The Surveillance, Epidemiology, and End Results Data Management System (SEER*DMS)
Change Control Board (CCB) Committee
Teleconference Summary
November 9, 2017
3:00 p.m. to 4:30 p.m. EST

Representatives from NCI, IMS, the Scientific Consulting Group, Inc. (SCG), and 11 SEER registries participated in the SEER*DMS CCB Committee conference call on November 9, 2017. Participating registries included:

Alaska
Connecticut
Georgia
Detroit
Hawaii
Iowa
Minnesota
New Jersey
New York
Utah
Greater California

NCI: Marina Matatova, Alyssa Wang

IMS: Linda Coyle

SCG: Kathy Brown-Huamani, rapporteur

ACTION ITEMS

- Kevin Ward agreed to integrate comments and questions on the Claims' Working Group (WG) document into a revised document.
- The WG chairs should review their WG member list on the portal because only those on the list will get notifications. Chairs can use the New Features button at the bottom of the page to add members.
- Marina Matatova agreed to distribute an email with NCI updates.
- Linda Coyle agreed to distribute the link to the WG member list, which is on the portal.
- Representatives from the New Mexico and Kentucky registries will review the Claims WG work plan and provide comments.
- Georgia, Hawaii, Connecticut, and New York registries will review the MU2 work plan and provide comments.
- The Iowa registry representative agreed to submit a Squish ticket concerning the slowing down and crashing of their test servers due to the download of large data files.
- A Squish issue will be posted requesting feedback from registries regarding the SEER Join Point methodology.
- Registries should propose agenda topics for the January 2018 CCB meeting.

REVIEW OF 1-YEAR PLANS FOR EACH WORK GROUP

Since the July 2017 face-to-face meeting, the WG chairs, NCI, and IMS held administrative meetings to draft documents defining the WG objectives. These documents defined the 2017–2018 projects for each WG and described short- (next 10–12 months) and long- (beyond 12 months) term goals and objectives. Draft documents for each WG were posted in Squish for comment. During the November 9, 2017 call, WG chairs presented their documents and participants discussed them.

Claims Workgroup (see Squish 5686) **Kevin Ward**

The Claims WG draft document contained short-, medium-, and long-term goals. A separate comments document will be drafted. The WG plans to complete all short-term and most medium-term goals within 1 year. Short-term goals were organized into four categories: onboarding, codes and formularies, workflow, and quality control.

Short-Term Goals

Onboarding

- Assess the current landscape, gain understanding of the status of local agreements in each SEER area, to allow integration of Unlimited data. The Claims WG documented the challenges experienced by registries receiving Unlimited data.
- Develop a checklist for registries to use to evaluate the data that they receive. The WG developed some solutions to those challenges that can be employed by other registries, including a checklist that will be integrated into SEER*DMS to walk registries through an initial evaluation of Claims data.
- Discuss the impact of different sources and the timelines for receiving claims data from those sources. The WG reviewed data sources, processes, and timelines for receiving data from different sources, which also differs across registries. Some data are received quarterly, some only annually.
- Implement viewer to track incoming claims by source/provider. The WG is working with IMS to develop a Dashboard/Viewer that will allow registries to track claims data over time and identify anomalies. Unusual patterns in the receipt of claims data (e.g., low numbers) can trigger a conversation with Unlimited to determine why those patterns are occurring.

Codes and Formularies

- Explore what additional data can be accurately gleaned from ICD-10 codes available in 2015 (metastases, progression, second primary, etc.). Most drug code lists and formularies have been developed by NCI and will reside on the NCI web site. Overall, ICD-10 codes will have greater specificity than the ICD-9 codes regarding tumor characteristics and treatment.
- Incorporate drug agent lists for all oncologics (chemo-, hormone, and immunotherapy).
- Develop a comprehensive list of radiation therapy codes.
- Develop methods to address agents with reference dates (e.g., considered chemotherapy at one point and immunotherapy at another). Some treatment codes have been reclassified over time (e.g., some previously considered chemotherapy is now considered immunotherapy). The drug coding system will allow registries to define treatments as they were defined at the point in time that treatment was received.

Workflow

- Finalize decisions on claims matching (patient only vs. CTC). Registries will need to match claims to CTCs as they are received or match them at the patient level according to the time period when the claims occurred.
    - If patient only, timing rules still would allow treatment augmentation.
    - If CTC, need to optimize MPH assessment rules for matching claims to CTCs. MPH rules have been relaxed for Claims data, which often do not include histology or laterality information. These rules might have to relax further if the WG and NCI determine that claims data should be matched to CTCs.
- Finalize DMS workflow for the processing of prospective claims.
    - The Claims WG will need to determine how to process prospective claims data, which could require the generation of AFLs and a decision about how to use those AFLs. Currently, all claims data received by registries are retrospective, so CTCs already exist for the patients.

- With prospective claims, auto consolidation will be necessary and auto consolidation processes will need to be developed. The process will need to determine which claims meet the criteria for auto consolidation. Claims will need to provide adequate information for linkage to a Patient Set.

Quality Control

- Create standard queries within SEER*DMS to check the quality of claims data.
- Evaluate claims data across registries to determine exactly what claims data to add to existing SEER data sets. The objective will be to use retrospective claims data to evaluate the degree to which claims data identify new cases and cases in need of follow back at each registry.

Medium-Term Goals

- Develop processes for follow back on incoming claims that do not match a Patient Set.
- Develop rules for automation:
  - Define all fields for automation.
  - Identify first versus later course therapy.
  - Capture agents in discreet fields.
- Establish processes for conducting quality control on automated data:
  - Develop test plans; a suggestion is to include epidemiologists and researchers to evaluate the usability of the data for research.
- Improve data access.
  - Develop data query capability for non-Structured Query Language users.
  - Facilitate internal access to the research file of claims data for analyses.

Long-Term Goals

- Expand implementation to other oncology practices and vendors, which would involve the mandatory reporting of administrative claims for oncology practices that are not submitting data.
- Develop methods for research use of the data (data documentation, data query, ability, to export, export format).
- Develop rules for automation for research use:
  - Complete decisions regarding population of additional discrete data fields that are relevant, such as comorbidity, change in agents/regimens, completion, and early termination of therapy.

Discussion

Kevin requested that questions be submitted to him after the meeting. The Seattle registry is not receiving data from Unlimited at this point, but is receiving some claims data from a large health management organization (HMO) in the area. Kevin recommended that the Seattle registry join the Claims WG. Information about the Seattle registry's processes for using their HMO data would be useful for other registries. The long-term goal is to expand beyond Unlimited data to include claims received directly from large insurance providers or other organizations within the catchment areas of registries. IMS has developed a flexible system within SEER*DMS; there is an underlying pre-record table that can read data from many different formats. Kevin commented that IMS soon would receive pharmacy data that are applicable to all registries, which might also be handled by Claims WG.

Marina asked whether pharmacy data collection is a separate goal from the expansion of data collection to other sources of claims data. Kevin indicated that the longer-term goal was to expand data collection to any medical record from sources other than Unlimited.

Registries provided their feedback regarding the Claims WG work plan. The Utah registry representative expressed support for the plan. The New Jersey registry already provided comments on the plan and California's comments previously were integrated. The Detroit representative indicated that this registry has no comments on the plan.

The Kentucky registry is not represented on the Claims WG. Frances Ross recommended that Dean Vaughn from the Kentucky registry join the Claims WG.

Auto-Consolidation Workgroup (see Squish 5750) **Frances Ross**

The Auto-consolidation WG outlined their 1-year goals. The work plan for this WG includes the development of a process for identifying tumor-related data fields that can be auto-consolidated. Consolidation is defined as the process of selecting the best information when two or more source records (linked at patient and cancer level) contain the same data items, but with different values.

The WG selected two data items to test the auto-consolidation process over the next year. Steps in the testing process will include:

- Referencing the SEER Coding Manual and other sources for coding instructions applicable to the consolidation of the data item.
- Identifying all data items that might be used in the decision making process for consolidating each data item.
- Developing logic rules that would identify the record with the best data value for the consolidated (CTC) record for each data item.
- Designing a test to determine how well the new logic rules identify the same data value as that stored in the SEER CTC, which was manually consolidated (e.g., match values on CTCs to values in source records).
- Implementing the test:
  - Using existing data in SEER*DMS databases to validate the testing protocol. One record per CTC/source record would be included in the output file.
- Analyzing the test results to determine necessary revisions.

Discussion

There were no additional comments on the Auto-Consolidation work plan. The California registry comments have been incorporated in the plan.

MU2 Workgroup (see Squish 5748) **Brent Mumphrey**

The MU2 WG has determined that CDAs can be imported into the SEER*DMS database. The 1-year goals for the WG include: (1) data analysis, (2) use case development, and (3) workflow and interface development.

For analysis, CDAs will be grouped into those matching or not matching an existing Patient Set. For CDAs not matching existing Patient Sets, the WG will examine how many verifiable new cases were found. The MU2 WG will determine how many new cases are found among the CDAs matching Patient Sets. Among CDAs matching an existing CTC, the WG will examine the number that provides new treatment information. For Patient Sets with multiple CDAs per CTC, the MU2 WG will determine whether the CDAs are true duplicates.

For workflow development, the MU2 WG plans to develop a workflow in SEER*DMS for each use case presented. Current use cases include the updating of follow-up information, addition of new treatment information, identification of metastases, recurrence, and comorbidities as well as rapid case ascertainment.

Concerning interface development, the MU2 WG will analyze the current methods for viewing the CDA data in SEER*DMS and determine what interface is best for viewing the data. The WG also will seek information on fields that the users need to see and those that can be removed.

Within the 1-year timeline, the MU2 WG will complete data analysis. The January 2018 to April 2018 timeframe relates to workflow development with use cases of follow up information and case finding, as well as interface analysis development. For May 2018 and beyond, the MU2 WG will engage in use case workflow development for additional treatment, metastases, recurrence, and comorbidities as well as rapid case ascertainment.

Discussion

The Connecticut registry noted that they were contacted by a physician practice interested in testing MU2 data. The Iowa registry tested 867 CDAs from 2015 (and prior years) and identified 36 new cases from dermatology laboratories. Iowa registry representative indicated that registry staff would analyze additional MU2 data. The New Jersey registry was experiencing long download times for large files. The download of case reports from one practice took 42 hours. The problem is large CDAs rather the number of records.

The California registry submitted comments in Squish. Ben Wormeli wants to join the MU2 WG. All current WG memberships are posted on the portal.

SEER*DMS CHANGES

Worklist Filters (see Squish 5705)

Worklist filters were designed when a single record was the focus of most worklist tasks. Tasks related to consolidation and visual editing are now Patient Set based. IMS has made changes at the backend in the interface so the filters work within the new workflow. The following modifications were proposed:

- Enhance the record filters so that they work for any record that needs to be consolidated in a Consolidate task.
- Change the records and CTC filters that include the fields (1) event date (diagnosis date for a CTC), (2) site, (3) data type, and (4) region. A record requires work if it has not yet been consolidated. A CTC requires work if it needs visual editing or is failing edits.
- Add filters for CTCs that require work.

Discussion

Registries submitted comments regarding these proposed changes. Specific comments indicated that additional clarification on the date fields are needed. Linda asked the registries whether they agreed that the worklist filters should be changed so that they work for each pending record. The Iowa registry staff are unable to find tasks using their current filters because they produce inaccurate information on the worklist. Complicating the filter problem is the fact that different people are assigned to different tasks. Changes will enable users to find records, but they will not be displayed immediately. Modifications

would immediately change how the filters work, however. Linda agreed to discuss the problem of worklist filters and multiple assignments to a task on a future call.

The Iowa registry also appears to experience problems with tasks being re-rerouted to the wrong queue so that editors see items on a worklist that do not belong there. Either scripts or people can assign tasks. To evaluate the underlying cause of the problem, Linda suggested creating an autolog of task assignments that note who was assigned the task, system, polisher, and person who re-routed the task. Seattle, Louisiana, Utah, New York, and Minnesota agreed that the autolog of tasks assignments would be useful.

Imports of Electronic Images (see Squish 5568) **Brent Mumphrey**

The Louisiana registry is receiving imaging reports from AIM CNS tumor software. Louisiana wishes to include imaging reports in SEER*DMS and create a separate import for these records. Linda recommended that separate imports be flagged for the type of record. Distinguishing imaging reports from pathology reports is necessary. The type of workflow required for an imaging report is unclear. The Louisiana registry staff will need to examine their SEER data and decide how to integrate imaging data.

Discussion

Valerie Otto said that the Utah Cancer Registry is receiving imaging data that have not been used yet. Hawaii and Georgia registries expect to receive imaging data and were interested in an import for the electronic imaging reports. The Seattle registry is attempting to set up imaging data feeds although the registry likely will not receive imaging reports in the near future.

NCI UPDATES **Linda Coyle**

Highlights and Discussion

- The effort to support North American Association of Central Cancer Registries (NAACCR) XML as an import for registries using SEER*Abs is ongoing. This work is set for completion by 2020. The New Mexico registry already is using NAACCR XML for production work. The New York registry representative asked if there are edits in XML or if the files are being translated back into a flat file for editing. Linda replied that flat file translation is not relevant in SEER*DMS because data will be imported and loaded into record table. The Iowa registry representative asked if the XML import standard is in its final version. The standard has a well-defined production version, but there may be new versions.
- Updates to the SEER*DMS Data Quality Profile completeness estimate to the new SEER Join Point methodology is scheduled for the 2018 and 2019 submissions. The goal for the 2019 submission has not been defined but might involve discussions with NCI and the Minnesota and New York registries. Alternatively, a linear regression approach might be used to set this goal.
- IMS added fields to registry User Accounts so that supervisors can manage CCB WG lists. Users will control who receives CCB notifications and will know who the members of the WG are. A tab has been added to the user accounts to allow managers to assign members to the CCB WG. IMS can set flags for people already members of the CCB WG as defined by managers. From this point forward, WG managers will need to manage in the user account themselves. If managers deactivate an account, that person will no longer receive WG notifications.

Other Topics

Topics for this month or a future meeting should be sent to Linda or Marina.

Next Meeting

Marina requested that registries submit topics for the January 2018 meeting agenda. She will present updates on NCI's usability efforts and project management-related topics. Alyssa Wang will update the WG regarding current communication strategies. Feedback from the registries is needed to improve these strategies.