input Current 020000) or manually updated (CS Version Input Current 020001, 020100, or 020200) after first conversion), subsequently converted from V0202 to V0203 and manually updated (CS Version Input Current 020302 after second conversion and update), and now converted to V0204 without manual update.
- CS Version Input Current 020420:  
Cases originally coded in 02000001/020100/020200 (CS Version Input Current 020001/020100/020200 on initial data entry), converted to V0203 without manual update (CS Version Input Current 020301 after conversion), and now converted to V0204 without manual update.
- CS Version Input Current 020423:  
Cases originally coded in 02000001/020100/020200 (CS Version Input Current 020001/020100/020200 on initial data entry), converted to V0203 with manual update (CS Version Input Current 020302 after conversion and update), and now converted to V0204 without manual update.
- CS Version Input Current 020430:  
Cases originally coded in V0203 and now converted to V0204 without manual update

CS Version Input Current values of 020410, 020412, 020413, 020420, 020423, and 020430 should not be kept or entered into any record that undergoes manual review after conversion.

The following table summarizes the assignment of CS Version Input Current values for records converted to V0204.

CS Version Original	CS Version Input Current					
	Converted to CS 020000 without update	Converted to CS 020000 with update	Converted to CS 020300 without update	Converted to CS 020300 with update	Converted to CS 020400 without update	Converted to CS 020400 with update
0009XX 01XXXX	020000		020300		020410	
0009XX 01XXXX	020000		020300			020440
0009XX 01XXXX	020000			020302	020413	
0009XX 01XXXX	020000			020302		020440
0009XX 01XXXX		020200	020301		020412	
0009XX 01XXXX		020200	020301			020440
0009XX 01XXXX		020200		020302	020413	
0009XX 01XXXX		020200		020302		020440
020001 020100 020200			020301		020420	
020001 020100 020200			020301			020440
020001 020100 020200				020302	020423	
020001 020100 020202				020302		020440
020302					020430	
020302						020440

## **3 REVIEW SPECIFICATIONS**

### **3.1 OBSOLETE DATA REVIEWED AND CHANGED V0204**

Case review requirements for V0204 are noted on the conversion specification spreadsheets. All cases including codes with an OBSOLETE DATA REVIEWED AND CHANGED V0204 tag MUST

be manually reviewed to complete the conversion from V0203 to V0204. (See discussion in section 3.3 below for review and recoding which must be completed BEFORE the conversion process.) Review notes are included for each code, indicating reason for review and suggested replacement codes. As noted in the Conversion document, software providers should provide each user with a list of cases meeting review criteria, so the registrar can complete the review of these cases manually. The user should also be provided with copies of the spreadsheets for reference during case review.

The reviewer should always verify the accuracy of the original codes, and consider other codes than those suggested if appropriate for the case information. Note that in many cases two data fields must be recoded if the change in code involves a data description that has been moved to a different data item; for example, certain nodes have been reclassified from regional to distant nodes. Note also that in many cases a combination code may be the appropriate choice, and the reviewer must consider the meaning of other codes involved in the combination that may not be tagged for review.

### **3.2 OBSOLETE DATA RETAINED AND REVIEWED V0204 and OBSOLETE DATA RETAINED V0204**

Review is STRONGLY RECOMMENDED for all cases with the tag OBSOLETE DATA RETAINED AND REVIEWED V0204. Review is SUGGESTED for all cases with the tag OBSOLETE DATA RETAINED V0204, depending on registry resources and needs for data use and analysis. Cases with these codes will continue to derive stage values, but the codes should not be used for new cases originally coded in V0204, nor for cases with a diagnosis date of 2012 and later. Review notes are included for each code, indicating reason for review and suggested replacement codes.

### **3.3 Special Review Situations**

Three special review situations are described on the conversion spreadsheets.

- EsophagusGEJunction/Stomach, Site-Specific Factor 25, Schema Discriminator:

Review and recoding of cases with site code C161-C166, C168, C169 and SSF 25 code 010, assigned to either the EsophagusGEJunction or Stomach schema, must be carried out BEFORE the case is sent through the conversion program. This code is tagged OBSOLETE DATA REVIEWED AND CHANGED V0204. Primary site code C160 should have been used for all cases with SSF 25 code of 010, and the conversion specifications are to convert code 010 to code 982. However, if code 010 appears in a case with a different primary site code, and the primary site code is not changed before conversion, after conversion to code 982 the record may no longer be assigned to a schema. The

conversion documents contain information about identifying and recoding these cases to avoid the problem of no schema assignment, but this must be done in advance of the conversion process. If a case is reassigned to a different schema as a result of recoding the primary site code, the case needs to be reabstracted using correct codes in the new schema.

Listing of these EsophagusGEJunction/Stomach cases will not be part of the conversion utility program provided by CDC/NPCR (see section 5.1). The list will need to be prepared locally based on the specifications in the spreadsheet.

- Scrotum and Skin, Site-Specific Factor 12, High Risk Features:

The definition of high risk features was corrected in V0203, with lymph-vascular invasion removed from the list and the value of depth of invasion changed from  $\geq 4$  mm to  $>2$ mm. The highest number if coded (5 for Scrotum, 6 for Skin) was changed to one less (4 for Scrotum, 5 for Skin) in the conversion to V0203, but review was not specified for cases with less than the highest number of features coded. Review is now specified for cases with a CS Version Input Original code less than V020302, to account for cases that were originally coded in an earlier version of CS and those that may have been updated from an earlier version of CS.

- CorpusCarcinoma, Histology 8950 and 8951

Histologies 8950 and 8951 have been reassigned to the CorpusCarcinoma schema from the CorpusSarcoma schema. The codes and code definitions are not the same for these two schemas for the data items CS Extension, CS Lymph Nodes, and CS Site-Specific Factor 1, FIGO Stage. The conversion documents specify that these data items, for all cases with CS coding with an ICD-O-3 histology code of 8950 or 8951 reassigned to the Corpus Carcinoma schema in V0204, must be listed for review.

## 4 SOFTWARE IMPLEMENTATION

### 4.1 Source Files

All files required for adding V0204 to the registry system are available through the CS Web page <http://www.cancerstaging.org/cstage/software/index.html>, including documentation of the application interface, a Windows–compiled library, XML and HTML tables, and ready-to-compile source code for other operating systems.

### 4.2 Application Interface (API)

The API documentation for the API (see *CS API Version 2.doc*) describes the tools available to use the XML tables as a resource, reducing the need for external tables to support the library.

### 4.3 Derivation of Stage Values

The algorithm will return derived stage values to the calling program via the datacard. The returned stage values include both AJCC 6 and AJCC 7 stage values as well as SS1977 and SS2000 values. The calling program is responsible for copying these values from the CS datacard into the patient record.

If the algorithm is reapplied to a case, but the rederivation of one or more of the stage values is not possible, the old values in those derived stage fields, including the corresponding derived stage flags, should be replaced with blanks.

### 4.4 Version Stamping

The CSv2 library contains only the latest version number and returns it to the calling program via the API; it is up to the calling program to determine when and how the version number should be stamped into each of these three fields.

- CS Version Input Original – this field should be filled by the calling program the first time the CS tables are accessed for data entry.
- CS Version Input Current – this field should be filled by the calling program whenever a case has been updated with newer codes from a CS version newer than that used to stamp the CS Version Input Original field. For new cases, CS Version Input Current should be stamped with the current version from the CStage.DLL. For converted cases, CS Version Input Current is filled by the conversion program, and is only changed when the case is updated or a new conversion is applied.
- CS Version Derived – this field should be filled by the calling program whenever the algorithm can successfully calculate a derived stage value (SEER Summary Stage or AJCC Derived Stage Group) from the current values in the input fields.

If a case abstracted under a previous version of CS has an empty field for CS Version Input Original, the algorithm has no way to know under which version it was abstracted; it knows only the current version. Such cases should be corrected before updating to V0204 in order to have the correct CS Version recorded for the initial coding system.

See section 6 for information on version stamping for NPCR requirements.

## 5 DATA CONVERSION FROM V0203 TO V0204

### 5.1 Conversion Program

A conversion program to convert V0203 to V0204 records is available from CDC/NPCR. Details of that program and how it can be implemented are available directly from CDC. A link to the program and documentation on the CDC/NPCR Web page are posted on the CS Web site, <http://cancerstaging.org/cstage/software/index.html>. The conversion program from CDC/NPCR will not identify the EsophagusGEJunction/Stomach cases which may need recoding before conversion.

## 5.2 Conversion Specifications

The complete conversion specifications are in two documents posted on the CS Web site. The specifications are those that will be used in the CDC/NPCR conversion program mentioned in section 5.1. TNM components, TNM descriptors, and stage value for AJCC 6 and AJCC 7, and Summary Stage 1977 and Summary Stage 2000 values must be rederived on all cases as part of the conversion process. TNM components, TNM descriptors, and stage value for AJCC 6 and AJCC 7, and Summary Stage 1977 and Summary Stage 2000 values must be rederived again on all cases which require manual review and recoding as identified during the conversion process.

## 5.3 Old Cases, New Version

Pre-2012 cases coded in V0203 do not need to be recoded with the new V0204 code values, other than records with codes returned for review from the conversion program or records failing Edits after conversion.

## 5.4 New Cases, Old Version

All cases diagnosed on or after 1/1/2012 should be coded according to the V0204 data entry codes. Because the production release of V0204 is scheduled for December 5, 2011 following a beta test period, it is anticipated that registries will have V0204 available at or soon after that date, in time for abstracting 2012 cases. However, for any registry that begins coding 2012 cases before receiving the CS upgrade, all 2012 cases coded under CS V0203 must be reviewed and recoded when V0204 becomes available.

## 6 NPCR REQUIREMENTS for SEER Summary Stage 2000

For diagnosis years 2012 and later, NPCR permits the use of SEER Summary Stage 2000 [NAACCR data item 759] in those cases where collection of CSv2 data items is not feasible due to lack of data or staffing and time constraints at the local or central registry. Over-ride CS 20 has been designated as a special-purpose flag to identify cases where SEER Summary Stage 2000 [NAACCR data item 759] is directly coded and reported in lieu of Derived SS2000 [NAACCR data item 3020], in accordance with NPCR reporting requirements. The Over-ride CS 20 value of "1", set by the user, identifies a record with NAACCR data item 759 used to report Summary Stage 2000 as permitted by NPCR requirements only; Over-ride CS 20 is left blank for all other cases.

NPCR requires the collection of specified CS data items on all records, even those flagged by Over-ride CS 20: CS Tumor Size, Regional Nodes Examined, Regional Nodes Positive, CS SSF 25, CS Version Input Original, CS Version Input Current, and Site-Specific Factors for certain schemas as identified in NPCR published standards. Other CS input fields may remain blank and derived fields must remain blank in these flagged records. With the implementation of CS V0203, blanks were not permitted by any standard setter in CS data entry fields. With the implementation of CS V0204, blanks are allowed by NPCR only (except as noted above) in records where the Over-ride CS 20 flag is set to "1".

Note: When both SEER Summary Stage 2000 [NAACCR data item 759] and Derived SS2000 [NAACCR data item 3020], are being reported, the Over-ride CS 20 field should be left blank. The Over-ride CS 20 flag is not required to be set to “1” for all records with SEER Summary Stage 2000 coded. The flag should be set only where the directly coded SEER Summary Stage 2000 is reported in lieu of the CS coded and derived fields.

Stamping of records with CS Version Input Original and CS Version Input Current must be based on the subset of CS data input fields required by NPCR for flagged cases. The return of derived stage values and error messages from the CS algorithm should be suppressed where the Over-ride CS 20 flag is set and the CS input fields (except as noted above) are blank.

SEER and COC standards do not permit the reporting of SEER Summary Stage 2000 [NAACCR data item 759] in lieu of Derived SS2000 [NAACCR data item 3020]. Registries meeting these standards may not use Over-ride CS 20 as a data flag for this purpose.

## **7 DOCUMENTATION**

### **7.1 NAACCR Data Dictionary**

The NAACCR Standards Volume II Version 12.2 defines the CS data fields. The Data Dictionary provides the NAACCR item number, position within the NAACCR version 12.2 report layouts, and requirements for collection as established by the standard setters, SEER, COC, NPCR, and CCCR. This volume is available on the NAACCR website at [http://www.naacccr.org/index.asp?Col\\_SectionKey=7&Col\\_ContentID=133](http://www.naacccr.org/index.asp?Col_SectionKey=7&Col_ContentID=133).

### **7.2 CSv2 User Documentation**

The CSv2 User Documentation is presented in two formats, PDF and HTML online help, both available for downloading from <http://www.cancerstaging.org/cstage/manuals/index.html>. Both formats include two parts.

- Collaborative Stage Data Collection System Coding Instructions: Part I Section 1: General Instructions, contains a description of the system and coding instructions for each CS data element. Part I Section 2: Lab Tests and Tumor Markers, Site-Specific Factor Notes, contains schema-specific notes, illustrations, and descriptions of the data collected in the new CSv2 site-specific factors. Both documents have been updated to reflect any changes in codes or coding instructions based on changes in V0204.
- Collaborative Stage Data Collection System Part II contains the tables used for coding each schema, with all notes included: above-table notes guiding abstracting and below-table notes identifying extra tables used by the algorithm to derive T, N, M, or stage values. The tables document the mapping to T, N, M and stage group for both AJCC 6 and AJCC 7 staging and for both SS1977 and SS2000. Part II may also be generated from the Cstage.DLL using API functions. The V0204 changes to codes, coding instructions, and stage groupings are contained in the HTML tables.



Software vendors may include the HTML online help version of the CS user documentation within their registry software releases.

The CSv2 User Documentation is intended to be used in electronic form, accessible on the abstractor's computer. A printed version of the PDF manuals or replacement pages may be available for purchase from the National Cancer Registrars Association. Contact NCRA for details.

### 7.3 CSv2 Software

The CSv2 software is published on the AJCC website for downloading by software vendors and central registries, at <http://www.cancerstaging.org/cstage/software/index.html>. The software contains:

- DLLs
  - cstage.dll -- For the build/version number, either right-click on the file in Windows Explorer and check the version properties tab or call the API function CStage\_get\_version()
  - Wrapper dlls used for VB.Net and Java
- Include files: csapi.h and collab.h
- API sample programs
  - Test-o-Matic
  - VB.NET, C, Java, and C#
- XML & HTML tables used in the build

The .DLL files contain the source code for the CS algorithm, composed of routines to read the CS XML tables and derive stage values. The XML tables are the fundamental repository of the CS system, containing table notes, codes, descriptions, mappings to the derived values, and specifications for data processing. Test-o-Matic is a downloadable interactive program useful for entering individual case information and reviewing CS output, both for testing the performance of the algorithm and the accuracy of data coding.

### 7.4 CAnswer Forum

The CAnswer Forum, maintained by the AJCC and COC, <http://cancerbulletin.facs.org/forums/>, is the official resource for questions about the use of CSv2. The CAnswer Forum contains a forum for Collaborative Stage questions, which are answered by the CTAP, volunteer registrars with CS expertise and access to AJCC curators when needed. The forums support public comment or dialogue on answers posted by CTAP members.

## 8 EDITS

### 8.1 NAACCR Version 12.2 Edits Metafile

The NAACCR Version 12.2 Edits Metafile is scheduled for release concurrently with the release of V0204. All current CS edits have been updated to accommodate changes with new and obsolete codes. CS Edit Override fields are available in the NAACCR record and may be used in new edits in the 12.2 Edits Metafile. Over-ride CS 20 has been designated as a special-purpose flag for

cases reporting SEER Summary Stage 2000 [NAACCR data item 759] as permitted by NPCR, as described in section 6. Edits enforce the correct use of the flag for NPCR-required data items. CS input data elements will be edited if they are included in a record with coded SEER Summary Stage 2000 [NAACCR data item 759] and Over-ride CS 20 set to “1”; the derived stage components will fail edits if they are not returned as blank.

The CS Edits Workgroup continues to review CS codes and instructions to identify coding situations where edits can be used to assist registrar abstracting. As noted in previous implementations, given the increase in the amount of data collected in CSv2, the expanded table notes instructing coding, the complexity of data relationships across the tables, and time required for coders to use and identify problems, it is anticipated that the identification and construction of CSv2 edits will extend well beyond the initial implementation of the system. The 12.2 Edits Metafile contains new edits that apply to cases collected under previous versions of CSv2, and all edits that could be applied to V0204 cases have not yet been written.

## **8.2 Edits Processing**

The CS Mapping Team recommends that, in an interactive setting, Edits be run on case information after input of CS data items but before the CS algorithm is called to derive stage output values. The Edits program generates more specific error messages than the CS algorithm; processing edits before CS calculation assists the abstractor in identifying coding problems that may not be readily identifiable from a manual review of CS input codes.

## **8.3 Edits at Conversion**

All available CS edits should be run both before and after data conversion. The NAACCR 12.1 Edits Metafile should be used before conversion, and the NAACCR 12.2 Edits Metafile should be used after conversion. All edit errors should be reviewed and edit failures corrected.

All codes marked with an OBSOLETE tag in the CSv2 XML/HTML tables will fail with a CS Version Input Original value equal to or later than the version number specified in the OBSOLETE data tag.

All cases with a diagnosis date 1/1/2012 and later will fail with a CS Version Input Current value less than 020440. This edit, “CS Version Input Current, Date of DX (CS)”, is useful in identifying any 2012 cases that may have been entered in V0203 and require updating to V0204 codes.

Note that ALL cases with codes marked OBSOLETE DATA REVIEWED AND CHANGED V0204 will fail standard edits and will not derive a stage value until the records are reviewed and coding updated.

## **9 CENTRAL REGISTRY CONSIDERATIONS**

Central registries should review the conversion specifications and release notes, including information about new, obsolete, and review codes, to determine if they affect any of their reporting requirements. They should communicate any change in requirements to reporting facilities and software vendors supplying reporting facilities.

Central registries should determine if record review and updating is conducted in-house, or if reporting facilities will be requested to submit the updated records. Central registries should make certain that information about V0204 coding changes is conveyed to data quality staff and reporting facility staff, to assure a common understanding and application of V0204 codes.

Central registries should also communicate to reporting facilities and vendors when they will be able to accept V0204 data, the date after which they will no longer accept V0203 data, and their expectations for submission of review cases.

Central registries must be cognizant of the CS Version Input Current value of the records in their databases and records being submitted from reporting facilities, as the timing of software upgrades and database conversions may vary among facilities and between facilities and the central registry. Each CSv2 algorithm is designed to run on data either originally coded in or converted to that CS version, and CS-related edits are also version-specific. A central registry will encounter significant problems in trying to consolidate data across records coded in CS V0202 and CS V0203 or CS V0203 and CS V0204 for example, or in applying the CS V0204 algorithm to a mix of CS V0203 and CS V0204 cases, or in using the NAACCR 12.1 metafile (with edits for CS V0203) on records submitted in the NAACCR 12.2 format (with edits for CS V0204). If a central registry receives V0204 records before conversion of its database, it should consider segregating the CS fields from routine data processing. If a central registry, after conversion of its database to V0204, continues to receive V0203 records from a reporting facility, those records should be converted to V0204 before being entered into the central database.

Central registries should determine whether and how they will implement NPCR requirements for reporting SEER Summary Stage 2000 [NAACCR data item 759] in lieu of Derived SS2000 [NAACCR data item 3020], and inform reporting facilities not subject to SEER and/or COC requirements, and the providers of software used by these facilities. Central registries must modify software developed in-house, or work with their outside providers, to assure that CS version stamping is correctly applied to flagged records and that derived stage components and CS error messages because of incomplete CS data are suppressed. Central registries may be required to modify data processing procedures, to separate flagged and non-flagged records, if the CS algorithm is routinely run on all records to rederive stage values. See section 6 for further information.

## **9.1 Data Preparation for Existing Records**

To ensure correct conversion and derivation of CS data, all records in the registry database with staging information diagnosed from 1/1/2004 through the date of implementation must be processed through the V0204 algorithm; AJCC 6 stage and SS1977 and SS2000 will be rederived on all cases where sufficient information is available. It is strongly recommended that registries reedit all cases using the most recent 12.1 Edits Metafile before implementation of V0204. As noted, edit development is a continuous process, and all CSv2 cases in the registry database may not have passed through all available CS edits developed for V0202 and V0203. Fewer conversion problems are anticipated with edited data. It is strongly recommended that registries identify any

cases which fail stage derivation before conversion and correct problems, so that complete, intact cases are sent through the conversion process.

Registries may elect to rederive AJCC 6 and/or SS1977 and SS2000 on their cases after they have been edited using the V0203 algorithm, as a second step to ensure that cases are correctly coded before data conversion. Registries should also identify and correct coding on all cases which contain a blank in any of the CS Version fields to ensure correct processing by the conversion program.

As described in section 3.3 above, central registries must identify and recode all cases with 010 in SSF 15 for EsophagusGEJunction and Stomach schemas where the site code is not C160, BEFORE running the conversion program. If a case is assigned to another schema by this recoding, the entire case will have to be reviewed for correct coding in the new schema.

## **9.2 Data Conversion**

Central registries must run the data conversion program on cases diagnosed from 2004 through the latest received records in the registry database. If the registry continues to receive records with a CS Version Input Current value less than 020400 after conversion, the registry must run the conversion program on these records before incorporating them into the registry database.

## **9.3 Data Review After Conversion**

If record review will be carried out in-house, registries should develop a plan for manual review and change of cases with data items identified in the conversion specifications as OBSOLETE DATA REVIEWED AND CHANGED V0204, OBSOLETE DATA RETAINED AND REVIEWED V0204, and OBSOLETE DATA RETAINED V0204. The volume of these cases is small for the V0204 conversion.

Except for cases with code 010 in SSF 25 for EsophagusGEJunction/Stomach, the cases with the tag OBSOLETE DATA REVIEWED AND CHANGED V0204 should result in an ERROR value upon first processing by the V0204 algorithm. They require manual review, and then reprocessing to derive stage values. The automated conversion program will produce a list of cases requiring manual review and conversion. The CS Version Input Current field should be set to 020440 when cases are recoded using the V0204 tables.

After the data review has been completed and the cases updated, all CS derived fields should be calculated via the algorithm for all cases with CS information. All cases with failure to derive any component after conversion should be reviewed and coding updated. Derived components include AJCC 7 T, N, and M; T, N, and M descriptors; and stage group for cases diagnosed 1/1/2010 and later; AJCC 6 T, N, and M; T, N, and M descriptors; and stage group for cases diagnosed 1/1/2004 and later; and Summary Stage 1977 and Summary Stage 2000 for cases diagnosed 1/1/2004 and later.

## **9.4 Submission Testing**

Central registries should consider requesting test files from all vendors for review and approval of placement of data fields and correctness of data conversions before accepting routine submissions

of data converted to V0204. Careful inspection of initial submissions can identify problems in the vendor's translation of codes from registry software into the NAACCR layout, and also problems in the central registry's processing programs moving data from the NAACCR layout into the registry database. Reviewing the first data submissions which contain converted data may identify systematic problems in either the facility's or the central registry's conversion of V0203 coding to V0204, which can be rectified at an early stage when the registries are still focused on the conversion process.

## **9.5 Edits**

Edits updated to V0204 are available in the NAACCR 12.2 metafile. The central registry should install and use the NAACCR 12.2 Edit Metafile after conversion to V0204. The central registry should carry out any desired customization of the NAACCR 12.2 Edit Metafile as soon as practicable, and re-edit all converted cases using the NAACCR 12.2 Metafile. As noted in Section 7.1, the development and review process for CS edits will continue into the future, and central registries should develop a plan that facilitates upgrades of edit metafiles as they are posted by the NAACCR Edits Committee.

Early release of a customized metafile to vendors will facilitate implementation of those edits within the reporting registries.

## **9.6 Quality Monitoring**

In addition to aggressively supporting all training opportunities for central registry and reporting registry staff members, central registries should develop a program of monitoring data quality focused on the changes from V0203 to V0204. Subjects might include the three special reviews noted for EsophagusGEJunction, Skin and Scrotum, and CorpusCarcinoma. As they perform quality monitoring, central registries should be aware of data relationships that could be subject to edits, registrar problems or misconceptions that could be amenable to training, and system faults that could be rectified. Central registries should communicate any such findings to the general registry community and in particular to the CSv2 Project Management Team.

## **10 VENDOR/SOFTWARE DEVELOPER CONSIDERATIONS**

Vendors were faced with a number of presentation issues in the conversion from CSv1 to CSv2. One issue that continues to be reported is the handling of OBSOLETE codes on data entry. The CS Mapping Team has strongly recommended that OBSOLETE codes be presented only on request; that they NOT be presented on routine abstracting. This recommendation is reiterated.

Given the importance of the information contained in table notes to guide accurate coding, the recommendation to include all notes and code descriptions in the tables for coders is also reiterated.

Vendors that supply software used by central registries, or by facilities required to meet NPCR reporting requirements as directed by central registries (but not SEER or COC requirements), must consult with the central registries to determine if they will accept the reporting of SEER Summary

Stage 2000 [NAACCR data item 759] in lieu of Derived SS2000 [NAACCR data item 3020]. Where applicable, software must be modified to support the coding of Over-ride CS 20 with “1” as a flag to return values in CS Version Input Original and CS Version Input Current, but to suppress the output of CS derived stage components and CS error messages on blank input fields. See section 6 for further information.

### **10.1 Timing of CS Version Stamping**

The conversion documents specify the CS Version Input Current values that are to be assigned to converted records.

After conversion, the return of CS Version Input Original and CS Version Input Current must be tied to data entry rather than stage derivation. See the extensive discussion of versioning in Section 2.5, and further considerations for version stamping in sections 4.4 and 6.

CS Version Derived should be stamped upon successful completion of derived Summary and/or AJCC stage. For quality purposes, the software should require derivation of stage upon completion of the case, so that algorithm output can be reviewed by the registrar and input values modified as necessary while the registrar is working with the source materials.

### **10.2 Timing of Data Entry for 2012 Cases**

It is strongly recommended that registries not begin abstracting 2012 cases until after they have received their software upgrades with the V0204 algorithm installed. The vendor can assist those registries that begin abstracting 2012 cases before the V0204 upgrade, by listing in the reports from the conversion process all cases with a diagnosis date 1/1/2012 and later and a CS Version Input Original value less than 020440.

### **10.3 Edits**

If possible, the NAACCR 12.2 Edit Metafile should be included with the software release to the registries, along with instructions for processing converted cases with the metafile. As noted in Section 7.1, the development and review process for CS edits will continue into the future, and vendors should develop a plan that facilitates upgrades of edit metafiles as they are posted by the NAACCR Edits Committee. Vendors should make sure that their software data collection and reporting programs accommodate the 20 CS Edit Override fields.

### **10.4 Submission Format Testing**

It is recommended that vendors submit test files to the central registries of their clients before software is released to registries, to ensure that submitted data are formatted properly and conversion has been undertaken correctly. Once a vendor has been certified, it is likely that all client registries of that vendor will be compliant with the new reporting rules.

## 10.5 Data Conversion Instructions and Review

### Before Conversion:

It is recommended that vendors provide instructions for data preparation of all cases, including identification and correction of any records with a blank in any CS Version data field, including derived item fields, and instructions for processing all cases with the latest version of the NAACCR 12.1 Edits Metafile. Vendors should provide assistance to registries in identifying EsophagusGEJunction and Stomach cases with primary site code other than C160 and SSF25 code 010, so that these cases can be appropriately recoded before going through the conversion process. (See Section 3.3 for details.)

### After Conversion:

It is recommended that vendors make available to client registries a list of all cases which fail the conversion process because a schema was not identified.

It is recommended that a report be made available to the registry, after data conversion and processing by the CS algorithm, of cases which have failed to derive any CS derived value. Derived components include AJCC 7 T, N, and M; T, N, and M descriptors; and stage group for cases diagnosed 1/1/2010 and later; AJCC 6 T, N, and M; T, N, and M descriptors; and stage group for cases diagnosed 1/1/2004 and later; and Summary Stage 1977 and Summary Stage 2000 for cases diagnosed 1/1/2004 and later.

It is recommended that vendors make available to client registries a list of all cases which meet the review specifications as detailed in the conversion spreadsheets. In conjunction with this list, it is recommended that vendors make available the spreadsheets with information and instructions on reviewing cases, with specific information for each code linked to each case.

The registry should review and recode all cases described above, which fail schema selection and/or stage derivation and which require review, using V0204 codes, and then manually rerun the algorithm to derive an AJCC 6 and 7 stage and SS1977 and SS2000 values. The CS Version Input Current value should be updated to 020440 upon manual update of converted records, indicating that the case has been recoded using valid codes in V0204.

It is recommended that vendors make available to client registries a list of any 2012 cases that may have been abstracted in V0203 for recoding. The registry software should provide an easy method for registries to mark such cases for resubmission to the central registry after they have been recoded and stage rederived.

## 10.6 Vendor Information

Vendor information is available on the CS website at <http://www.cancerstaging.org/cstage/software/index.html>. Vendor questions or general questions regarding the CS release can be submitted to [CSv2@facs.org](mailto:CSv2@facs.org).

## **11 FACILITY REGISTRY CONSIDERATIONS**

### **11.1 Release Notes**

The registry should carefully review all release notes and changes in the CS Coding Instructions, Part I, Sections 1 and 2, to become familiar with coding changes and to determine if facility coding policies must be updated. Before abstracting a case in each schema, each registrar should carefully review all table notes and codes to make sure that abstracting is based on current definitions and instructions rather than recall of previous information.

### **11.2 Timing of Data Abstracting and Reporting**

The facility registry should communicate its plans with the central registry for reporting cases abstracted in V0204. Both partners should be aware of the time constraints under which each operates, when the facility registry will be able to and desires to submit 2012 cases, and when the central registry will be able to receive and process cases in the NAACCR 12.2 layout with V0204 data included. The registry should develop an understanding with the central registry about responsibility for review, recoding, and reporting of cases converted to V0204.

### **11.3 Data Preparation**

Facility registries should undertake the same data preparation steps as the central registries. With vendor support if required, all cases with stage coded and derived in V0203 should be re-edited with the most recent version of the NAACCR 12.1 metafile. All cases failing stage derivation before conversion should be identified and problems corrected so that complete, intact cases are sent through the conversion and review process. If the registry has started abstracting 2012 cases before the software upgrade is available, the registry should develop its own plan to make sure that these cases are identified, recoded if necessary after data conversion, and resubmitted to the central registry.

As described in section 3.3 above, central registries must identify and recode all cases with 010 in SSF 15 for EsophagusGEJunction and Stomach schemas where the site code is not C160, BEFORE running the conversion program. If a case is assigned to another schema by this recoding, the entire case will have to be reviewed for correct coding in the new schema.

### **11.4 Data Review**

After data conversion and processing of converted cases by the CS algorithm, the registry should receive a report of cases identified for review, cases which failed processing because the schema was not identified, and cases not listed for review which failed derivation of any CS derived value. The registry must review these cases, recode the obsolete codes using current V0204 codes, and rerun the algorithm to derive stage values. The conversion specifications identify the situations where manual review is required. The registry should process all converted records with the NAACCR 12.2 Edits Metafile to identify any possible problems with the data conversion.

After data conversion, the registry should also review the CS Version Input Current values to determine if the software program appears to be assigning these values correctly, based on understanding of past data coding practices within the registry. The registry should investigate any



apparent discrepancies in CS Version assignment with the software vendor before submitting data to the central registry.

### **11.5 Data Submission**

The facility registry should recode in V0204 and resubmit to the central registry any 2012 cases coded before software upgrade to V0204, any converted cases that have not already been submitted to the central registry, and any new cases entered in V0204. The facility should submit reviewed and updated cases to the central registry if so requested. The facility registry should not submit its entire database of converted cases from 2004 forward unless specifically requested to do so by the central registry. If the facility registry discovers a systematic problem in its data conversion that may affect the validity of previously reported cases, or require a correction to the converted data within the central registry, the facility registry should communicate such problems to the central registry and participate in joint planning for resolution. The facility registry should identify the first data submission to the central registry using the updated software program, to alert the central registry to closely review the submission for correctness in data placement and coding. The facility registry should identify the first data submission to the central registry that contains data from converted cases as well as new cases, to alert the central registry to closely review the submission for the accuracy of the conversion process.

### **11.6 Questions**

All questions and inquiries about how to use the CSv2 system or changes in V0204 should be submitted to the CAnswer Forum at <http://cancerbulletin.facs.org/forums/>.