SEER*DMS Auto-Consolidation Work Group
Teleconference Summary
April 18, 2019
3:00 to 4:00 p.m. EDT

Representatives from the NCI, IMS, the Scientific Consulting Group, Inc. (SCG), and 10 cancer registries participated in the SEER*DMS Auto-Consolidation Workgroup (WG) conference call on April 19, 2019. Participants included:

REGISTRIES:
Alaska
California Central
Detroit
Idaho
Iowa
Minnesota
New Jersey
New Mexico
New York
Seattle

NCI: Peggy Adamo, Lois Dickie, Serban Negoita

IMS: Suzanne Adams, Linda Coyle, Fabian Depry

SCG: Carolyn Fisher, rapporteur

Westat: Laura Lourenco

Action Items

Participants agreed to the following action items:

- IMS developers will begin to test the Lymphovascular Invasion **(**LVI) auto-consolidation logic.
- Linda agreed to create a Squish issue on a known over the unknown rule for the Mets at DX fields (i.e., data items 1112–1117).
- IMS will provide information on the frequency that Site-Specific Data Items (SSDIs) are being updated via the auto-consolidation rules.
- Linda agreed to develop a query to search for the number of SSDIs that are coded unknown, but have a known value in the record.
- IMS will provide data searches with examples to SEER*DMS registries to help the WG define metrics.
- The NCI will conduct a review of the NAACCR metafiles data to identify any existing edits for conflicts in derived and direct coding rules for the Summary Stage 2018 field.
- Linda agreed to check with Jennifer Stevens about SEER*Edits updates prior to developing rules for the Sentinel Lymph Nodes data items.
- A subgroup of WG members will meet to discuss the next data items to auto-consolidate.

IMS Updates

Linda updated the WG on IMS activities including running edits on import files, LVI Logic, auto-consolidation changes, Type of Reporting Source rule changes, new auto-consolidation fields, and the system task for testing proposed rules. The WG needs to makes decisions regarding rules for known over unknowns.

Running Edits on Import Files for Source Validation

IMS will add a simple, streamlined system task that SEER*DMS registries can use to obtain statistics about failed edits on source files. A workflow will not be developed at this time, but the information might be useful to the WG Source Record Validation Subgroup.

Auto-Consolidation Changes

IMS developers are working on the auto-consolidation module to support logic that considers all linked records in the rules, not just the incoming record. Currently, auto-consolidation is based on a comparison of an incoming record with the already consolidated data.

LVI Auto-Consolidation Logic

Suzanne reported that the LVI logic is being finalized and noted the remaining issues with the neoadjuvant therapy data item. In the *SEER Program Coding and Staging Manual 2018* (SEER Manual), the only LVI cases that do not fit the newly defined logic are cases with priority codes 0 and 9. The options for handling these cases include: (1) manual review of cases coded 0 versus those coded 9 in their entirety; (2) selection and review of cases that had neoadjuvant therapy based on a defined criteria; or (3) a combination of options 1 and 2.

Discussion

Ms. Lisa Pareti at the Louisiana registry representative has expressed concerns about the neoadjuvant therapy data item but was not present for this call. Linda indicated that she would like input from Ms. Pareti.

A NAACCR group is working to develop a neoadjuvant data item, which will capture whether neoadjuvant therapy was given. The implementation date likely will not occur until 2021 or later. Participants agreed that manually reviewing these cases until then might be the best option. The alternative option of using sequence number variables might create data quality problems. The WG's LVI rules are conservative, would be implemented across registries, and manually consolidated fields could be reviewed and automated at a later date.

Participants discussed the possibility of developing a tool to determine whether neoadjuvant therapy had been given based on the record/file dates from the pathology reports. The development of such a tool could be more complicated than anticipated and delay implementation of LVI auto-consolidation rules. Treatment and sequence number data have been reviewed and reasons for conflicting data have been identified in some cases.

Participants discussed benign/ borderline brain and CNS tumor brain default codes. The SEER Manual and the *Standards for Oncology Registry Entry (STORE)* coding logic agree that the default code for cases diagnosed in 2018 or later should be set to 8. The LVI code of 8 is used when LVI is not applicable for that site. Coding logic from different sources used in the years prior to 2018 can be problematic. The 2016 and 2017 SEER Manuals instruct registrars to have a default LVI code of 0 for borderline/brain and CNS cases, and most registries have been following this instruction. The fact that guidance prior to 2014 is unavailable further complicates the matter. The WG could decide to auto-consolidate 2018 and later cases or also define a consistent way to code pre-2018 cases. The ways that these codes (0 vs 9 vs 8) affect research should be considered. The Iowa registry was using a default code of 0 for benign borderline

brain prior to 2018 and prefers to continue this practice. The California, New Mexico, New Jersey, and Minnesota registries use code 0 as per the 2016 and 2017 guidelines. At the Detroit registry, the SEER Abstracting Tool (SEER*Abs) defaults to a code of 8 for benign borderline brain cases diagnosed in 2018 or later and the registry is coding benign borderline brain as 8 when they are diagnosed prior to 2018. The Detroit registry may require a conversion for cases coded to 8 that should be 0. Linda asked about a standard and whether registries should choose how to code this data field. Participants were not clear about the Commission on Cancer (CoC) hospital guidelines for coding benign/borderline brain and CNS cases.

System Task for Testing Proposed Rules

IMS has been writing, testing, and formalizing a system task, which IMS expects to release to the registries May 2019. Registries will be able to use this system task to test the proposed auto-consolidation rules.

Type of Reporting Source

IMS is continuing to work with registries to adjust old data to fit the new rules. These efforts are ongoing.

Known Over Unknown Rules

Linda clarified that there are currently in production known over unknown auto-consolidation rules for SSDIs.

The known over unknown rules for Mets at DX fields are pending. Questions still need to be answered regarding Primary Payer and Marital Status at diagnosis. Linda proposed meeting with Bobbi Jo Matt and Frances Ross for a discussion about these data items. She also asked registries to provide comments about this issue via Squish.

Discussion

Participants asked about ways to determine which values had been polished and updated. IMS can write a query to collect information on the frequency of SSDI updates via auto-consolidation and on the number of SSDIs that are unknown but have a known value in the record.

Defining Metrics

Participants discussed metrics to evaluate auto-consolidation processes and outcomes. Metrics to measure how often the auto-consolidation logic f is being applied would be a start. Neither the SEER*DMS user logs nor the audit logs provide information about polisher or system changes. Linda explained that retrieving this information will not be a simple task. No changes are made to incoming records that are correct as is and do not need to be auto-consolidated. One metric IMS has used is the number of times that manual consolidation and the auto-consolidation rules disagree.

The auto-consolidation logic only is applicable when a source record is added to the existing CTC, which could be a metric to evaluate the system changes. Metrics are needed to identify when a CTC is built from one abstract and then has values changed because of a second abstract. The metric would be the total number of changes divided by the total number of runs. Linda pointed out that only the changes, not the interim iterations, are stored in SEER*DMS. The next step will be for IMS to provide data searches with examples to SEER*DMS registries.

Next Data Fields for Auto-Consolidation

The WG reviewed the spreadsheet of SEER Phase 1 fields needing auto-consolidation and the potential groupings of those fields.

Discussion

Linda noted that logic for Primary Payer at DX had been developed for the New York registry and could be leveraged.

In response to participant queries, Serban confirmed that the instructions for the Summary Stage 2018 field regarding EOD had not been sent to registries and likely will be reviewed in 2020. The NCI could review the NAACCR metafiles for existing edits to conflicts in derived and direct coding rules for Summary Stage 2018. Developing rules for EOD Primary Tumor, EOD Regional Nodes, and EOD Mets can be considered, but is not a priority.

Participants identified the following tentative data item groupings:

- Regional Nodes Positive, Regional Nodes Examined, EOD Regional Nodes, and date of Regional Lymph Node Dissection.
- Sentinel Lymph Nodes Examined and Sentinel Lymph Nodes Positive, Date of Sentinel Lymph Node Biopsy.
- EOD Primary Tumor, EOD Regional Nodes, and EOD Mets.
- EOD Mets and Mets at DX Brain.

Since Sentinel Lymph Nodes are only required for breast and melanoma cancer sites, participants asked whether SEER*Edits permitted those fields to remain unfilled for other sites. Jennifer Stevens (IMS) might be able to answer this question.

Linda and Bobbi decided to convene an administrative meeting of the WG Chairs to discuss the next data items to auto-consolidate. WG members should send suggestions regarding these data items. Participants suggested developing rules for the Mets at DX field series.

Upcoming Auto-Consolidation Work Group Calls

The next Auto-Consolidation WG call is scheduled for May 16, 2019, and will focus on source code validation.