

A thick dark blue vertical bar runs down the left side of the page. A blue arrow-shaped graphic points to the right from the bar, containing the date.

1/31/2015

Data Item Consolidation Manual

Several thin, light blue curved lines originate from the bottom left corner and sweep upwards and to the right, creating a sense of movement or growth.

NAACCR Data Item
Consolidation Work Group

Contents

| | |
|--|-------------------------------------|
| DATA ITEM CONSOLIDATION WORK GROUP MEMBERS | 5 |
| INTRODUCTION | 7 |
| Background | 7 |
| Scope of Work | 7 |
| Operational Issues | 9 |
| Ensuring Data Quality with Automation..... | 9 |
| Work Flow Processing | 10 |
| Weighting Source Records | 10 |
| Class of Case for Central Registries | 10 |
| Recommended Next Steps | 13 |
| Demographic Data Items | 15 |
| Considerations..... | 15 |
| Confidential items | 15 |
| Critical Items..... | 15 |
| Collected but not Transmitted Items | 15 |
| Demographic Data Items..... | 17 |
| Race 1 [160] | 17 |
| Race 2 [161], Race 3 [162], Race 4 [163], Race 5 [164] | 19 |
| Spanish Origin [190]..... | 21 |
| Sex [220] | 23 |
| Date of Birth [240] | 24 |
| Date of Birth Flag [241] | 25 |
| Birthplace [250] | Error! Bookmark not defined. |
| Birthplace - State [252] | 26 |
| Birthplace - Country [254]..... | 27 |
| Last Name [2230] | 28 |
| First Name [2240] | 29 |
| Middle Name [2250] | 30 |
| Suffix [2270], Alias [2280], and Maiden Name [2390]..... | 31 |

| | |
|---|----|
| Occupation | 32 |
| Industry | 34 |
| Current No & Street [2350]..... | 36 |
| Current City [1810] | 37 |
| Current State [1820] | 38 |
| Current Postal Code [100]..... | 39 |
| Current Country [1832]..... | 40 |
| Address Current Supplemental [2355] | 41 |
| Current Telephone [2360] | 42 |
| Social Security Number [2320] | 43 |
| Vital Status [1760]..... | 44 |
| Place of Death [1940]..... | 45 |
| Place of Death - State [1942] | 46 |
| Place of Death - Country [1944]..... | 47 |
| DC State File Number [2380] | 48 |
| Date of Death Canada [1755] | 49 |
| Date of Death - Canada Flag [1756] | 50 |
| Tumor Data Items..... | 51 |
| Address at Dx: No & Street [60]..... | 52 |
| Address at Dx: City [#70]..... | 53 |
| Address at Dx State [80] | 54 |
| County at DX [90]..... | 55 |
| Address at DX Postal Code [100] | 55 |
| Address at DX Country [102]..... | 57 |
| Address at DX Supplemental [2335] | 58 |
| Date of Diagnosis [390]..... | 58 |
| Primary Site [400] | 60 |
| Laterality [410]..... | 61 |
| Grade [440] | 67 |
| Diagnostic Confirmation [490]..... | 68 |
| Type of Reporting Source [500] | 69 |
| Histologic Type ICD-O-3 [522]..... | 70 |

| | |
|--|----|
| Behavior Code ICD-O-3 [523] | 71 |
| Stage Data Items | 72 |
| Treatment | 78 |
| Surgery Primary Site (2003+) [1290]..... | 79 |
| Scope Regional Lymph Node Surgery (2003+) [1292] | 81 |
| Surgery Other Regional/Distant Sites (2003+) [1294] | 82 |
| Rad- Regional RX Modality [1570] | 83 |
| Chemotherapy [1390]..... | 85 |
| Hormone Therapy [1400] | 86 |
| Immunotherapy [1410]..... | 87 |
| Hematologic Transplant and Endocrine Procedures [3250]..... | 88 |
| Other Treatment [1420] | 89 |
| APPENDIX: List of Rule Definitions for Data Item Consolidation | 90 |

DATA ITEM CONSOLIDATION WORK GROUP MEMBERS

Nicole Aagaard
Rocky Mountain Cancer Data Systems
nicole.aagaard@utah.edu

Kim Best
Rocky Mountain Cancer Data Systems
kim.best@utah.edu

Lois Dickie
National Cancer Institute
dickielo@mail.nih.gov

Michelle Esterly
Pennsylvania Cancer Registry
mesterly@pa.gov

Theresa Hinman
New York State Cancer Registry
tmh01@health.state.ny.us

Mei-chin Hsieh
Louisiana Tumor Registry
mhsieh@lsuhsc.edu

John Keighley
Kansas Cancer Registry
jkeighle@kumc.edu

Gary Levin
Florida Cancer Data System
glevin@med.miami.edu

Nancy Lozon
Metropolitan Detroit Cancer Surveillance System
lozon@karmanos.org

Cheryl Moody
California Cancer Registry
cmoody@ccr.ca.gov

Gail Noonan
Manitoba Cancer Registry
gail.noonan@cancercare.mb.ca

Sheena Batts
Kentucky Cancer Registry
sbatts@kcr.uky.edu

Linda Coyle
Information Management Services, Inc.
coyle@imsweb.com

Sharmen Dye
sdye@wvmedical.com

Sandra Gamber
Elekta Impac Software
Sandra.Gamber@elekta.com

Jim Hofferkamp
NAACCR
jhofferkamp@naaccr.org

Tiffany Janes
Seattle Cancer Surveillance System
tjanes@fhcrc.org

Mary Jane King
Ontario Cancer Registry
maryjane.king@cancercare.on.ca

Mary Lewis
Centers for Disease Control and Prevention
bkf5@cdc.gov

Bobbi Matt
State Health Registry of Iowa
bobbi-matt@uiowa.edu

Linda Mulvihill
Centers for Disease Control and Prevention
LMulvihill@cdc.gov

Lisa Paret
Louisiana Tumor Registry
lparet@lsuhsc.edu

Steven Peace
Florida Cancer Data System
speace@med.miami.edu

Emily Reed
Kentucky Cancer Registry
ereed@kcr.uky.edu

Frances Ross
Kentucky Cancer Registry
fer@kcr.uky.edu

Wendy Scharber
Registry Widgets
wendy@registrywidgets.com

Colleen Sherman
New York State Cancer Registry
cgs02@health.ny.gov

Shannon Vann
NAACCR
svann@naaccr.org

Susan Perez
Texas Cancer Registry
susan.perez@dshs.state.tx.us

Deirdre Rogers
Mississippi Cancer Registry
dbrogers@umc.edu

Randi Rycroft
Colorado Central Cancer Registry
randi.rycroft@state.co.us

Jennifer Seiffert
Centers for Disease Control and Prevention
hzs4@cdc.gov

Deborah Towell
Oregon State Cancer Registry
deborah.j.towell@state.or.us

Judy Williams
Elekta Impac Software
judy.williams@elekta.com

INTRODUCTION

Background

The North American Association of Central Cancer Registries (NAACCR) Record Consolidation Committee was an *ad hoc* committee first appointed in April 1997 charged with examining the principles and approaches to record consolidation in use at that time. Guidelines and descriptions for performing this function in central cancer registries were summarized in a 1999 report to NAACCR titled *Central Cancer Registry Record Consolidation: Principles and Processes*. This report provided some standardized definitions of terms, as well as an outline of the overall processes involved in record consolidation: patient linkage, tumor linkage, and determination of best data values.

The Committee went on to conduct a test using a file of 656 actual source records (de-identified) from hospitals, pathology laboratories, physician offices, and death certificates. The file was given to 11 central cancer registries to subject to their routine record consolidation operations, as well as to a panel of 3 cancer registry experts who performed manual consolidation. The resulting consolidated files were compared with the expert panel's file and with each other on counts of tumors by patient and counts by type of cancer. The results of this study may be found in the report to the NAACCR Board titled *Creation of a Record Consolidation Test File*. The Committee was dissolved once these activities were completed.

Subsequent NAACCR Work Groups have studied approaches to the related activities of tumor linkage and automated determination of multiple primaries. This work resulted in a "Same Sites Table" and a "Same Histology Table" for use in automating these processes. The final step in record consolidation—determination of the final best data values—is the focus of the current Data Item Consolidation Work Group (DIWG).

Scope of Work

The purpose of the DIWG is to document consensus best practices for consolidating discrepant values for the same patient and tumor from different reporting sources into a single best value for analysis purposes in the central cancer registry. The DIWG was charged with documenting different practices that are used in instances for which there is no consensus.

Before data item consolidation can occur in a central cancer registry, several other essential steps must have taken place:

1. **Application of appropriate data edits to, and sometimes visual review of, source records.** Data to be consolidated should be as complete and accurate as the registry can make them so that the most complete and accurate source record values contribute to consolidation. Data should pass, at a minimum, standard-setter required data edits. If the central registry performs visual review of coded values against text, this step should be completed before consolidation. (Visual review may also be needed during the consolidation or post-processing phases.)
2. **Patient linkage.** The incoming source records must be linked to any existing records for the same patient in the database.

3. **Tumor Linkage.** Once linked to a patient, the incoming source records must be linked to any existing records for the same tumor.

Other NAACCR committees and work groups have addressed these steps. The DIWG is addressing the subsequent step of determining a single value for each data item. Determining a single value is only necessary when there are discrepant values reported in different source records; the goal of the DIWG is to document best practices for choosing the optimal value to save in a consolidated record when multiple source records describing the same patient and cancer contain some discrepant data values.

The goal of data item consolidation is the selection of the best value when discrepant data values are present, based on all submitted codes and text, in accordance with coding rules in published standards. Historically, the gold-standard model for the consolidation process has been 100% manual review with selection of best values by a trained cancer registrar (i.e., a certified tumor registrar [CTR]). The process entails reviewing all codes and text from all source records and applying all existing relevant coding rules and standards (FORDS, SEER Manual, ICD-O-3, Multiple Primary and Histology Coding Rules, etc.) to select a “best value” for each data item. The DIWG believes that this model needs to be replaced by one that recognizes the advantages of automated processes.

DIWG Work Group members’ experience has shown that simple variables can successfully be consolidated via automated rules-based decision making, often without the need for manual review. On the other hand, complex, inter-related, and compound variables such as component variables for staging, multiple variables used to code surgical treatment, and treatment sequencing variables are more challenging and more often require manual review by a trained cancer registrar to determine the “best and final” code value. Human and machine decision making are not always equivalent. A trained registrar might have knowledge that is either unknown to the computer system or impractical to automate. For example, a registrar might use knowledge about the patient’s address, such as residence on an Indian reservation, to resolve a race code discrepancy, while programming an algorithm to use address to consolidate race would not be feasible. The registrar’s use of local knowledge could also introduce bias, while a computer algorithm will be consistently applied to each record.

The DIWG is keenly aware that many central registries do not have the staff or resources to review each incoming source record and each abstracted variable, given the volume of their incoming cancer reports and anticipated increases in the number of case reports received because of direct reporting by physicians under Meaningful Use. Thus in many registries, for reasons of both data quality and practicality, the consolidation process has evolved to incorporate automated rules-based consolidation for many person and tumor variables.

The DIWG believes that using an approach combining automated and manual methods will streamline the consolidation process, reduce resources needed for manual review, reduce local bias in final decision making, create reproducible and consistent results across registries, and improve “best data” selection, overall. Most registries will or already do apply a mix of automated and manual methods to achieve a final consolidated record. The “gold standard model” needs to evolve to recognize this reality. Best practices also need to acknowledge that the level of manual review compared to automated

decision making can vary widely based on the registries' purposes, philosophies, operational approaches, and available resources. It is not the purpose of the DIWG to establish a standard for the amount of automation to use in consolidation.

The DIWG has documented, for each data element, a list of applicable logical rules in a proposed order of application, usually ending with a step to review manually if the prior steps do not result in a single value. Central registries can use the documentation to assess their current practices and consider changes to their systems. It is expected that each registry will make its own decision about the applicability of the rules to their registry based on the purpose, philosophy, operational approach, and available resources. The DIWG tried to limit the data items included in the initial document to those data items required by all standard setters.

Once the consolidated values are determined, the DIWG recommends that, at a minimum, standard-setter required data edits should be run against the consolidated data to ensure that consolidated records are internally consistent, valid, and logically accurate.

Processing of subsequent source records for the same tumor from a facility/source that has already reported the cancer is beyond the scope of the DIWG, although it may have an impact on the approach that a registry chooses to implement for data item consolidation. This falls into the category of follow-up processing or correction processing.

Operational Issues

A number of operational issues that individual registries must address when designing a record consolidation procedure are discussed in the following paragraphs.

Ensuring Data Quality with Automation

As mentioned above, the goal of data item consolidation when discrepant values are present is selection of the best value based on all submitted codes and text, in accordance with coding rules in published standards.

Any automated data item consolidation procedures should be designed to achieve or approach this goal as nearly as possible. Given that some level of automation is necessary and desirable, a threshold for acceptable accuracy should be set and record consolidation procedures should be periodically assessed to determine if an acceptable level of accuracy is being achieved. Data quality may be measured by its fitness for use. Therefore, when setting the threshold for data quality, it is important to keep in mind the ultimate uses for these data as well as the cost and efficiencies gained by automated processing. Thresholds will likely vary by data item, with those used for patient linkage and tumor identification having the highest threshold.

The 1999 *Central Cancer Registry Record Consolidation: Principles and Processes* report includes brief descriptions of the balance points of "accuracy vs. specificity" and "automation vs. staff review." These balance points are moderated by cost and timeliness. Each central registry must decide individually how best to achieve accuracy (possibly foregoing specificity) in the most efficient and cost-effective manner.

Work Flow Processing

Two options have been identified for an overall approach to processing incoming records for consolidation. One is that all incoming records are consolidated automatically using electronic algorithms and immediately added to the database. During this process, status codes (or “flags”) are assigned to records needing manual review; records are subsequently reviewed after they have been added to the database. In the other option, incoming records are processed and consolidated individually (using a combination of automated and manual procedures) and then added to the database. Source records that require further manual review are sent to a “pending” file while awaiting the information needed by the reviewer to determine best data values. They are not added to the central registry database until all values have been resolved. Registries often take advantage of both options, automatically consolidating and adding records in a batch mode when possible, but holding conflicting records in a pending file for interactive processing when manual review is needed.

Another aspect of the processing work flow involves which records are compared during data item consolidation. Are incoming source records compared to all associated source records for a given tumor or are incoming abstracts compared only to the current consolidated values in the current consolidated record? There are advantages and disadvantages to each approach. For example, for the data item Date of Birth, one option may be to take the most frequently reported date as most likely to be accurate. The consolidation algorithm would need all source records available to use this rule. On the other hand, if two discrepant records had already been manually reviewed and consolidated to the correct value confirmed to be correct, then the registry would want to continue to use that confirmed value and not compare to all values or re-review. This method would require a tracking mechanism in order to determine which values had already been manually reviewed and confirmed.

Finally, all data items in a source record may not be considered equally critical. Some items may be determined to be dependent or subordinate to other items, and therefore the values are carried into the consolidated record based on another item’s consolidated value or source record.

Weighting Source Records

It was stated above that incoming records must be complete and accurate at some minimum level before being submitted to record consolidation processing. However, it is acknowledged that the completeness and quality of information in a cancer source record varies widely by the source of that record. Several data items within a source record carry information about the source of the record: the reporting facility identifier, type of reporting source, class of case, etc. Central cancer registries may have additional data about reporting sources that could be used to maximize accuracy in consolidated records during automated processing (i.e., Commission on Cancer (CoC) approval status for hospitals, or large research institutions vs. small community hospitals). These characteristics of the source records may be used very effectively for automated record consolidation.

Class of Case for Central Registries

The “Class of Case” code reflects a facility’s role in the management of the cancer that is being entered into that facility’s registry and that may be reported to external organizations (such as the CoC or the state central registry). The Class of Case categories describe the relationship of the cancer patient to the

reporting facility: whether or not the patient was diagnosed or treated there or both, and whether or not it was before or after the facility's registry reference date. As such, Class of Case categories can be very useful to a central registry for record consolidation purposes. They can be used to prioritize information from multiple reporting facilities.

Typically, cancers diagnosed and/or treated with first course therapy at a reporting facility are given priority over the same cases reported by facilities as "diagnosed elsewhere," "diagnosed prior to a reference date," or "diagnosed at a staff physician's office" or "lab specimen only" cases. Facilities diagnosing the cancer or providing initial treatment are expected to have more detailed and possibly more accurate information regarding diagnosis, stage at diagnosis, and treatment.

Although the following information can be used as a guide, the central registry's processing procedures, evaluation of data received, and reporting requirements should be considered when implementing Class of Case priorities in consolidation practices.

Although consolidation practices are based on the assumption that the incoming data are correct and edited, it is understood that they are correct *from the perspective of the reporting facility*. That is why the Class of Case code is so valuable—it indicates to the central registry when the patient with this cancer was seen along the continuum of care (at diagnosis, during workup, for treatment, for recurrence, etc.) at this reporting facility. It also indicates if it was before or after the facility's reference date. Even so, special situations can arise that necessitate additional review.

Class of Case code 99 (Case of Unknown Relationship to the Facility) is an example. It is unlikely ACoS-approved facilities would submit cases with Class of Case coded to 99; however, facilities not accredited by the ACoS may submit them, mainly due to confusion over definitions. Additionally, Class of Case code 99 may have been used by central registries before it was defined by the ACoS and it may have a special meaning in specific central registries. Depending on when and how Class of Case and Scope Regional Lymph Node Surgery = [9] code 99 are used in a central registry, this scenario could impact consolidation practices.

Another situation impacting consolidation priority is Class of Case code 43 (Pathology Lab Specimens Only). Because Specimen Date is routinely used for Date of Diagnosis in the absence of additional information in the lab reports, there may be a small percentage of cases from labs where the specimen is actually the metastatic site (although this is not stated on the lab report). If the lab case links to an existing patient in the database, the lab case may be identified as a separate primary because the primary site is different than the existing tumor on the database. This scenario impacts multiple primary determination as well as consolidation.

The lab specimen date may be the initial diagnosis the majority of the time; therefore, the lab cases could be considered a higher priority along with the other Class of Case codes indicating initial diagnosis and/or treatment at that facility. Registries should establish procedures for the rare instances when lab cases should not be prioritized.

Class of Case can also be used in conjunction with the data item Type of Reporting Source when considering consolidation rules.

As one approach for Central Registries, the Class of Case codes have been categorized into the following groups:

| Class of Case Codes Categorized By Group | | |
|---|-------------|---|
| Class of Case Group | Code | Code Description |
| Diagnosis Only/no Treatment at Reporting Facility | | |
| 0 | 00 | Initial diagnosis at the reporting facility AND all treatment or a decision not to treat was done elsewhere. |
| 0 | 30 | Initial diagnosis and all first course treatment elsewhere AND reporting facility participated in diagnostic workup (for example, consult only, treatment plan only, staging workup after initial diagnosis elsewhere). |
| Diagnosis and all or Part First Course Treatment at Facility | | |
| 1 | 10 | Initial diagnosis at the reporting facility or in a staff physician's office AND part or all of first course treatment or a decision not to treat was at the reporting facility, NOS. |
| 1 | 11 | Initial diagnosis in staff physician's office AND part of first course treatment was done at the reporting facility. |
| 1 | 12 | Initial diagnosis in staff physician's office AND all first course treatment or a decision not to treat was done at the reporting facility. |
| 1 | 13 | Initial diagnosis at the reporting facility AND part of first course treatment was done at the reporting facility; part of first course treatment was done elsewhere. |
| 1 | 14 | Initial diagnosis at the reporting facility AND all first course treatment or a decision not to treat was done at the reporting facility. |
| 1 | 34 | Type of case not required by CoC to be accessioned (for example, a benign colon tumor) AND initial diagnosis AND part or all of first course treatment by reporting facility. |
| 1 | 35 | Case diagnosed before program's reference date AND initial diagnosis AND all or part of first course treatment by reporting facility. |
| 1 | 42 | Non staff physician or non-CoC accredited clinic or other facility, not part of reporting facility, accessioned by reporting facility for diagnosis and/or treatment by that entity. |
| Diagnosed Elsewhere/all or Part First Course Treatment at Facility | | |
| 2 | 20 | Initial diagnosis elsewhere AND all or part of first course treatment was done at the reporting facility, NOS. |
| 2 | 21 | Initial diagnosis elsewhere AND part of first course treatment was done at the reporting facility; part of first course treatment was done elsewhere. |
| 2 | 22 | Initial diagnosis elsewhere AND all first course treatment or a decision not to treat was done at the reporting facility. |
| 2 | 36 | Type of case not required by CoC to be accessioned (for example, a benign colon tumor) AND initial diagnosis elsewhere AND all or part of first course treatment by reporting facility. |
| 2 | 37 | Case diagnosed before program's reference date AND initial diagnosis elsewhere AND all or part of first course treatment by facility. |
| Diagnosis and First Course Treatment Elsewhere | | |

| | | |
|---|----|--|
| 3 | 31 | Initial diagnosis and all first course treatment elsewhere AND reporting facility provided in-transit care; or hospital provided care that facilitated treatment elsewhere (for example, stent placement). |
| 3 | 32 | Diagnosis AND all first course treatment provided elsewhere AND patient presents at reporting facility with disease recurrence or persistence (active disease). |
| 3 | 33 | Diagnosis AND all first course treatment provided elsewhere AND patient presents at reporting facility with disease history only (disease not active). |
| Diagnosed by Autopsy at Reporting Facility | | |
| 5 | 38 | Initial diagnosis established by autopsy at the reporting facility, cancer not suspected prior to death. |
| Diagnosis and First Course Treatment at Staff Physician Office | | |
| 6 | 40 | Diagnosis AND all first course treatment given at the same staff physician's office. |
| 6 | 41 | Diagnosis and all first course treatment given in two or more different staff physician offices. |
| Pathology or Lab Specimen | | |
| 7 | 43 | Pathology or other lab specimens only. |
| Death Certificate Only | | |
| 8 | 49 | Death certificate only. |
| Unknown Relationship to Facility/Code Defined by Central Cancer Registry | | |
| 9 | 99 | Non-analytic case of unknown relationship to facility. |

Another approach would be to establish a hierarchy for each individual Class of Case code; this hierarchy may differ when consolidating different data items.

For data item consolidation, the DIWG recommends that a Class of Case hierarchy be invoked based on these established groups. Hierarchies can also be established within each group to prioritize codes within the groupings. A Class of Case hierarchy is provided for each data item in the Tumor Data Item section of this report. The hierarchy may be prioritized differently based on the type of data items. For example, the hierarchy or prioritization may be different for treatment data items as compared to diagnostic or staging data items.

Given these above operational issues, the remainder of this report will detail, item by item, automated processing steps that may be taken to determine the best data value for the consolidated cancer record. Individual registries may choose which steps are best suited for their operations, or implement other steps that work best in their unique environment. These are presented as consensus guidelines, with advantages, disadvantages, and rationales for specific choices.

Recommended Next Steps

- Test different approaches with comparison of consolidation results
 - Individual data item consolidation
 - Grouping inter-related variables for consolidation
 - “Carry-along” variables
 - Establish a threshold of “what is best” to aid in cut-off for what requires manual review
- Introduce the concept of “Best Source” tracking at the variable level (or group of variables) beyond Class of Case or Type of Reporting Source
- Maintain existing business rules

- Changes to reporting requirements
- Introduction of new data Items
- Introduction of new EDITS
- Changes to existing data Items (definition or codes)
- Changed EDITS
- Enhance consolidation rules to incorporate new data sources and new data types
 - Medical claims records including procedure codes (HCPCS, CPT, ICD-10-PCS)
 - Pharmacy records
 - Other sources
- Enhance consolidation rules to incorporate new or revised coding systems and/or standards
 - ICD-10-CM
 - ICD-10-PCS
 - CPT
 - HCPCS
 - ICD-O updates
 - Pharmacy and SEER*Rx

Demographic Data Items

Considerations

The demographic data items characterize the person with cancer, not the tumor or the treatment for the disease. Therefore, many of these data items are confidential as well as critical for linking reports about the same patient. Many of these items are required to be collected by the central registry but are never transmitted to any standard setter. Other demographic data items are critical for incidence and mortality reporting (i.e., race and sex) even though they are not confidential. The demographic data items are unique in that there are many external electronic files available to the central registry that can validate the information in these data items, such as birth and death records. However, these files are not submitted by reporting sources to the central registry; they are not in NAACCR format, and thus they are not part of these record consolidation rules except as mentioned under “Manual Review.” Finally, some of the demographic data item values are filled in by the central registry itself, such as occupation and industry codes, and may not require consolidation rules. Each of these considerations has an impact on the record consolidation rules and is further addressed below.

Confidential items

The demographic items that fall into the Confidential portion of the NAACCR record layout include all of the patient name fields, Social Security Number (SSN), Address at Diagnosis Street and Address Current Street, Telephone Number, and Death Certificate State File Number. Patient Last Name, First Name, and SSN are critical to linking multiple cancer source records about the same person, and for linking the patient to other external data files, such as birth and death records, Social Security Administration files, Center for Medicare and Medicaid files, and National Death Index files. For these critical items, there are few acceptable consolidation rules; discrepancies are generally manually reviewed due to the importance of these values being correct and to the widespread availability of external validation sources.

Other demographic data items are confidential, but not as critical. The additional name fields (Middle Name, Maiden Name, and Alias) are useful for patient linkage; Street Address at Diagnosis is needed for geocoding; Street Address Current, Telephone Number, and DC State File number might only be useful to the central registry’s internal operations.

Critical Items

In addition to the critical confidential items listed above, the items Date of Birth, Race, Spanish Origin, Sex, Address at Diagnosis State, and Vital Status are also critical for reporting cancer incidence and mortality rates. Again, there are few acceptable consolidation rules; discrepancies are generally manually reviewed due to the importance of these values and to the availability of external data sources.

Collected but not Transmitted Items

As with the confidential data items, some of the occupation and industry data items that are collected are not transmitted to any standard setter. Resolving discrepancies in the occupation and industry fields is very labor intensive. Unless special project funds are available and local uses demand it, most

registries do not have the resources required. Also, occupation and industry codes (not text) are generated by the central registries themselves, and may only be reported through state data exchange records. Therefore, extensive consolidation rules are not presented. However, more elaborate consolidation rules could be considered by a registry with more resources available and the need for these data.

Address Items

Current address data items are included in the demographic section. Address at diagnosis items are found in the tumor data items because they are specific to each cancer diagnosis.

Address fields must be grouped together for consolidation in case multiple reporting sources report a different address. If a different address is reported and Addr at Dx--No & Street is selected from Facility A, but Addr at Dx -- City is selected from Facility B, not only will the consolidated address at Dx data items be mismatched, an issue will result in geocoding as well.

It is also important to acknowledge that the consolidation of address is specific to a point in time. For Address at Dx, Date of 1st Contact or Date of Diagnosis should be considered so the earliest most complete address at the time of initial diagnosis is selected for consolidation. For Current Address data items, Date of Last Contact should be considered so the latest most complete address is selected for consolidation.

If Census Tract Certainty has been assigned to one of the records, registries may want to check the value. For example: If Certainty=4 (indicating the geocode vendor could not locate the address) the record in the registry database could have an incorrect house #, street name or even city name.

We should also remember that sometimes an UNKNOWN address @ DX is the CORRECT one and shouldn't be overwritten by an incoming KNOWN address. For example, a patient living in Korea, diagnosed with stomach cancer and had surgery in KOREA....subsequently came to the U.S. for additional treatment. ADDRESS @ DX is in KOREA, not the U.S. We don't have specific street and city name but the country KOREA is coded in Addr @ DX...and UNKNOWN values should be considered correct and shouldn't be overwritten by an incorrect 'known' address;

Demographic Data Items

Race 1 [160]

Data Item Category: Demographic

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|---------------------------|----------------------------------|---------------|---------------|---|
| | | 99 | 01-98 | Unknown defined as 99 Select any known race value 01-98 over 99 (unknown). |
| Step 2.1 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | | 98 | 03-97 | Select any specific race value 01-97 over 98 (other race). |
| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | | 15 | 16, 17 | Select value 16 (Asian Indian) or 17 (Pakistani) over 15 (Asian Indian or Pakistani). |
| Step 2.3 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | | 96 | 04-06, 08-17 | Select any specific Asian race code 04-06, 08-17 over code 96 (Asian, NOS). |
| Step 2.4 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | | 97 | 07, 20-32 | Select any specific Pacific Islander code 07, 20-32 over 97 (Pacific Islander, NOS). |
| Step 3 | Most Frequent | | | Comment |
| | | | | Take the most frequently occurring value other than 99 and 98. Values from the same reporting facility are counted only once. If 99 or 98 are the only values, then those codes are used. |
| Step 4 | Manual Review | | | Comment |
| | | | | Consolidate manually if multiple specific race codes remain. |
| Input Fields | | | | |
| Race 1 | | | | |
| Output Fields | | | | |
| Race 1 | | | | |
| Record Selection Criteria | | | | |
| | | | | |
| Secondary Consolidation | | | | |
| | | | | |

Suggested QC Evaluation

| | |
|---------------------------------------|---|
| If consolidated value of Race 1 is 98 | Review for misunderstanding of Ethnicity versus Race or see SEER Appendix D |
|---------------------------------------|---|

Issues to Consider

| | |
|---|------------------------|
| If consolidated value of Race 1 is 01 and the value of Race 2 is 02-32, 96-98 | Flag for Manual Review |
|---|------------------------|

| | |
|--|------------------------|
| If consolidated value of Race 1 is equal to any of Race 2, Race 3, Race 4, or Race 5 | Flag for Manual Review |
|--|------------------------|

| | |
|---|------------------------|
| If one of Race 2, Race 3, Race 4, or Race 5 is 07 | Flag for Manual Review |
|---|------------------------|

| | |
|--|------------------------|
| If consolidated value of Race 1 is 98 and any of Race 2-5 is coded 02-32, 96, 97 | Flag for Manual Review |
|--|------------------------|

| | |
|--|------------------------|
| If consolidated value of Race 1 is 97 and any of Race 2-5 is coded 07, 20-32 | Flag for Manual Review |
|--|------------------------|

| | |
|---|------------------------|
| If consolidated value of Race 1 is 96 and any of Race 2-5 is coded 04-06, 08-17 | Flag for Manual Review |
|---|------------------------|

Be sure to adjust Spanish Surname if Race is changed to a Non-Hispanic race code (Native American, Filipino, etc.)

Race 2 [161], Race 3 [162], Race 4 [163], Race 5 [164]

Data Item Category: Demographic

| Known Over | | | | |
|---|----------------|----------------------|----------------------|---|
| Step 1 | Unknown | Unknown Value | Known Value | Comment |
| | | 99, 88 | 01-32, 96-98 | Unknown defined as 99, 88 Select any known race value 01-32, 96-98 over 99 (unknown) or 88 (no other race documented). |
| Unknown Race 2-5 if Race 1 Unknown | | | | |
| Step 2.1 | Unknown | Unknown | No Other Race | Comment |
| | | 99 | 88 | If consolidated value of Race 1 is not 99 Select 99 if consolidated Race 1 is 99. |
| First Record Reported | | | | |
| Step 2.2 | | | | Comment |
| | | | | Keep first code reported if multiple specific race codes remain. Rationale: this is not a critical variable and may not warrant the time required for manual review. |
| Manual Review | | | | |
| Step 3 | | | | Comment |
| | | | | Consolidate manually if multiple specific race codes remain, and local demands and resources permit. |
| Step 4 | | | | |
| | | | | Comment |
| Input Fields | | | | |
| | | | | Race 1, 2, 3,4, 5 |
| Output Fields | | | | |
| | | | | Race 2, 3, 4, 5 |
| Record Selection Criteria | | | | |
| | | | | |
| Secondary Consolidation | | | | |
| | | | | |
| Suggested QC Evaluation | | | | |
| | | | | |

Issues to Consider

| | |
|--|------------------------|
| If consolidated value of Race 1 is 01 and the value of Race 2 is 02-32, 96-98 | Flag for Manual Review |
| If consolidated value of Race 1 is equal to any of Race 2, Race 3, Race 4, or Race 5 | Flag for Manual Review |
| If one of Race 2, Race 3, Race 4, or Race 5 is 07 | Flag for Manual Review |
| If consolidated value of Race 1 is 98 and any of Race 2-5 is coded 02-32, 96, 97 | Flag for Manual Review |
| If consolidated value of Race 1 is 97 and any of Race 2-5 is coded 07, 20-32 | Flag for Manual Review |
| If consolidated value of Race 1 is 96 and any of Race 2-5 is coded 04-06, 08-17 | Flag for Manual Review |

Be sure to adjust Spanish Surname if Race is changed to a Non-Hispanic race code (Native American, Filipino, etc.)

Spanish Origin [190]

Data Item Category: Demographic

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--|--|---------------|------------------------|----------------------|
| | | 9 | 0-8 | Unknown defined as 9 |
| Select known value 0-8 over 9 (unknown). | | | | |
| Step 2.1 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | | 7 | 0-6, 8 | |
| Select specific Spanish code 0-6, 8 over 7 (Spanish surname only).* | | | | |
| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | | 6 | 1-5, 8 | |
| Select specific Spanish code 1-5, 8 over 6 (Spanish, NOS). | | | | |
| Step 2.3 | Spanish Over Non-Spanish | Non-Spanish | Spanish | Comment |
| | | 0 | 1-6,8 | |
| Select specific Spanish code 1-6, 8 over 0 (Non-Spanish). | | | | |
| Step 2.4 | Less Specific Spanish, NOS if Multiple Discrepant Codes Remain | Spanish, NOS | Specific Spanish Codes | Comment |
| | | 6 | 1-5,8 | |
| Select Spanish, NOS if multiple specific Spanish codes remain. Rationale: if this is not a critical variable, it may not warrant the time required for manual review. | | | | |
| Step 3 | Manual Review | | | Comment |
| Consolidate manually if multiple specific codes remain and local demands and resources permit. | | | | |
| Step 4 | | | | Comment |
| May be able to use Birthplace code to find Spanish Origin. | | | | |

Input Fields

Spanish Origin

Birthplace

Output Fields

Spanish Origin

Special Consideration

For cases diagnosed prior to 2005, select 8 over 5

NOTE for Section 2.1: Many times one abstractor or visual editor identifies that the Surname is on the Spanish surname list, but another abstractor submitting the case from a different source may not have looked up the Surname to check it.

Suggested QC Evaluation

If consolidated Spanish Origin is 1-8, Race 1=98
and Race 2-5=88

Review for misunderstanding of
Ethnicity/Race

If Race 1 is 03, 05, and 06 Race 2-6=88 and Spanish
Origin is not 0

Review for possible recode of Spanish Origin
to 0

If Race 1=01, Race 2-9=88, (Birthplace-Country is
BRA or PRT or Place of Birth is 341 or 445) ,
and Spanish Origin not 0

Review for possible recode of Spanish Origin
to 0

Sex [220]

Data Item Category: Demographic

| Step 1 | Known Over Unknown | | Comment |
|--|----------------------------|--------------|--|
| | Known 1-6 | Unknown 9 | Select the record with sex coded 1-6 over 9 (unknown). |
| Step 2 | Most Frequent | | Comment |
| | | | Select the most frequently occurring value other than 9 (unknown). |
| Step 3 | Use Other Electronic Lists | | Comment |
| Compare name to sex-stratified first name lists. | | | If available, use lists of highly likely male first names or female first names to determine sex code. |
| Step 4 | | | Comment |
| Refer for manual review. | | | Sex is a critical field for incidence reporting. Manual review of discrepancies is warranted. |
| Input Fields | | | |
| Sex | | | |
| First Name | | | |
| Output Fields | | | |
| | | | |
| Record Selection Criteria | | | |
| | | | |
| Secondary Consolidation | | | |
| | | | |
| Suggested QC Evaluation | | | |

A Sex/Name utility is available at

<http://www.naaccr.org/StandardsandRegistryOperations/SexCodeUtility.aspx>

Date of Birth [240]

Data Item Category: Demographic

| Step 1 | Known Over Unknown | Unknown |
|--|--|---|
| | | Any date with blanks for month or day or year. Select a known date of birth over any unknown date. |
| Step 2 | Most Frequent | Comment |
| | Select most frequently reported date of birth among known dates. | Most frequently reported Date is most likely correct. |
| Step 3 | | Comment |
| | Refer for manual review. | Date of birth is a critical field for patient linkage. Manual review of discrepancies is warranted. |
| Input Fields | | |
| Output Fields | | |
| Record Selection Criteria | | |
| Secondary Consolidation | | |
| Suggested QC Evaluation | | |
| A Sex/Name utility is available at http://www.naacr.org/StandardsandRegistryOperations/SexCodeUtility.aspx | | |
| Issues to Consider | | |

Date of Birth Flag [241]

Data Item Category: Demographic

| Step 1 | Comment |
|--------|--|
| | Take value from record from which Date of Birth is selected. |

| Step 2.1 | Comment |
|----------|---------|
|----------|---------|

| Step 2.2 | Comment |
|----------|---------|
|----------|---------|

| Step 3 | Comment |
|--------|---------|
|--------|---------|

| Step 4 | Comment |
|--------|---------|
|--------|---------|

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Birthplace - State [252]

Data Item Category: Demographic

Step 1 Select Valid Code for Birthplace

If the place of birth is known, make sure place of birth is valid code.

Step 2.1 Known Over Unknown

Select any known code over Birthplace State = Unknown (ZZ).

Step 2.2 Known Over Unknown

Select known code over YY (outside US and Canada [CD], country unknown).

Step 2.3 Known Over Unknown

Select known code over XX (outside US and CD, country known).

Step 3 Place of Birth Known

1. Select specific state or province over US or CD.
2. Select most frequently reported known Place of Birth.

Step 4 Place of Birth Known

If multiple known codes remain, select the Place of Birth from the earliest report.

Input Fields

| | |
|-----|--------------------|
| 250 | Birthplace |
| 252 | Birthplace State |
| 254 | Birthplace Country |

Output Fields

| | |
|-----|--------------------|
| 250 | Birthplace |
| 252 | Birthplace State |
| 254 | Birthplace Country |

Record Selection Criteria

All records included

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Birthplace - Country [254]

Data Item Category: Demographic

Step 1 Select Valid Code for Birthplace

If the place of birth is known, make sure place of birth is valid code.

Step 2.1 Known Over Unknown

Select any known code over Birthplace Country = Unknown (ZZU).

Step 2.2 Known Over Unknown

Select known code over ZZX (outside USA and CAN, country unknown).

Step 3 Place of Birth Known

1. Select specific country over NOS codes: ZZN, ZZC, ZZS, ZZP, ZZE, ZZF, ZZA.
2. Select most frequently reported known Birthplace Country.

Step 4 Place of Birth Known

If multiple known codes remain, select the Place of Birth from the earliest report.

Input Fields

| | |
|-----|--------------------|
| 250 | Birthplace |
| 252 | Birthplace State |
| 254 | Birthplace Country |

Output Fields

| | |
|-----|--------------------|
| 250 | Birthplace |
| 252 | Birthplace State |
| 254 | Birthplace Country |

Record Selection Criteria

All records included

Secondary Consolidation

Suggested QC Evaluation

Last Name [2230]

Data Item Category: Demographic

| Step 1 | Comment |
|---|--|
| Compare Last Name of incoming record to Last Name of record in central DB; if match, accept w/o review. | |
| Step 2.1 | Comment |
| If there is a discrepancy between incoming record Last Name and central record Last Name, reject for manual review. | It is essential to have the correct name in order to match the record to death certificates and other data linkages. |
| Step 2.2 | Comment |
| Follow-up back to hospitals and/or consult other resources (linkages, internet, etc.) to determine correct name. | |
| Step 3 | Comment |
| | |
| Step 4 | Comment |
| | |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

First Name [2240]

Data Item Category: Demographic

Step 1

Comment

Compare FN of incoming record to FN of record in central DB; if match, accept without review.

Step 2.1

Comment

Select full name over first initial.

Select full name if first character matches a first initial only.

Step 2.2

Comment

Select a formal name over a nickname.

User may create a lookup table of common nicknames and their formal names (e.g., Bob and Robert).

Step 2.3

Comment

Select from other electronic resources (i.e., linkage databases, batch files, etc.) to determine correct name.

User may create lookup tables from SSDI or birth certificates or other resources.

Step 3

Comment

If there is a discrepancy between incoming record FN and central record FN, reject for manual review.

It is essential to have the correct name in order to match the record to death certificates and other data linkages.

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Middle Name [2250]

Data Item Category: Demographic

| Step | Rule | Comment |
|---------------|--|---|
| Step 1 | Known Over Unknown | |
| | Known Middle Name overwrites blank Middle Name. | Select any characters over blank. |
| Step 2 | Known Over Unknown | |
| | Disregard NMI or NMN. | Select any characters over 'NMI' or 'NMN' which many use to mean 'no middle initial' or 'no middle name.' Select any characters over 'UNK', 'UNKNOWN'. Select any initial over 'U'. |
| Step 3 | More Specific Over Less Specific | |
| | Select full name over first initial. | Select full name if first character matches a first initial only. |
| Step 4 | | |
| | For Middle Name discrepancy, newest record overwrites current record. | Middle Name is not essential for matching; manual review of this item is not optimal use of resources. |
| Step 5 | | |
| | If there is a discrepancy between two records, reject for manual review. | Option for registries per local needs and resources. |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Review for First Name and Middle Name that are the same

Issues to Consider

Suffix [2270], Alias [2280], and Maiden Name [2390]

Data Item Category: Demographic

| Step 1 | Known Over Unknown | Comment |
|--------|--------------------|---------|
|--------|--------------------|---------|

Select any character string over blank.

| Step 2 | Most Recent | Comment |
|--------|-------------|---------|
|--------|-------------|---------|

For discrepancies in these items, select latest record over current record.

Suffix is not essential for matching; manual review of this item is not optimal use of resources.
Latest record in could reflect more current information.

| Step 3 | | Comment |
|--------|--|---------|
|--------|--|---------|

If there is a discrepancy between two records, reject for manual review.

Option for registries per local needs and resources.

| Step 4 | | Comment |
|--------|--|---------|
|--------|--|---------|

| Step 5 | | Comment |
|--------|--|---------|
|--------|--|---------|

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

MD, DDS, MR, MRS, MS, etc should not be consolidated. Registrars erroneously fill this field with this information rather than JR, SR, III, ESQ, etc.

Occupation

Data Item Category: Demographic

Data Item Name [#]: Usual Occupation Text [310]

| Step 1 | | Comment | |
|--|---------------------------|--|---------|
| If patient is under 14 do not consolidate; leave blank. | | | |
| Step 2 | Unknown Values | Known Values | Comment |
| Known over unknown. | Blank (see Issue 3 below) | Non-blank (see Issue 3 below) | |
| Step 3 | | Comment | |
| Keep first received if there are discrepant known values. [Do Not Consolidate If Discrepant.] | | Once a known value is present in the consolidated record, do not automatically update over time. | |
| Step 4 | | Comment | |

Input Fields

Usual Occupation Txt [310]

Age at Diagnosis [230]

Output Fields

Usual Occupation Txt [310]

Record Selection Criteria

Applies only to patients with age greater than 014.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

- 1 Code is assigned, usually by standardized software, based on text in item 310, Text--Usual Occupation. Text is of uncertain and often poor quality.
- 2 Resolving discrepancies in text and manually reviewing text that fails to automatically code is very labor intensive. Unless special project funds are available, most registries will not have the resources required. Therefore, extensive consolidation directives are not warranted. However, more elaborate consolidation directives could be considered by a registry with more resources available.
- 3 Codes to be considered as Unknown can include codes for "Unknown," "Retired," "Unemployed," etc. If codes for these or other similar terms have differed over time, then a consolidation rule might need to be written to use Census Occ/Ind Sys 70-00 [330] along with the code.
- 4 Consideration should be given to whether the occupation and industry codes should be selected from the same record.

Industry

Data Item Category: Demographic

Data Item Name [#]: Usual Industry Text [320]

Step 1 Age under 14

Comment

If patient is under 14 do not consolidate; leave blank.

Step 2 Known over Unknown

Unknown Values

Known Values

Comment

Blank (see Issue 3 below)

Non-blank (see Issue 3 below)

Step 3 First Received

Comment

Keep first received if there are discrepant known values. [Do not consolidate if discrepant.]

Once a known value is present in the consolidated record, do not automatically update over time.

Step 4

Comment

Input Fields

Usual Industry Text [320]

Output Fields

Usual Industry Text [320]

Record Selection Criteria

Applies only to patients with age greater than 014.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

- 1 Code is assigned, usually by standardized software, based on text in item 310, Text--Usual Occupation. Text is of uncertain and often poor quality.
- 2 Resolving discrepancies in text and manually reviewing text that fails to automatically code is very labor intensive. Unless special project funds are available, most registries will not have the resources required. Therefore, extensive consolidation directives are not warranted. However, more elaborate consolidation directives could be considered by a registry with more resources available.
- 3 Codes to be considered as Unknown can include codes for "Unknown," "Retired," "Unemployed," etc. If codes for these or other similar terms have differed over time, then a consolidation rule might need to be written to use Census Occ/Ind Sys 70-00 [330] along with the code.
- 4 Consideration should be given to whether the occupation and industry codes should be selected from the same record.

Current No & Street [2350]

Data Item Category: Demographic - Address Current

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|--------------------|---------------|---------------|---|
| | | Unknown | Valid address | This field is tied to associated fields for current address: city, state, postal code, country and telephone. |

| Step 2.1 | Most Current | Comment |
|----------|--------------|---------|
|----------|--------------|---------|

Current address should be taken from record with most current date of last contact - if full address is present.

| Step 2.2 | Less Specific | More Specific | Comment |
|---|---------------------------------|--|---------|
| More specific information consists of full address information. | Street name with missing number | Inclusion of apartment number - if appropriate | |

| Step 3 | Manual Review | Comment |
|--------|---------------|---------|
|--------|---------------|---------|

Manual review needed if any elements of address are missing on record with most current date of last contact.

| Step 4 | Comment |
|--------|---------|
|--------|---------|

| Input Fields |
|--------------|
|--------------|

| Output Fields |
|---------------|
|---------------|

| Record Selection Criteria |
|---------------------------|
|---------------------------|

| Secondary Consolidation |
|-------------------------|
|-------------------------|

| Suggested QC Evaluation |
|-------------------------|
|-------------------------|

| Issues to Consider |
|--------------------|
|--------------------|

Current City [1810]

Data Item Category: Demographic - Address Current

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|--|---------------|-------------|---|
| | If only one record, take information from that record. | Unknown | City name | This field is tied to associated fields for current address: city, state, postal code, country and telephone. |

| Step 2 | Comment |
|--------|---|
| | Current city should be taken from record with most current date of last contact - if full address is present. |

| Step 3 | Manual Review | Comment |
|--------|--|---------|
| | Manual review is needed if any elements of address are missing on record with most current date of last contact. | |

| Step 4 | Comment |
|--------|---------|
| | |

| Input Fields |
|--------------|
| |

| Output Fields |
|---------------|
| |

| Record Selection Criteria |
|---------------------------|
| |

| Secondary Consolidation |
|-------------------------|
| |

| Suggested QC Evaluation |
|-------------------------|
| |

| Issues to Consider |
|--------------------|
| |

Current State [1820]

Data Item Category: Demographic - Address Current

| | | Unknown | Known Value | Comment |
|----------------------------------|--|----------------------|---|---|
| Step 1 | Known Over Unknown | Value | Known Value | Comment |
| | If only one record, take information from that record | ZZ | Valid state, province or XX or YY if foreign resident | This field is tied to associated fields for current address: no & street, city, postal code, country and telephone. |
| Step 2.1 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | Current state should be taken from record with most current date of last contact - if full address is present. | | | |
| Step 2.2 | | | | Comment |
| | More specific information consists of a specific state vs. a country, if not foreign. | US or CD | State, province or territory | |
| Step 3 | Manual Review | | | Comment |
| | Manual review needed if any elements of address are missing on record with most current date of last contact. | | | |
| Input Fields | | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Secondary Consolidation | | | | |
| Suggested QC Evaluation | | | | |
| Issues to Consider | | | | |

Current Postal Code [100]

Data Item Category: Demographic - Address Current

| Known Over | | | | |
|---------------------------|---|--|---|--|
| Step 1 | Unknown | Unknown Value | Known Value | Comment |
| | If only one record take information from that record. | 99999 or 999999999 for US and CD; 88888 or 888888888 for foreign | Valid US or Canadian postal code | This field is tied to associated fields for address current: no & street, city, state, postal code, country. |
| More Specific | | | | |
| Step 2.1 | Over Less Specific | Less Specific | More Specific | Comment |
| | Postal code should be taken from record with most current date of last contact. | Only take more specific postal code if it agrees with other elements of address current. | | |
| More Specific | | | | |
| Step 2.2 | Over Less Specific | Less Specific | More Specific | Comment |
| | Only take more specific postal code if it agrees with other elements of address current. | Five-digit postal code for US addresses | Nine-digit extended postal code for US residences | |
| Step 3 | Manual Review | | | Comment |
| | Manual review needed if any elements of address are missing on record with most current date of last contact. | | | |
| Step 4 | | | | Comment |
| | | | | |
| Input Fields | | | | |
| | | | | |
| Output Fields | | | | |
| | | | | |
| Record Selection Criteria | | | | |
| | | | | |
| Secondary Consolidation | | | | |
| | | | | |
| Suggested QC Evaluation | | | | |
| | | | | |
| Issues to Consider | | | | |
| | | | | |

Current Country [1832]

Data Item Category: Demographic - Address Current

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|----------------------------------|---|------------------------------|---|---|
| | If only one record take information from that record. | ZZU | Valid value from appendix B of NAACCR Volume II or appendix E of FORDS 2013 | This field is tied to associated fields for current address: no & street, city, state, postal code and telephone. |
| Step 2.1 | Most Current Date | | | Comment |
| | Current county should be taken from record with most current date of last contact - if full address is present. | | | |
| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
| | More specific information is listed as 'Where the Detail is Known' on in appendix E of FORDS 2013 | Table of less specific codes | Preferred code | |
| Step 3 | Manual Review | | | Comment |
| | Manual review needed if any elements of address are missing on record with most current date of last contact. | | | |
| Step 4 | | | | Comment |
| Input Fields | | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Secondary Consolidation | | | | |
| Suggested QC Evaluation | | | | |
| Issues to Consider | | | | |

Address Current Supplemental [2355]

Data Item Category: Demographic - Address Current

| Known Over | | | | |
|---------------------------|---|---------------|--|--|
| Step 1 | Unknown | Unknown Value | Known Value | Comment |
| | If only one record, take information from that record. | Blank | Facility name or other information useful for geocoding | This field is tied to associated fields for current address: no & street, city, state, postal code, country and telephone. |
| Step 2 | | | | Comment |
| | Current address - supplemental should be taken from record with most current date of last contact. | | Blank is a valid value for this field and can be over-written. | |
| Step 3 Manual Review | | | | Comment |
| | Manual review needed if any elements of address are missing on record with most current date of last contact. | | | |
| Step 4 | | | | Comment |
| | | | | |
| Input Fields | | | | |
| | | | | |
| Output Fields | | | | |
| | | | | |
| Record Selection Criteria | | | | |
| | | | | |
| Secondary Consolidation | | | | |
| | | | | |
| Suggested QC Evaluation | | | | |
| | | | | |
| Issues to Consider | | | | |
| | | | | |

Current Telephone [2360]

Data Item Category: Demographic - Address Current

| Known Over | | | | |
|----------------------------------|--|----------------------|------------------------|---|
| Step 1 | Unknown | Unknown Value | Known Value | Comment |
| | If only one record take information from that record. | Blank | Valid telephone number | This field is tied to associated fields for current address: no & street, city, state, postal code and country. |
| More Specific | | | | |
| Step 2.1 | Over Less Specific | Less Specific | More Specific | Comment |
| | Current telephone should be taken from record with most current date of last contact. | | | |
| More Specific | | | | |
| Step 2.2 | Over Less Specific | Less Specific | More Specific | Comment |
| | More specific information consists of full address information. | Missing area code | Full telephone | Depending on locality, area code can be inferred from address. |
| Step 3 | Manual Review | | | Comment |
| | Manual review needed if any elements of phone number are missing on record with most current date of last contact. | | | |
| Step 4 | | | | Comment |
| Input Fields | | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Secondary Consolidation | | | | |
| Suggested QC Evaluation | | | | |
| Issues to Consider | | | | |

Social Security Number [2320]

Data Item Category: Demographic

| Unknown | | | | |
|--|----------------|--|--|--|
| Step 1 | Over Invalid | Invalid Value | Unknown Value | Comment |
| Select unknown over invalid value. | | 1. First three digits = 000; 2. First three digits = 666; 3. Fourth and fifth digits = 00; 4. Last four digits = 0000; 5. First digit = 9 (except when first digit of 999999999) | 999999999 | Some values in the range are invalid and fail edits. |
| Known | | | | |
| Step 2 | Over Unknown | Unknown Value | Known Value | Comment |
| Select known valid value over unknown. | | 999999999 | Valid value in range 001010001 899999999 | |
| Select | | | | |
| Step 3 | Verified Value | | | Comment |
| Select verified value. | | | | If system has a flag that a value is reviewed and correct, select it. |
| More | | | | |
| Step 4 | Most Frequent | Less Frequent | Frequent | Comment |
| Select most frequent valid value. | | | | Assumes you are reviewing all source records. |
| Manual Review | | | | |
| Step 5 | Manual Review | | | Comment |
| Refer for manual review. | | | | This is a critical field, it is usually possible to find correct SSN. Select correct SSN from other sources (linkages, hospital records, etc.) |
| Input Fields | | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Secondary Consolidation | | | | |
| Suggested QC Evaluation | | | | |
| Issues to Consider | | | | |

Vital Status [1760]

Data Item Category: Demographic

Step 1

If central record is 1 (Alive), and new record is 0 or 4 (Dead), AND new record has the same or later DLC, Vital status = 0 or 4 (Dead).

Comment

Vital status must be consolidated in conjunction with Date Last Contact.

Step 2

If central record is 0 or 4 (Dead), and new record is 1 (Alive), AND new record has earlier or same or unknown DLC, Vital status= Dead.

Comment

Step 3

If central record is 0 or 4 (Dead), and new record is 1 (Alive), AND new record has later DLC, refer for manual review.

Comment

Before referring for manual review, perform linkage with other electronic databases for confirmed deaths, if possible.

Step 4

If central record is 1 (Alive), and new record is 0 or 4 (Dead), AND new record has earlier or unknown DLC, refer for manual review.

Comment

Before referring for manual review, perform linkage with other electronic databases for confirmed deaths, if possible.

Step 5

Comment

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Step

Process

Suggested QC Evaluation

Issues to Consider

Place of Death [1940]

Data Item Category: Demographic

Step 1

Comment

Consolidate Vital status first. If Vital status = Alive, set value to 997.

Step 2.1

Comment

Consolidate Vital status first. If Vital status = Dead, select value from DCO record (class of case 49).

Place of Death defined as where DC filed. If one record is a DCO, select it.

Step 2.2

Comment

Consolidate Vital status first. If Vital status = Dead, select 999 over 997.

Step 2.3

Comment

Consolidate Vital status first. If Vital status = Dead, select any other value over 999 or 997.

Step 3

Comment

Consolidate Vital status first. If Vital status = Dead, select value from death certificate record.

Use external electronic vital status file for correct value, if available.

Step 4

Comment

Consolidate Vital status first. If Vital status = Dead, select value from record with latest Date of Last Contact.

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Step

Process

Suggested QC Evaluation

Issues to Consider

Place of Death - State [1942]

Data Item Category: Demographic

Step 1

Consolidate Vital status first. If Vital status = Alive, set value to blank.

Comment

Step 2

Consolidate Vital status first. If Vital status = Dead, select value from DCO record (class of case 49).

Best information is from death certificate, so use DCO case first.

Comment

Step 2.1

Consolidate Vital status first. If Vital status = Dead, select ZZ over blank.

Comment

Step 2.2

Consolidate Vital status first. If Vital status = Dead, select any valid value over ZZ.

Comment

Step 3

Consolidate Vital status first. If Vital status = Dead, select value from death certificate record.

Select value from external electronic DC file, if available.

Comment

Step 4

Consolidate Vital status first. If Vital status = Dead, select value from record with latest Date of Last Contact.

Comment

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Place of Death - Country [1944]

Data Item Category: Demographic

| Step 1 | Comment |
|--|---|
| Consolidate Vital status first. If Vital status = Alive, set value to blank. | |
| Step 2 | Comment |
| Consolidate Vital status first. If Vital status = Dead, select value from DCO record. | Best information from death certificate; take DCO record first. |
| Step 2.1 | Comment |
| Consolidate Vital status first. If Vital status = Dead, select ZZU over blank. | |
| Step 2.2 | Comment |
| Consolidate Vital status first. If Vital status = Dead, select any valid value over ZZU. | |
| Step 3 | Comment |
| Consolidate Vital status first. If Vital status = Dead, select value from death certificate record. | Use external electronic DC file, if available. |
| Step 4 | Comment |
| Consolidate Vital status first. If Vital status = Dead, select value from record with latest Date of Last Contact. | |
| Input Fields | |
| | |
| Output Fields | |
| | |
| Record Selection Criteria | |
| | |
| Secondary Consolidation | |
| | |
| Suggested QC Evaluation | |
| | |
| Issues to Consider | |
| | |

DC State File Number [2380]

Data Item Category: Demographic

Step 1

Comment

Select any non-blank value over blank.

Step 2.1

Comment

Select any value not = to unknown over unknown.

Step 2.2

Comment

Select value from state death certificate file. Use external electronic DC file, if available.

Step 3

Comment

Select value from reporting source with access to death certificate file. Likely source for these data is other State central registry records received through data exchange agreements.

Step 4

Comment

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Date of Death Canada [1755]

Data Item Category: Demographic Date of Death Canada

Step 1

Comment

Consolidate Vital Status first. If Vital status=1 (Alive), then Date of Death Canada must be blank (or appropriate default value).

Step 2.

Comment

Consolidate Vital Status first. If Vital status = 0 or 4 (Dead), then select blank date value over an invalid date value.

Step 3

Comment

Consolidate Vital Status first. If Vital status = 0 or 4 (Dead), then select any valid date over a blank date value.

Step 4

Comment

Consolidate Vital Status first. If Vital status = 0 or 4 (Dead), then select the valid date from the record with the highest priority of Type of Reporting Source where Vital Stat is also 0 or 4 (see notes).

Step 5

Comment

Else, manually review the source documents, other sources to choose the correct Date of Death CA.

Input Fields

Output Fields

Issues to Consider

Notes

Type of Reporting Source Hierarchy:

- 7 Death Certificate only
- 6 Autopsy only
- 1 Hospital inpatient
- 8 Other hospital outpatient unit
- 2 Radiation center
- 4 Physician office
- 5 Nursing home
- 3 Lab only

Date of Death - Canada Flag [1756]

Data Item Category: Demographic

Step 1

Comment

Take value from record from which Date of Death Canada is selected.

Step 2.1

Comment

Step 2.2

Comment

Step 3

Comment

Step 4

Comment

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Tumor Data Items

The Tumor Data Item Subgroup includes suggested consolidation rules for the following tumor data items: Address at Dx data items, Date of Diagnosis, Primary Site, Histologic Type, Behavior, Grade, Laterality, Diagnostic Confirmation, and Type of Reporting Source. The goal is to provide central registries with a set of recommended rules for each item that can be automated within any system.

Class of Case should be considered when consolidating tumor-specific data items because facilities diagnosing and treating the patient typically have detailed information to support the primary site and histologic diagnosis.

The following two Class of Case hierarchy methods are recommended as best practices in using Class of Case in tumor data item consolidation procedures:

- Class of Case Group Hierarchy: 1, 6, 2, 7, 0, 3, 5, 9, 8. (Refer to table, [Class of Case Codes Categorized](#))

Registries are encouraged to evaluate the two methods and incorporate the approach that works best within their consolidation procedures and database structure. Consideration should be given to registry-specific use of Class of Case as well as potential situations to review or flag for data quality review (e.g., lab specimen of metastatic site).

Address at Dx: No & Street [60]

Data Item Category: Demographic - Address at DX

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|---|--------------------|---------------|---------------|--|
| If only one record take information from that record. | | Unknown | Valid address | This field is tied to associated fields for address at dx: number & street, supplemental, city, state, postal code, country. |

| Step 2.1 | Class of Case | | | Comment |
|--|---------------|--|--|--|
| Number & street should be taken from record with class of case that includes diagnosis | | | | If more than one class of case = dx, must consolidate class of case first. |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|---|----------------------------------|---------------------------------|--|---------|
| More specific information consists of full address information. | | Street name with missing number | Inclusion of apartment number - if appropriate | |

| Step 3 | Manual Review | | | Comment |
|--------|---------------|--|--|---|
| | | | | Other than Address at Dx Supplemental, manual review needed if any elements of address are missing on record with earliest date of first contact. |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

Address at Dx: City [#70]

Data Item Category: Tumor

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|--------------------|---------------|-------------|--|
| | | Unknown | City name | This field is tied to associated fields for address at dx: no & street, city, state, postal code, country. |

| Step 2.1 | Class of Case | | | Comment |
|----------|---------------|--|--|--|
| | | | | If more than one class of case = dx, must consolidate class of case first. |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|----------|----------------------------------|---------------|---------------|--|
| | | | | Manual review needed if any elements of address are missing on record with earliest date of first contact. |

| Step 3 | Manual Review | | | Comment |
|--------|---------------|--|--|---------|
| | | | | |

| Step 4 | | | | Comment |
|--------|--|--|--|---------|
| | | | | |

| Input Fields |
|--------------|
| |

| Output Fields |
|---------------|
| |

| Record Selection Criteria |
|---------------------------|
| |

| Secondary Consolidation |
|-------------------------|
| |

| Suggested QC Evaluation |
|-------------------------|
| |

| Issues to Consider |
|--------------------|
| |

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

Address at Dx State [80]

Data Item Category: Tumor

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|---|--------------------|---------------|---|--|
| If only one record take information from that record. | | ZZ | Valid state, province or XX or YY if foreign resident | This field is tied to associated fields for address at dx: no & street, city, state, postal code, country. |

| Step 2.1 | Class of Case | Comment |
|----------|--|---------|
| | State should be taken from record with class of case equal to diagnosis. | |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|--|----------------------------------|---------------|------------------------------|--|
| More specific information consists of a specific state or province vs. country; if US or CD. | | US or CD | State, province or territory | If more than one class of case = dx, must consolidate class of case first. |

| Step 3 | Manual Review | Comment |
|--------|--|---------|
| | Manual review needed if any elements of address are missing on record with earliest date of first contact. | |

| Step 4 | Comment |
|--------|---------|
| | |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

County at DX [90]

Data Item Category: Tumor

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|---------------|---|----------------------|---|--|
| | If only one record take information from that record. | 998 or 999 | Code valid value from FIPS for US county - Canada use 998 | This field is tied to associated fields for address at dx: no & street, city, state, postal code, country. |

| Step 2.1 | Class of Case | Comment |
|-----------------|---|--|
| | County should be taken from record with class of case equal to diagnosis. | If more than one class of case = dx, must consolidate class of case first. |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|-----------------|---|----------------------|----------------------|----------------|
|-----------------|---|----------------------|----------------------|----------------|

| Step 3 | Manual Review | Comment |
|---------------|--|----------------|
| | Manual review needed if any elements of address are missing on record with earliest date of first contact. | |

| Step 4 | Comment |
|---------------|----------------|
|---------------|----------------|

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

Address at DX Postal Code [100]

Data Item Category: Tumor

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|---|--|--|--|
| | If only one record take information from that record. | 99999 or 999999999 for US and CD; 88888 or 888888888 for foreign | Valid US, Canadian, or known foreign postal code | This field is tied to associated fields for address at dx: no & street, city, state, postal code, country. |

| Step 2.1 | Class of Case | Comment |
|----------|---|---------|
| | Postal code should be taken from record with class of case equal to dx. | |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|----------|--|---|---|--|
| | Only take more specific postal code if it agrees with other elements of address at dx. | Five-digit postal code for US addresses | Nine-digit extended postal code for US residences | If more than one class of case = dx, must consolidate class of case first. |

| Step 3 | Manual Review | Comment |
|--------|--|---------|
| | Manual review needed if any elements of address are missing on record with earliest date of first contact. | |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

Address at DX Country [102]

Data Item Category: Tumor

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|---|---------------|---|--|
| | If only one record take information from that record. | ZZU | Valid value from appendix B of NAACCR v13or Appendix E of FORDS | This field is tied to associated fields for address at dx: no & street, city, state, postal code, country. |

| Step 2.1 | Class of Case | Comment |
|----------|--|--|
| | Country should be taken from record with class of case equal to diagnosis. | If more than one class of case = dx, must consolidate class of case first. |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|----------|--|------------------------------|----------------|---------|
| | More specific information is listed in "Where the detail is known" on the table. | Table of less specific codes | Preferred code | |

| Step 3 | Manual Review | Comment |
|--------|--|---------|
| | Manual review needed if any elements of address are missing on record with earliest date of first contact. | |

| Step 4 | Comment |
|--------|---------|
| | |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

Address at DX Supplemental [2335]

Data Item Category: Tumor

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|---|---------------|---|--|
| | If only one record take information from that record. | Blank | Facility name or other information useful for geocoding | This field is tied to associated fields for address at dx: no & street, city, state, postal code, country. |

| Step 2.1 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|----------|-----------------------------------|---------------|--|---------|
| | More specific over less specific. | | Blank is a valid value for this field and can be over-written. | |

| Step 2.2 | More Specific Over Less Specific | Less Specific | More Specific | Comment |
|----------|--|---------------|---------------|--|
| | Supplemental should be taken from a record with class of case equal to dx. | | | If more than one class of case = dx, must consolidate class of case first. |

| Step 3 | Manual Review | | | Comment |
|--------|--|--|--|---------|
| | Manual review needed if any elements of address are missing on record with earliest date of first contact. | | | |

| Step 4 | | | | Comment |
|--------|--|--|--|---------|
| | | | | |

Input Fields

Output Fields

Suggested QC Evaluation

Issues to Consider

Take records from [class of case hierarchy](#). If still have a tie - use earliest date of first contact.

Date of Diagnosis [390]

Data Item Category: Tumor

| Step 1 | Known over Unknown | Unknown Value | Known Value | Comment |
|--------|--------------------|--------------------|-------------|--|
| | | Blank Day or Month | Valid Date | Full date over partial date or unknown date. |

| Step 2 | Class of Case Hierarchy | Comment |
|--------|-------------------------|---------|
| | Class of Case Hierarchy | |

| Step 3 | Comment |
|--|--|
| Earliest Date of Diagnosis (final selector). | Use source record with earliest Date of Diagnosis. |

| Step 4 | Comment |
|--------|---------|
| | |

Input Fields

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

If Class of Case = 99, manual review.

Issues to Consider

Primary Site [400]

Data Item Category: Tumor

| Step 1 | | Class of Case Hierarchy | | Comment |
|---|--|----------------------------|-------------|---|
| Step 1 | | Class of Case Hierarchy | | Comment |
| Class of Case Hierarchy for an explanation of Class of Case for Central Registries. | | | | |
| Step 2 | | Unknown Value | Known Value | Comment |
| Step 2 | | Unknown Value | Known Value | Comment |
| Known over unknown subsite. | | Cxx.9 | Cxx.0-Cxx.8 | Unknown defined as 9; eliminate 9 if any other value 0-8. |
| Step 3 | | Use Multiple Primary Rules | | Comment |
| Step 3 | | Use Multiple Primary Rules | | Comment |
| | | | | Possibly use manual review flag. |
| Step 4 | | Comment | | |
| Step 4 | | Comment | | |
| Input Fields | | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Secondary Consolidation | | | | |
| Suggested QC Evaluation | | | | |
| If Class of Case = 99, manual review. | | | | |
| Issues to Consider | | | | |

Laterality [410]

Data Item Category: Tumor

Step 1

Known over Unknown

Comment

Step 2

Laterality pairs tables

If consolidated histology Hematopoietic or Kaposi Sarcoma, select laterality from case that primary site was selected from.
See Laterality Pairs Table A for Wilms, Retinoblastoma, epithelium ovarian, or inflammatory breast primaries.
See Laterality Pairs Table B for primary sites not listed on the Paired Organ Table (see Table D).
See Laterality Pairs Table C for primary sites listed on the paired Organ Table (Table D).

Comment

Step 3

Visual Review

Comment

Input Fields

Output Fields

Record Selection Criteria

Select Laterality from cases where Primary Site matches.

Secondary Consolidation

Suggested QC Evaluation

| Table A: Wilms, Retinoblastoma, epithelial ovarian or inflammatory breast primaries | | | |
|--|---------------------|--------------------------------|----------------------|
| Laterality 1 | Laterality 2 | Consolidated Laterality | Manual Review |
| 0 | 0 | 0 | No |
| 0 | 1 | 1 | No |
| 0 | 2 | 2 | No |
| 0 | 3 | 3 | No |
| 0 | 4 | 4 | No |
| 0 | 5 | 0 | No |
| 0 | 9 | 9 | No |
| 1 | 1 | 1 | No |
| 1 | 2 | 4 | Yes |
| 1 | 3 | 1 | No |
| 1 | 4 | 4 | No |
| 1 | 5 | 1 | Yes |
| 1 | 9 | 1 | No |
| 2 | 2 | 2 | No |
| 2 | 3 | 2 | No |
| 2 | 4 | 4 | No |
| 2 | 5 | 2 | Yes |
| 2 | 9 | 2 | No |
| 3 | 3 | 3 | No |
| 3 | 4 | 4 | No |
| 3 | 5 | 5 | No |
| 3 | 9 | 3 | No |
| 4 | 4 | 4 | No |
| 4 | 5 | 4 | No |
| 4 | 9 | 4 | No |
| 5 | 5 | 5 | No |
| 5 | 9 | 5 | No |
| 9 | 9 | 9 | No |

Some central registries may prefer to assign a value based on manual review for cases marked “yes” in the manual review column rather than autocoding a value and marking the case for later review.

| Table B: Primary Site not listed on the Paired Organ Table (Table D) | | | |
|---|---------------------|--------------------------------|----------------------|
| Laterality 1 | Laterality 2 | Consolidated Laterality | Manual Review |
| 0 | 0 | 0 | No |
| 0 | 1 | 0 | No |
| 0 | 2 | 0 | No |
| 0 | 3 | 0 | No |
| 0 | 4 | 0 | No |
| 0 | 5 | 0 | No |
| 0 | 9 | 0 | No |
| 1 | 1 | 1 | No |
| 1 | 2 | 1 | Yes |
| 1 | 3 | 1 | No |
| 1 | 4 | 1 | Yes |
| 1 | 5 | 1 | Yes |
| 1 | 9 | 1 | No |
| 2 | 2 | 2 | No |
| 2 | 3 | 2 | No |
| 2 | 4 | 2 | Yes |
| 2 | 5 | 2 | Yes |
| 2 | 9 | 2 | No |
| 3 | 3 | 3 | No |
| 3 | 4 | 3 | Yes |
| 3 | 5 | 3 | No |
| 3 | 9 | 3 | No |
| 4 | 4 | 4 | No |
| 4 | 5 | 4 | Yes |
| 4 | 9 | 4 | No |
| 5 | 5 | 5 | No |
| 5 | 9 | 5 | No |
| 9 | 9 | 9 | No |

Both Surveillance Epidemiology and End Results (SEER) and CoC allow hospital registries to assign laterality to sites other than those in the paired organ table.

Some central registries may prefer to change all primaries not listed on the paired organ table (Table D) to a laterality of 0 (not a paired organ).

Some central registries may prefer to assign a value based on manual review for cases marked “yes” in the manual review column rather than autocoding a value and marking the case for later review.

| Table C: Primary Site listed on the paired organ table (Table D) | | | |
|---|---------------------|--------------------------------|----------------------|
| Laterality 1 | Laterality 2 | Consolidated Laterality | Manual Review |
| 0 | 0 | 0 | No |
| 0 | 1 | 1 | No |
| 0 | 2 | 2 | No |
| 0 | 3 | 3 | No |
| 0 | 4 | 4 | No |
| 0 | 5 | 5 | No |
| 0 | 9 | 9 | No |
| 1 | 1 | 1 | No |
| 1 | 2 | 1 | Yes |
| 1 | 3 | 1 | No |
| 1 | 4 | 1 | Yes |
| 1 | 5 | 1 | Yes |
| 1 | 9 | 1 | No |
| 2 | 2 | 2 | No |
| 2 | 3 | 2 | No |
| 2 | 4 | 2 | Yes |
| 2 | 5 | 2 | Yes |
| 2 | 9 | 2 | No |
| 3 | 3 | 3 | No |
| 3 | 4 | 3 | Yes |
| 3 | 5 | 5 | Yes |
| 3 | 9 | 3 | No |
| 4 | 4 | 4 | No |
| 4 | 5 | 4 | Yes |
| 4 | 9 | 4 | No |
| 5 | 5 | 5 | No |
| 5 | 9 | 5 | No |
| 9 | 9 | 9 | No |

Some central registries may prefer to assign a value based on manual review for cases marked “yes” in the manual review column rather than autocoding a value and marking the case for later review.

| Table D: SEER Paired Site Table (SEER Manual 2013 pg. 69) | |
|--|---|
| ICD-O-3 Code | Site or Subsite |
| C079 | Parotid gland |
| C080 | Submandibular gland |
| C081 | Sublingual gland |
| C090 | Tonsillar fossa |
| C091 | Tonsillar pillar |
| C098 | Overlapping lesion of tonsil |
| C099 | Tonsil, NOS |
| C300 | Nasal cavity (excluding nasal cartilage, nasal septum) |
| C301 | Middle ear |
| C310 | Maxillary sinus |
| C312 | Frontal sinus |
| C340 | Main bronchus (excluding carina) |
| C341-C349 | Lung |
| C384 | Pleura |
| C400 | Long bones of upper limb, scapula, and associated joints |
| C401 | Short bones of upper limb and associated joints |
| C402 | Long bones of lower limb and associated joints |
| C403 | Short bones of lower limb and associated joints |
| C413 | Rib, clavicle (excluding sternum) |
| C414 | Pelvic bones (excluding sacrum, coccyx, symphysis pubis) |
| C441 | Skin of the eyelid |
| C442 | Skin of the external ear |
| C443 | Skin of other and unspecific parts of the face |
| C445 | Skin of the trunk |
| C446 | Skin of upper limb and shoulder |
| C447 | Skin of the lower limb and hip |
| C471 | Peripheral nerves and autonomic nervous system of upper limb and shoulder |
| C472 | Peripheral nerves and autonomic nervous system of the lower limb and hip |
| C491 | Connective, subcutaneous, and other soft tissues of upper limb and shoulder |
| C492 | Connective, subcutaneous, and other soft tissues of the lower limb and hip |
| C500-C509 | Breast |
| C569 | Ovary |
| C570 | Fallopian tube |
| C620-C629 | Testis |

| | |
|-----------|--|
| C630 | Epididymis |
| C631 | Spermatic cord |
| C649 | Kidney, NOS |
| C659 | Renal pelvis |
| C669 | Ureter |
| C690-C699 | Eye and adnexa |
| C700 | Cerebral meninges, NOS (Effective with cases diagnosed 1/1/2004) |
| C710 | Cerebrum (Effective with cases diagnosed 1/1/2004) |
| C711 | Frontal lobe (Effective with cases diagnosed 1/1/2004) |
| C712 | Temporal lobe (Effective with cases diagnosed 1/1/2004) |
| C713 | Parietal lobe (Effective with cases diagnosed 1/1/2004) |
| C714 | Occipital lobe (Effective with cases diagnosed 1/1/2004) |
| C722 | Olfactory nerve (Effective with cases diagnosed 1/1/2004) |
| C723 | Optic nerve (Effective with cases diagnosed 1/1/2004) |
| C724 | Acoustic nerve (Effective with cases diagnosed 1/1/2004) |
| C725 | Cranial nerve, NOS (Effective with cases diagnosed 1/1/2004) |
| C740-C749 | Adrenal gland |
| C754 | Carotid body |

Grade [440]

Data Item Category: Tumor

| Step 1 | Unknown | Known | Comment |
|---|-----------------------|-----------------------------------|----------------|
| Known over Unknown | 9 | 1-8 | |
| Step 2 | Specific Codes | Comment | |
| Hierarchy | 1-8 | Take higher code over lower code. | |
| Step 3 | Comment | | |
| No manual review required. | | | |
| Input Fields | | | |
| | | | |
| Output Fields | | | |
| | | | |
| Record Selection Criteria | | | |
| | | | |
| Secondary Consolidation | | | |
| | | | |
| Suggested QC Evaluation | | | |
| | | | |
| Issues to Consider | | | |
| Unknown primaries should be coded with an unknown grade. Instructions for Grade were updated for 2014, so registries may want to consider manual review rather than selecting highest code until reporters are familiar with grade coding changes or until grade values can be validated (in addition to passing edits). | | | |

Diagnostic Confirmation [490]

Data Item Category: Tumor

Step 1

Comment

Hierarchy-Diagnostic Confirmation Codes

Priority Order

3. Positive Histology Plus Positive Immunophenotyping and/or positive genetic studies (use only for hematopoietic or lymphoid neoplasms 95903-99923, diagnosed 1/1/2010 and later)
1. Positive Histology
2. Positive Cytology, no positive Histology
4. Positive Lab confirmation, method not specified
5. Positive Lab Test/ Marker study
6. Direct Visualization without microscopic confirmation
7. Radiology and/or other imaging techniques, no microscopic confirmation
8. Clinical Diagnosis Only (other than 5, 6 or 7)
9. Unknown whether or not microscopically confirmed

Step 2

Final Selector

Comment

Value not decided, take value from either first or last submitted source record.

Last Submitted Source Record or First Submitted Source Record.

Input Fields

Diagnostic Confirmation

Class of Case

Output Fields

Record Selection Criteria

Only clean records using all source records available that are a demographic and primary match.

Secondary Consolidation

Suggested QC Evaluation

When unknown code (9) for Diagnostic Confirmation is assigned through automated consolidation process, recommend QC.

Issues to Consider

RX Summ-Primary Site Surgery code could be used for extra rule.

Type of Reporting Source [500]

Data Item Category: Tumor

| Step 1 | Known Over Blank | Comment |
|--------|------------------|---------|
|--------|------------------|---------|

| | | |
|------------------|--|--|
| Known over Blank | | There is no value of 9. Take known value over blank. |
|------------------|--|--|

| Step 2 | Hierarchy-Reporting Source Codes | Comment |
|--------|----------------------------------|---------|
|--------|----------------------------------|---------|

Priority Order

1. Hospital inpatient; Managed health plans with comprehensive, unified medical records
2. Radiation Treatment Centers or Medical Oncology Centers (hospital-affiliated or independent)
8. Other hospital outpatient units/surgery centers
4. Physician's office/private medical practitioner (LMD)
3. Laboratory only (hospital-affiliated or independent)
5. Nursing/convalescent home/hospice
6. Autopsy only
7. Death certificate only

| Step 3 | Comment |
|--------|---------|
|--------|---------|

| Step 4 | Comment |
|--------|---------|
|--------|---------|

Input Fields

Type of Reporting Source

Output Fields

Record Selection Criteria

Record value must not be blank.

Secondary Consolidation

Suggested QC Evaluation

Manual review when value is blank.

Issues to Consider

Histologic Type ICD-O-3 [522]

Data Item Category: Tumor

| Step 1 | Unknown Value | Known Value | Comment |
|---|--|----------------------|---|
| Known over unknown | 9999 | Any other value | Unknown defined as 9999. Eliminate 9999 if any other value. |
| Step 2 | Less Specific | More Specific | Comment |
| More Specific Over Less Specific Codes | 8000, 8010 | Value > 8010 | Eliminate 8000 or 8010 if any other code > 8010. |
| Step 3 | | | Comment |
| Manual Review | | | Manual Review if Class of Case = 99. |
| Step 4 | Class of Case Hierarchy | Comment | |
| See Class of Case Hierarchy for an explanation of Class of Case for Central Registries. | | | |
| Step 5 | Comment | | |
| Manual Review | Consolidate manually if multiple codes remaining. Ex. Different Histology codes from two analytic sources. | | |
| Input Fields | | | |
| Output Fields | | | |
| Record Selection Criteria | | | |
| Secondary Consolidation | | | |
| Suggested QC Evaluation | | | |
| Issues to Consider | | | |

Behavior Code ICD-O-3 [523]

Data Item Category: Tumor

| Step 1 | Comment |
|---------------------------|---|
| Linked with Histology | Consolidate value based on the selected histology code (# 522). |
| Step 2 | Comment |
| | |
| Step 3 | Comment |
| | |
| Step 4 | Comment |
| | |
| Step 5 | Comment |
| | |
| Input Fields | |
| | |
| Output Fields | |
| | |
| Record Selection Criteria | |
| | |
| Secondary Consolidation | |
| | |
| Suggested QC Evaluation | |
| | |
| Issues to Consider | |
| | |

Stage Data Items

Because standard-setter requirements for collection of stage data items will be changing through 2016, data item consolidation of stage data items is a work in progress. As additional information is shared with the registry community, consolidation will be reviewed and updated guidelines for stage consolidation will be released.

Although Collaborative Stage is effective, two options have been identified to address consolidation of Stage data items and are described below:

1) Same Source

In this option, all stage information is selected from one source record based on the record selected for SEER Summary Stage. CS Schema numbers are reviewed to determine the record selection. A two-tiered approach is utilized to consolidate stage as described below:

1. Record Selection:

- (a) One CS Schema number used by all source records—all analytic records as defined by central registries (see [Class of Case discussion](#)) are considered.
- (b) If source records contain multiple CS Schema numbers—select the record(s) that match the CS Schema number that is consistent with the consolidated Primary Site, consolidated Histology and consolidated Behavior including SSF25 Schema Discriminator as necessary.
 - (i) If there is more than one record that matches, or if there is no matching record:
 1. If hematopoietic (Histology > 9582), all analytic records with the same CS schema as the record matching the consolidated Histology are considered.
 2. If not hematopoietic (Histology < 9590), all analytic records with the same CS Schema as the record matching the consolidated Primary Site are considered.

2. Logic for Consolidating Stage Data Items:

- (a) For Urinary sites (C659, C669, C67*, C680 - C681), the record with the highest stage that matches the consolidated Primary Site is selected.
- (b) Select SEER Summary Stage 2000 using SS_2000 Decision Table (below).
 - (i) If more than one record has the same SS_[nnnn] value:
 1. Use the record with the same Consolidated Primary Site and same Consolidated Surgery.
 2. Choose record with Diagnosis Date is closest to the Consolidated Diagnosis Date.
 3. Choose record with earliest receipt date.

3. Advantages/Benefits:

- (a) Decision rules are simple and well documented.
- (b) Review flags handle difficult/problematic cases.
- (c) All cases are consolidated in accordance with consolidation rules eliminating bias from subjective decision making such as personal preference, report source, abstractor, or other subjective factor that might change over time.

- (d) By grouping the CS core items used to derive stage as a single “stage” and carrying along the associated cancer site-specific factors, the result is a group of CS core items, “stage” that derives correctly with all data captured from the best record at time of diagnosis. There is a reduction in CS inter-field edits that may occur as the grouped data have already passed standard EDITS. When selecting one CS core data item value from one source record and a different value from another source record—the data must be used to re-derive the stage(s) and consolidate individual site-specific factors – which may introduce new or unexpected inter-field errors within the CS items when the case goes through final EDITS processing, post-consolidation. “Berry-picking” individual data item values across multiple records may also result in over-staging cases or under-staging cases which may increase the number of cases requiring manual review, unnecessarily.
- (e) There is no need for schema-specific rules unless it is determined that special handling is required outside of normal CS derivation or specific rules for primary site, site group, or histology-driven CS Schema. This greatly reduces maintenance and the chance for update errors when new schema or changes to existing schema occur.
- (f) Manual intervention is only required when specific review flags are set during the automated process. Any case can be identified by review flag should it become necessary to retest criteria for manual review or provide more in-depth analysis of cases meeting criteria for one or more review flags.
- (g) Analytic case reports provide the most reliable data taken from source medical record reports at the time of diagnosis when stage at diagnosis is determined. By only including analytic cases in consolidation decision making, there is no confusion over stage at diagnosis versus stage or restage at recurrence or disease progression.
- (h) When there are new source records linked to the consolidated tumor record, the entire process is repeated taking into account any new data provided. SS and TNM are re-derived should any change be made in any CS data item.
- (i) Any change to consolidation logic and/or decision rules can be implemented and the CS algorithm re-run to re-derive any site-specific schema or all schema and update to the most current rules.
- (j) Relatively low maintenance is required to keep all related tables and decision rules current, including workflow for decision making.

4. Disadvantages/Limitations:

- (a) There may be some reduction in specificity of TNM elements (T1 versus T1a or T1b) and AJCC Stage by selecting based on summary stage differences and class of case.
- (b) Manual intervention is still required to consolidate stage items for cases meeting criteria for review (review flag = manual review)
- (c) Consolidating individual CS Core data elements and SSFs one-by-one may add specificity in some instances but increases likelihood that the consolidated record will fail edits when source record(s) did not, adding another level of manual intervention
- (d) Change management and maintenance could become problematic should major changes to CS be introduced.
- (e) This is a complex process, regardless.

SS_2000 Decision Table:

| SS_2000 1 | SS_2000 2 | Consol SS_2000 |
|------------------|------------------|-----------------------|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 0 | 2 | 2 |
| 0 | 3 | 3 |
| 0 | 4 | 4 |
| 0 | 5 | 5 |
| 0 | 7 | 7 |
| 0 | 8 | 8 |
| 0 | 9 | 0 |
| 1 | 1 | 1 |
| 1 | 2 | 2 |
| 1 | 3 | 3 |
| 1 | 4 | 4 |
| 1 | 5 | 5 |
| 1 | 7 | 7 |
| 1 | 8 | 8 |
| 1 | 9 | 1 |
| 2 | 2 | 2 |
| 2 | 3 | 3 |
| 2 | 4 | 4 |
| 2 | 5 | 5 |
| 2 | 7 | 7 |
| 2 | 8 | 8 |
| 2 | 9 | 2 |
| 3 | 3 | 3 |
| 3 | 4 | 4 |
| 3 | 5 | 3 |
| 3 | 7 | 7 |
| 3 | 8 | 8 |
| 3 | 9 | 3 |
| 4 | 4 | 4 |
| 4 | 5 | 4 |
| 4 | 7 | 7 |
| 4 | 8 | 8 |
| 4 | 9 | 4 |
| 5 | 5 | 5 |
| 5 | 7 | 7 |
| 5 | 8 | 8 |
| 5 | 9 | 5 |
| 7 | 7 | 7 |
| 7 | 8 | 8 |
| 7 | 9 | 7 |
| 8 | 8 | 8 |
| 8 | 9 | 8 |
| 9 | 9 | 9 |

The following data items are selected from the record that provided the consolidated SS_2000:

| | | |
|-------------------------|--------------------|----------------------|
| CS Schema Number | CS SSF11 | AJCC6 M |
| CS Tumor Size | CS SSF12 | AJCC6 M Descriptor |
| CS Extension | CS SSF13 | AJCC6 Stage Group |
| CS Tumor Size Ext Eval | CS SSF14 | AJCC7 TUMOR_CS_LVI |
| CS Lymph Nodes | CS SSF15 | AJCC7 T |
| CS Lymph Nodes Eval | CS SSF16 | AJCC7 T Descriptor |
| Regional Nodes Positive | CS SSF17 | AJCC7 N |
| Regional Nodes Examined | CS SSF18 | AJCC7 N Descriptor |
| CS Mets at Dx | CS SSF19 | AJCC7 M |
| CS Mets Eval | CS SSF20 | AJCC7 M Descriptor |
| CS SSF1 | CS SSF21 | AJCC7 Stage Group |
| CS SSF2 | CS SSF22 | AJCC FLAG |
| CS SSF3 | CS SSF23 | AJCC SS1977 Flag |
| CS SSF4 | CS SSF24 | AJCC SS2000 Flag |
| CS SSF5 | CS SSF25 | TNM Clin T |
| CS SSF6 | CS Version 1st | TNM Clin N |
| CS SSF7 | CS Version Current | TNM Clin M |
| CS SSF8 | CS Version Last | TNM Clin Stage Group |
| CS SSF9 | AJCC6 T | TNM Clin Descriptor |
| CS SSF10 | AJCC6 T Descriptor | TNM Clin Staged By |
| | AJCC6 N | TNM Edition Number |
| | AJCC6 N Descriptor | |

Because this option involves selecting all stage data items from the same source record, the following review flags have been developed so the following scenarios can be identified and manually reviewed after the stage information has been consolidated in the database:

Review Flags:

| SS_[nnnn] Status | |
|------------------|--|
| Code | Description |
| 0 | SS_[nnnn] okay |
| 1 | SS_[nnnn] unknown |
| 2 | Consolidated Stage is UNKNOWN (from an analytic record), but a non-analytic record has a KNOWN SS_[nnnn] |
| 3 | A different (0, 1, (2,3,4,5), 7) SS_[nnnn] is available on an analytic record |
| 4 | Source records have a Summary Stage of (2, 3). Review to determine whether stage should be "4" |
| 5 | More than one CS Schema Number included for this tumor (analytic [0,1,2,6,9]) |
| 6 | A non-analytic record has a different stage than the consolidated SS_[nnnn] |
| 7 | Both Insitu [0] and Invasive [1,2,3,4,5,7] are available on analytic records [(0,1,2,6,9)] |

2) Individual Stage Items Consolidated

In this option, all stage data items including individual Collaborative Stage data elements are consolidated. The derived fields are not consolidated but are re-calculated based on the values in the individual data elements in the Consolidated Record.

- a) Consolidation rules are applied on a data item-by-data item basis whenever multiple values have been reported for the same data item allowing for best value selection.
- b) Schema-specific consolidation rules have been created for each CS input field including defined Site-Specific Factors for the following schema: Colon, Rectum, Breast, Prostate, Lung, HemeRetic, Bladder, Kidney Parenchyma, Thyroid, and Melanoma. Higher volume sites were prioritized.
- c) The application of consolidated rules for stage is a combination of automation and manual intervention. Due to the complexity of extent of disease, manual intervention is needed when there are two different specific values reported from multiple sources. The manual intervention or visual review can be completed as the records are processed or after the records have been consolidated through automation and added to the registry database as a quality assurance activity. For initial implementation of consolidation logic for Collaborative Stage, a conservative approach was taken limiting automated consolidation to codes for which consensus was most likely. Manual intervention is needed for the more complex decisions requiring review of text or CTR review.
- d) Advantages/Benefits:
 - (i) All information from all sources, potentially multiple analytic sources, is considered in determining the Derived Stage values. Example: Analytic Source A reports CS values equating to Regional to Lymph Nodes, then Analytic Source B reports CS values equating to Regional by Direct Extension. The derived SEER Summary Stage would be re-calculated to Regional to Both (Lymph Nodes and Direct Extension) based on the information from both sources.
 - (ii) Provides for accurate reporting of Derived Stage.
 - (iii) Non-specific or less specific CS values are eliminated through automation so manual intervention is only necessary when two specific values are reported by multiple sources.
 - (iv) Reduces manual review of Collaborative Stage data items when consolidation rules are implemented.
 - (v) While the code values are not defined the same across schema, there is a pattern within the Collaborative Stage structure for each data item.
 - (vi) The overall number of records to consolidate manually is reduced if CS items were the only data items requiring manual review and now may be resolved.
 - (vii) Fewer Collaborative Stage data items to review for consolidation.
- e) Disadvantages:
 - (i) Investment of time to draft the consolidation directives for many CS data items.
 - (ii) Consolidation logic needs to be schema-specific due to differences in code structure by schema.

- (iii) Complexity.
- (iv) Revisions necessary when new CS versions are released.

Treatment

The Treatment Data Item Subgroup includes suggested consolidation rules for the following Treatment data items: Surgery Primary Site, Scope Regional Lymph Node Surgery, Surgery Other Regional/Distant Sites, Rad--Regional RX Modality, Chemotherapy, Hormone Therapy, Immunotherapy, Hematologic Transplant and Endocrine Procedures. The goal is to provide central registries with a set of recommended rules for each item that can be automated within any system.

Treatment date items and their corresponding flag codes were not included in the consolidation rules. The Treatment Subgroup felt that the date of treatment and the treatment flag should come from the record from which the treatment code was selected. For example: Hospital A and Hospital B submit an abstract on the same tumor. However, the chemotherapy codes from each facility differ. After running each case through the consolidation rules it is determined that the code from Hospital A will be used. Therefore, the date of treatment and treatment flag from Hospital A will also be used in the consolidated record. Once all of the dates and corresponding date flags for systemic treatment (Chemotherapy, Hormone Therapy, Immunotherapy, Hematologic Transplant and Endocrine Surgery) have been established in the consolidated record, then the Date Systemic Therapy Started will be based on the earliest date full or partial date of the systemic therapy codes.¹ Date of First Course Treatment will be based on the earliest date between Date First Surgical Procedure, Date Systemic Therapy, Date Radiation Therapy, and Date Other Treatment.²

¹ Date Systemic Therapy. If Date Chemotherapy Started, Date Hormone Therapy Started, or Date Immunotherapy Started the earliest date should be used to code Date Systemic Therapy. Priority should be given first to the date with the earliest year, second to the date with the earliest month, and third to the date with the earliest day. If the month and year are the same for multiple dates, a known day takes priority over an unknown or blank day. If the year is the same for multiple dates, a known month takes priority over an unknown or blank month.

² Date of First Course Treatment. If Date First Surgical Procedure, Date Systemic Therapy, Date Radiation Therapy, Date Other Treatment have full dates, the earliest date should be used to code Date Systemic Therapy. Priority should be given first to the date with the earliest year, second to the date with the earliest month, and third to the date with the earliest day. If the month and year are the same for multiple dates, a known day takes priority over an unknown or blank day. If the year is the same for multiple dates, a known month takes priority over an unknown or blank month.

Surgery Primary Site (2003+) [1290]

Data Item Category: Treatment

| Step 1 | | | | Comment |
|---|--|---------------|---------------|---|
| Compare surgery primary site of incoming record to surgery primary site of record in central DB; if match then skip. For cases diagnosed in 2003+. | | | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 99 | 10-90 | |
| Eliminate 99 if any other value. | | | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 00 | 10-90 | |
| Step 4 | More Specific Over Less Specific Codes | Less specific | More specific | Comment |
| | | 90 | 10-80 | |
| | | 80 | 30s-70s | Resection, NOS |
| Step 5 | Higher Code Over Lower Code | Lower code | Higher code | Comment |
| | | 10s | 20s-80 | Specific surgery codes vary by cancer site. This is just a general logic. |
| | | 20s | 30s-80 | |
| | | 30s | 40s-70s | |
| | | 40s | 50s-70s | |
| | | 50s | 60s-70s | |
| | | 60s | 70s | |
| Step 6 Manual Review | | | | |
| Code 00 vs. code 99: required manual review. | | | | |
| There will be no hierarchical order within the same sub-category, because it varies by cancer site. It requires manual review. | | | | |

Output Fields

Record Selection Criteria

1. Surgery Primary Site information will be considered from class of case 10-22, 40-42.
2. Exclude consolidated Primary Site codes (C420, C421, C423, C424, C76x, C809) and/or consolidated Histology codes (9727, 9733, 9741-9742, 9750, 9760-9820, 9826, 9831-9967, 9975-9992). Cases with these codes should have Surgery Primary Site coded to 98.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Scope Regional Lymph Node Surgery (2003+) [1292]

Data Item Category: Treatment

| Step 1 | Known Over Unknown | Unknown Value | Known Value | Comment |
|--------|---|----------------------------------|---------------|---------|
| | | 9 | 1-7 | |
| | | Eliminate 99 if any other value. | | |
| Step 2 | Known Over None | None | Known Value | Comment |
| | | 0 | 1-7 | |
| Step 3 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 1 | 2-7 | |
| | | 2, 4, 5 | 6,7 | |
| | | 3 | 4-7 | |
| Step 4 | Manual Review | | | |
| | Code 0 vs. code 9: required manual review | | | |
| | Code 4 vs. code 5: required review | | | |
| | Code 6 vs. code 7: required review | | | |

Output Fields

Record Selection Criteria

1. Scope Regional Lymph Node Surgery information will be considered from class of case 10-22, 40-42.
2. Exclude consolidated Primary Site in (C420, C421, C423, C424, C70.0–C70.9, C71.0–C71.9, C72.0–C72.9, C75.1–C75.3, C76x, C809) or consolidated Histology in (9727, 9733, 9741-9742, 9750, 9760-9820, 9826, 9831-9967, 9975-9992) or (consolidated Primary Site in (C77.0–C77.9) and consolidated Histology in (M-9590-9726, 9728-9732, 9734-9740, 9750-9762, 9811-9831, 9940, 9948, 9971)). For these site/histologies set Scope of Regional Lymph Node Surgery = 9.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Surgery Other Regional/Distant Sites (2003+) [1294]

Data Item Category: Treatment

| Step 1 | | Comment | | |
|--|---|--|---------------|---------|
| | | Compare surgery other regional/distant sites of incoming record to surgery other regional/distant sites of record in central DB; if match then skip. | | |
| | | For cases diagnosed in 2003+. | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 9 | 1-5 | |
| | | Eliminate 9 if any other value. | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 0 | 1-5 | |
| Step 4 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 1 | 2-5 | |
| | | 2, 3, 4 | 5 | |
| Step 5 | Manual Review | | | |
| | Code 0 vs. code 9: required manual review | | | |
| | Code 2 vs. codes 3, 4: required review | | | |
| | Code 3 vs. codes 2, 4: required review | | | |
| | Code 4 vs. codes 2, 3: required review | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Surgery other regional/distant sites information will be considered from class of case 10-22, 40-42. | | | | |
| Secondary Consolidation | | | | |
| Issues to Consider | | | | |

Rad- Regional RX Modality [1570]

Data Item Category: Treatment

| Step 1 | Date of Diagnosis Prior to 1-01-2003 | | | Comment |
|----------|--|---------------|---------------|--|
| | | | | Exclude codes 80, 85 If date of diagnosis is prior to January 1, 2003 do not consolidate. |
| Step 2 | Incoming Record Match | | | Comment |
| | | | | Accept without review. |
| Step 3 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 99 | 00-62, 98 | Unknown defined as 99. Eliminate 99 if any other value. |
| Step 4 | Known Over None | None | Known Value | Comment |
| | | 00 | 20-98 | Eliminate 00 (no radiation treatment) if any other value (20-98). |
| Step 5.1 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 98 | 20-62 | Other, NOS defined as 98. Eliminate 98 if any other value (20-62). |
| Step 5.2 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 20 | 21-40 | Defined by treatment type EBRT. 20 defined as External Beam, NOS. Eliminate 20 if any value (21-40). |
| Step 5.3 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 41 | 42-43 | Defined by treatment type Radiosurgery. 41 defined as Radiosurgery, NOS. Eliminate 41 if any value (42-43). |
| Step 5.4 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 50 | 51-55 | Defined by treatment type Brachytherapy. 50 defined as Brachytherapy, NOS. Eliminate 50 If any value (51-55). |
| Step 5.5 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |

60 61-62 Defined by treatment type Systemic radiotherapy.

Eliminate 60 if any value (61-62).

| Step 5 | Radiation Treating Facility Over Other Facility | Comment |
|--------|---|---------|
|--------|---|---------|

Treating facility as defined by transmitted NAACCR data item 1550, type of reporting source NAACCR data item 500 is 2, or central registry facility id designated as radiation therapy service provided.
Treating facility value over any other value.

| Step 6 | Manual Review | Comment |
|--------|---------------|---------|
|--------|---------------|---------|

Consolidate manually if multiple types of radiation codes remain. (Ex. Prostate- EBRT or Systemic- Strontium-98).

Output Fields

Record Selection Criteria

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

- 1 The code value set for NAACCR data item 1570 is not current. A proposal to convene a task force for NAACCR Vol II data harmonization was submitted to the NAACCR (S & RD SC), with a priority objective of reviewing radiation treatment data items.
- 2 Consider standard of care by site to prioritize.
- 3 The data item Rad-Boost Rx Modality [3200] should be taken from the same record as Rad-Regional RX Modality [1570]

Chemotherapy [1390]

Data Item Category: Treatment

| Step 1 | | | | Comment |
|--|--|---------------|-----------------------|---|
| Compare chemo of incoming record to chemo of record in central DB; if match then skip. | | | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 99 | 01 ,02, 03, 82, 85-88 | |
| Eliminate 99 if any other value. | | | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 00 | 01 ,02, 03, 82, 85-87 | |
| Eliminate 00 (no chemo treatment) if any other value 01, 02, 03. | | | | |
| Step 4 | Known Over Planned | Planned | Known Value | |
| | | 88 | 01 ,02, 03, 82, 85-87 | |
| Step 5 | More Specific Over Less Specific Codes | Less specific | More Specific | Comment |
| | | 01 | 02, 03 | |
| 01 = Chemo, NOS; 02 = Single agent; 03 = Multiple agent | | | | |
| | | 86 | 82, 85, 87 | For 2003+ cases. |
| Step 6 | Manual Review | | | Comment |
| | Code 00 vs. code 88 or 99: required manual review | | | Need to know which cancers will usually not be administered chemotherapy. |
| | Code 02 vs. code 03: required manual review | | | |
| | Codes 01, 02, 03 vs. codes 82, 85-87: required manual review | | | |

Output Fields

Record Selection Criteria

Chemotherapy information will be considered from class of case 10-22, 40-42.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Hormone Therapy [1400]

Data Item Category: Treatment

| Step 1 | | | | Comment |
|--|---|---------------|---------------|--|
| Compare hormone of incoming record to hormone of record in central DB; if match then skip. | | | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 99 | 01, 82, 85-88 | |
| Eliminate 99 if any other value. | | | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 00 | 01, 82, 85-87 | |
| Step 4 | Known Over Planned | Planned | Known Value | |
| | | 88 | 01, 82, 85-87 | |
| Step 5 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 86 | 82, 85, 87 | For 2003+ cases. |
| Step 6 | Manual Review | | | Comment |
| | Code 00 vs. code 88 or 99: required manual review | | | Need to know which cancers will usually receive hormone. |
| | Code 01 vs. codes 82, 85-87: required manual review | | | |

Output Fields

Record Selection Criteria

Hormone therapy information will be considered from class of case 10-22, 40-42.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Immunotherapy (BRM) [1410]

Data Item Category: Treatment

| Step 1 | | | Comment | |
|--|---|---------------|---------------|---|
| Compare Immunotherapy of incoming record to Immunotherapy of record in central DB; if match then skip. | | | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 99 | 01, 82, 85-88 | |
| Eliminate 99 if any other value. | | | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 00 | 01, 82, 85-87 | |
| Step 4 | Known Over Planned | Planned | Known Value | |
| | | 88 | 01, 82, 85-87 | |
| Step 5 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 86 | 82, 85, 87 | For 2003+ cases. |
| Step 6 | Manual Review | | | Comment |
| | Code 00 vs. code 88 or 99: required manual review | | | Need to know which cancers will usually receive immunotherapy |
| | Code 01 vs. codes 82, 85-87: required manual review | | | |

Output Fields

Record Selection Criteria

BRM information will be considered from class of case 10-22, 40-42.

Secondary Consolidation

Suggested QC Evaluation

Issues to Consider

Code 00 vs. code 99: Reporting hospital usually assigned 00 in BRM field if patient did not receive BRM at their facility which is incorrect per CDC.

Sometimes class of case group 3 cases contain prior tx info especially if pt is being txd at facility for r/c; also for OOS cases especially MDA they submit cases where they provided TX as class of case group 3 cases instead of class of case group 2.

Hematologic Transplant and Endocrine Procedures [3250]

Data Item Category: Treatment

| Step 1 | | | | Comment |
|--|--|---------------|---|------------------|
| Compare Hematologic Transplant and Endocrine Procedures of incoming record to Hematologic Transplant and Endocrine Procedures of record in central DB; if match then skip. | | | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 99 | 10-40, 82, 85-88 | |
| Eliminate 99 if any other value. | | | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 00 | 10-40, 82, 85-88 | |
| Step 4 | Known Over Planned | Planned | Known Value | |
| | | 88 | 10-40, 82, 85-88 | |
| Step 5 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 86 | 11 or 12 over 10; 40 over 10-30; 82, 85, 87 | For 2003+ cases. |
| Step 6 | Manual Review | | | Comment |
| | Code 00 vs. code 88 or 99: required manual review | | | |
| | Code 10-40 vs. codes 82, 85-87: required manual review | | | |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Hematologic Transplant and Endocrine Procedures information will be considered from class of case 10-22, 40-42. | | | | |
| Secondary Consolidation | | | | |
| Issues to Consider | | | | |

Code 00 vs. code 99: Reporting hospital may have assigned 00 in Hematologic Transplant and Endocrine Procedures field if patient did not receive BRM at their facility which is incorrect per CDC.

Other Treatment [1420]

Data Item Category: Treatment

| Step 1 | | | | Comment |
|--|--|----------------------|----------------------|--|
| Compare Other Treatment of incoming record to Other Treatment of record in central DB; if match then skip. | | | | |
| Step 2 | Known Over Unknown | Unknown Value | Known Value | Comment |
| | | 9 | 1-8 | |
| Eliminate 9 if any other value. | | | | |
| Step 3 | Known Over None | None | Known Value | Comment |
| | | 0 | 1-8 | |
| Step 4 | Known Over Planned | Planned | Known Value | |
| | | 8 | 1-7 | |
| Step 5 | More Specific Over Less Specific Codes | Less Specific | More Specific | Comment |
| | | 1 | 2-7 | For 2003+ cases. |
| Step 6 | Manual Review | | | |
| | Code 0 vs. code 8 or 9: required manual review | | | Need to know which cancers will usually receive Hematologic Transplant and Endocrine Procedures. |
| Output Fields | | | | |
| Record Selection Criteria | | | | |
| Other treatment information will be considered from class of case 10-22, 40-42. | | | | |
| Secondary Consolidation | | | | |
| Suggested QC Evaluation | | | | |
| Issues to Consider | | | | |

APPENDIX: List of Rule Definitions for Data Item Consolidation

The following is an alphabetic listing of consolidation rules/commands commonly used in central registries. Consolidation rules may be used alone or in sequence to define an automated consolidation directive for each data item. Sequenced rules can be applied in user-defined order until a consolidation decision is reached or all rules are exhausted.

Consolidate to Combination Code if Specific Conditions Met

Updates the value on the consolidated record if a specific condition is met.

Example: Update If Meets Condition (40, [RX Summ--Transplnt/Endocr = 11] AND [RX Summ--Transplnt/Endocr = 30])

| <u>Abstract ID</u> | <u>RX Summ--Transplnt/Endocr</u> |
|--------------------|----------------------------------|
| 00009453 | 11 [BM Transplant] |
| 00010523 | 99 [Unknown] |
| 00020355 | 30 [Endocrine surgery] |

Selected consolidated value for RX Summ--Transplnt/Endocr: 40 [Combination of 11 and 30]
(Sets the consolidated record value for RX Summ--Transplnt/Endocr to '40' since this field contains '30' on record 00020355 and '11' on record 00009453. Values on other records do not impact consolidation decision.)

Earliest Date:

Selects the earliest chronological date.

Additional consolidation logic may be added to this consolidation rule to include most complete date so that abstracts containing full or more complete dates are given priority over those with partial or blank dates as this may indicate the reporting source with less complete information is estimating dates or does not have as much detail as another reporting source.

Eliminate By Presence:

Eliminates a value from consideration in the consolidated decision if other specified values are present.

The user must specify the conditions under which specific codes are eliminated.

Example: For Spanish/Hispanic Origin, eliminate a value of 7 if any other value 0-6 or 8.

| <u>Abstract ID</u> | <u>Hispanic</u> |
|--------------------|--------------------------|
| 00000001 | 1 [Mexican] |
| 00000002 | 7 [Spanish surname only] |

Selected Value for Spanish/Hispanic Origin: 1 [Mexican]

(Code 7 eliminated from consideration as the consolidated value since the other source record contains a value in the range of 0-6 or 8.)

The NAACCR code definitions for Spanish/Hispanic Origin included in this example are provided below:

- 0 Non-Spanish; non-Hispanic
- 1 Mexican (includes Chicano)
- 2 Puerto Rican
- 3 Cuban
- 4 South or Central American (except Brazil)
- 5 Other specified Spanish/Hispanic origin (includes European; excludes Dominican Republic)
- 6 Spanish, NOS Hispanic, NOS Latino, NOS There is evidence, other than surname or maiden name, that the person is Hispanic, but he/she cannot be assigned to any of the other categories 1-5.
Spanish surname only (Code 7 is ordinarily for central registry use only, hospital registrars may use code 7 if using a list of Hispanic surnames provided by their central registry; otherwise, code 9 'unknown whether Spanish or not' should be used.)
- 7 The only evidence of the person's Hispanic origin is the surname or maiden name and there is no contrary evidence that the person is not Hispanic.
- 8 Dominican Republic

Full Name Over Initial:

Selects a value with two or more characters over a single character value.

Example:

| <u>Abstract ID</u> | <u>First Name</u> |
|--------------------|-------------------|
| 00009453 | R |
| 00020355 | Robert |

Selected Value: Robert

Golden Reporter:

Golden Reporter is considered to be a reporting source reporting high-quality data.

For each field, one or more Golden Reporters can be identified. If an abstract is submitted by a Golden Reporter, then the consolidated field value will be selected from the source abstracts submitted by a

Golden Reporter. Golden Reporters can be pre-defined for each abstract field, and have a specific priority within the set of Golden Reporters for a specific field. If two abstracts are identified as having come from golden reporters, the consolidation value is taken from the abstract whose Golden Reporter ID occurs first in the series of codes in the Golden Reporter Table.

Intent: To enable the user to define, per data item, a facility from which values should always be taken. For example if the registry receives reports from a well-known cancer treatment center and always selects their reported histologic type, then the specified facility could be defined as a Golden Reporter for all morphology-associated data items.

Hierarchy:

A hierarchy can be assigned to a specific data item to select the value for the data item from the abstract that has the highest weight or most significant value in the hierarchy associated with the data item.

Example: Type of Reporting Source

The following priority order is recommended excluding death data items since priority would be given to state death files or other verified sources: 1, 2, 8, 4, 3, 5, 6, and 7

| <u>Abstract ID</u> | <u>Type of Reporting Source</u> |
|--------------------|---------------------------------|
| 00000001 | 3 [Laboratory only] |
| 00000002 | 1 [Hospital Inpatient] |

Selected Value for Type of Reporting Source: 1 [Hospital Inpatient]

The NAACCR code definitions for Type of Reporting Source are provided below:

- 1 Hospital inpatient; Managed health plans with comprehensive, unified medical records
- 2 Radiation Treatment Centers or Medical Oncology Centers (hospital-affiliated or independent)
- 3 Laboratory only (hospital-affiliated or independent)
- 4 Physician's office/private medical practitioner (LMD)
- 5 Nursing/convalescent home/hospice
- 6 Autopsy only
- 7 Death certificate only
- 8 Other hospital outpatient units/surgery centers

Highest Code:

Selects the numerically highest code.

Known Over Unknown:

Eliminates “Unknown” values when a “Known” value is present. Unknown values can be defined in software applications conforming to NAACCR data standard definitions. Unknown values may be item specific, for example the defined unknown value for *Birthplace--State is ZZ*, however most unknown value definitions meet the following defaults:

| | |
|---------------------|---------------------------------------|
| Date fields..... | Blank; portions of dates can be blank |
| Numeric fields..... | Blank, 9, 99, 999,.... |
| Text fields..... | Blank, Unknown, Unk, NR |

Example:

| | |
|---------------------|---------------------------------------|
| <u>Abstract ID:</u> | <u><i>RX Summ--Surg Prim Site</i></u> |
| 00009453 | 99 [unknown] |
| 00020355 | 50 [known valid surgery code] |

Selected consolidated value for *RX Summ--Surg Prim Site* is 50 since 99 is considered unknown.

Latest Date:

Selects the latest chronological date.

Lowest Code:

Selects the numerically lowest code.

Manual Review:

Requires manual consolidation if source values differ.

Manual Review if Meets Condition:

Manual review is necessary if a specified condition is present.

Example: For Primary Site, if Class of Case on any abstract = 99, manual review needed.

Most Complete Address by Date:

Address fields must be grouped together for consolidation in case multiple reporting sources report a different address. If a different address is reported and Addr at Dx--No & Street is selected from Facility A, but Addr at Dx -- City is selected from Facility B, not only will the consolidated address at Dx data items be mismatched, an issue will result in geocoding as well.

It is also important to acknowledge that the consolidation of address is specific to a point in time. For Address at Dx, Date of 1st Contact or Date of Diagnosis should be considered so the earliest most complete address at the time of initial diagnosis is selected for consolidation. For Current Address data items, Date of Last Contact should be considered so the latest most complete address is selected for consolidation.

Most Frequent:

Selects the value that occurs most often.

Example: The consolidation rule for *Social Security Number* includes Most Frequent.

| <u>Abstract ID</u> | <u>Sequence Number</u> | <u>SSN</u> | <u>Date of Last Contact</u> |
|--------------------|------------------------|------------|-----------------------------|
| 00000012 | 01 | 123456789 | 20091214 |
| 00000301 | 01 | 987654321 | 20110115 |
| 00001010 | 02 | 123456789 | 20121124 |

Selected consolidated value for *Social Security Number*: 123456789

- **Most Frequent(Collection)**

Allows the selection of most frequent occurrence to be modified by a specified grouping, so that equal values within the same group are counted only once. This instruction is used to reduce the possibility that multiple reports from the same reporting facility will add bias to the consolidation decision.

Example: For the consolidation rule Most Frequent (Reporting Facility)

A patient with two primary cancers is reported by three facilities with one facility reporting both primaries. Social Security Number and associated Reporting Facility values are:

| <u>Sequence Number</u> | <u>Reporting Facility</u> | <u>SSN</u> |
|------------------------|---------------------------|------------|
| 01 | 1111111111 | 123456789 |
| 01 | 2222222222 | 987654321 |
| 02 | 2222222222 | 987654321 |
| 02 | 3333333333 | 123456789 |

Selected consolidated value for Social Security Number: 123456789

(The Social Security Number value of 987654321 will be counted only once since it was reported multiple times by the same Reporting Facility.)

Most Recent:

Selects the consolidated value based on most recent date. While any date field could be used, Date of Last Contact is the most common and is often the default sequencing field.

Most Recent may be part of the consolidation rule used with the Physician--Follow-Up field since the most recent physician based on Date of Last Contact should be selected for consolidation.

Example: The consolidation rule for Physician--Follow-Up includes Most Recent (Date Last Contact)

| <u>Abstract ID</u> | <u>Physician-Follow-Up</u> | <u>Date of Last Contact</u> |
|--------------------|----------------------------|-----------------------------|
| 0000001 | 00011111 | 20121214 |
| 00000301 | 00055555 | 20130115 |
| 00001010 | 00004444 | 20131124 |

Selected consolidated value for Physician--Follow-Up: 00004444

Same Source:

Selects the value from the abstract used to select the designated data item. For example, If SameSourceAs(PSite) is used to consolidate Laterality then the Laterality value selected for the consolidated record will be selected from the abstract from which the consolidated primary site code was selected.