The Surveillance, Epidemiology, and End Results Data Management System (SEER*DMS)
Meaningful Use (MU2) Work Group
Teleconference Summary
July 17, 2018
1:00 p.m. to 2:00 p.m. EDT

Representatives from NCI, IMS, the Scientific Consulting Group, Inc. (SCG), and nine SEER registries participated in the SEER*DMS MU2 work group conference call on July 17, 2018. Participants included:

REGISTRIES:
California Central
Connecticut
Iowa
Louisiana (Brent Mumphrey, Chair)
Minnesota
New Jersey
New York
Seattle
Utah

NCI: Paul Fearn, Andrew Grothen, Marina Matatova, Kai Wong

IMS: David Angelaszek, Linda Coyle, Jennifer Stevens

Westat: Laura Lourenco

The Centers for Disease Control and Prevention (CDC): Wendy Blumenthal

SCG: Kathryn Brown-Huamani, rapporteur

Action Items

- NCI, IMS, and CDC participants agreed to schedule an administrative meeting to discuss approaches for handling duplicates.
- Brent agreed to create a Squish issue so the WG can review his draft workflow for casefinding.
- IMS will check on the type of CDA document used by radiation oncology facilities.
- Wendy agreed to examine the radiation CDAs to see what content is included. Marina agreed to send the CDAs to Wendy.
- Paul agreed to ask Elekta about providing guidance for radiation oncology facilities to share CDAs.
- The New York registry representative participating in the call agreed to share notes with the WG lead from that registry and ask her about her interest in participating in a test of methods for receiving radiation oncology data.
- Wendy agreed to double check to ensure that Mosaic was not recently certified for MU cancer.
- David agreed to set up data sets to test duplicate CDAs using various criteria for determining duplicates.
- NCI and IMS will work on an approach for handling duplicates.
- Registries with MU data should evaluate those data, including comparing vendors, to determine how the amount and quality of CDA data varies across registries.
- IMS will create a Squish issue for casefinding to allow WG members to provide additional feedback and add workflow steps. Participants agreed to discuss this feedback during the next MU2 WG call.
- Marina, Linda, and Brent agreed to work together on the creation of the Radiation Oncology subgroup.
- IMS will post a Squish issue outlining the goals of the radiation oncology data testing subgroup. Registries can respond with questions and to express interest in participating in the testing.
- Marina and Linda agreed to meet with Brent to discuss the MU2 WG presentation at the fall face-to-face meeting. The presentation should highlight WG activities and accomplishments to date as well as findings from analyses at different registries.

Radiation Data **Paul Fearn**

NCI is seeking ways to obtain more radiation oncology data through the two major vendors of those data, Varian and Elekta (Mosaic software), that cover the majority of the market. NCI SEER is working with Varian developers to create a SQL query that pulls data elements of interest in a standard way so that IMS can receive those data. These processes were tested in New York at the registry and a large reporting hospital. NCI SEER is working on a similar project with Elekta. The SQL approach likely will not work with Elekta so they are developing another approach using the CDA document. A test case for the CDA document approach was distributed to WG members so that they can help determine whether this method is viable for sending data from radiation oncology facilities to registries. If the method is viable, WG members will need to determine which facilities have radiation treatment information in CDAs as well as the capability to send data using the method described.

Linda would like answers to the following questions:

1. What mechanism can registries use to put data in the proper format?

2. Are there enough registries that include required details in CDA documents?

3. Does anyone want to try a new SQL query? A query has been developed that can be made available for testing by registries.

Discussion

Wendy asked about the type of CDA document that would be used in obtaining Elekta radiation data. CDA is the structure but can use the MU or Continuity of Care type of document. The MU document would include more detailed cancer treatment information.

The University of Washington developed a SQL query that works with the Mosaic database. The Seattle registry has access to this query and is willing to share it with interested registries. Louisiana has reporting facilities that use Elekta and is willing help test methods for receiving CDAs from radiation oncology facilities. Brent would like to know the level of expertise required for radiation facilities to run queries. These facilities might need guidance from Elekta. Paul can ask Elekta about providing guidance.

Marina asked registry participants to help identify radiation oncology facilities that could provide MU documents and test queries. She asked other participants if they might be willing to test the feasibility of CDA reporting by radiation oncology facilities at their registries. The New Jersey representative expressed interest in testing at his registry. The Connecticut representative also expressed interest in testing because the registry is seeking new ways to obtain information on cancer treatment. This registry, however, does not have access to Unlimited data. The New York registry might be interested in participating in testing, but the lead was not on this call. The representative participating in the call agreed to ask the WG lead from the New York registry about testing.

Wendy was willing to contact non-SEER (National Program of Cancer Registries) registries about testing ways to receive information from radiation oncology facilities. She agreed to discuss this option with NCI staff offline.

Mosaic is not certified for MU data but might be encouraged to get certified if testing shows that registries can receive Mosaic data.

Data Analysis **Brent Mumphrey**

Utah completed an analysis of MU2 data. The New Jersey registry also completed the analysis and sent results to Linda. The Iowa registry has started that analysis and expects to submit results in approximately 1 week. New Jersey performed the analysis for medical oncology groups. The WG representative for that registry plans to conduct another analysis of data from a single dermatology facility and will send results to Linda when the analysis is complete.

Duplicate CDAs **Brent Mumphrey**

CDC's Implementation Guide for Ambulatory Healthcare Provider Reporting to Central Cancer Registries indicates that CDAs should be cumulative but this is not always the case with MU2 data. For MU3 data, duplicates can be handled by creating an updated file document. Electronic Mapping, Reporting, and Coding Plus consolidate data when a report has exact matches on provider, patient, and tumor type. The CDC has consolidation rules for automation in the Central Cancer Registry database software.

Duplication of CDAs varies by vendor. IMS will explore ways to determine whether a report is a duplicate with appended information.

David began testing to see if duplicates could be identified via specific fields in the CDA. These fields included radiation notes (text), progress notes (text), procedures, and medications. If two submit dates appeared for the same patient, the algorithm would delete the earlier submit date. David tested this algorithm using CDAs from the Utah registry on a dev server. Once duplicates were removed, the number of CDAs decreased. David ran analyses to determine why some duplicates were not flagged as duplicates. He presented a table listing the reasons why some duplicates were not flagged for different fields. All radiation notes fields were "None" or blank. The other fields have several cases. The progress notes field revealed that minor changes in the notes created the appearance of duplicates. In many cases, the minor changes were the result of physicians going back and editing the notes. David was continuing to examine the procedures and medications fields to determine why they were not flagged as duplicates.

Linda recommended a data set for beta testing apparent duplicates. The first data set would include CDAs identified as duplicates using the most conservative criteria. The next data set would be apparent duplicates identified based on less conservative criteria.

Linda noted that the CDAs could be put into production before the duplication issue is fully resolved, with approval from Chuck May. The reason for putting them into production at this time would be to avoid having to keep re-loading the CDAs into beta every time they are evaluated. This re-loading is time consuming and increases the risk of errors because care has to be taken to ensure the correct data are loaded. Permissions can be used to limit access to staff participating in the evaluation. Having CDAs in the production database will not affect consolidated data. CDAs would be stored in a separate table. Once the duplicate algorithms are refined, duplicate CDAs can be removed from the production data sets.

Marina would like to discuss CDA duplication issues at an upcoming meeting between NCI SEER staff, Linda, Chuck May, and Brent before considering putting CDAs into production. A plan for production can be developed at this meeting, which then should be approved by the registries.

Brent added that the Louisiana registry staff loaded about a month's worth of CDAs and analyzed a small sample to evaluate the number that linked to CTCs, the number of new cases found, and the type of treatment information found. Of the CDAs that linked at both the patient and CTC level, approximately 20 included substantial treatment information that was not in the registry's abstracts. The CDAs also

included information on recurrence and progression, tumor size, receptor status, and subsequent treatment. Results of similar analyses by other registries with CDA data should be entered in template in Squish issue 6170.

Discussion

David clarified that he primarily examined the coded part of the CDA. The narrative portion was examined for the notes fields. Marina expects to be able to provide an update about moving the CDAs into production before the August call. Participants agreed with Linda that they would prefer not to have to re-load CDAs every time they reset the task server.

A data quality evaluator asked how the addition of MU data would enhance registry treatment and other data. She asked about a way to link uploaded MUs to a specific field that a registrar could code.

Participants asked if MU data would be treated like Claims data in SEER*DMS or have a different workflow. Linda explained that the plan was to use MU data in casefinding and augmenting treatment information. The CDAs are linked when they are loaded into the registry database but are not yet processed automatically. The plan is to begin using MU data to update "date of last contact" then determine whether and how to use those data for casefinding. Next, the WG will determine whether to automate the addition of treatment information from CDAs.

Casefinding Workflow **Brent Mumphrey**

The Louisiana registry drafted a workflow for casefinding. The casefinding process and proposed workflow would involve the following steps:

- CDA messages received from facilities are loaded into SEER*DMS on the autoloader daily.
- SEER*DMS attempts to match incoming records at the patient and CTC levels.
- Cases that matched at the patient level are considered possible new cases to include in the workflow.
- Identify unlinked CDA messages for which a specific amount of time has passed since date of diagnosis. The Louisiana registry staff viewed only CDA messages where at least 18 months had passed since the CDA date of diagnosis and the record still was not linked to the CTC. Queries would need to be created to identify these cases for all registries.
- Review the most recent cases and work backwards. After investigating and following back on the CDAs, the case would be considered reportable and would be abstracted or would be considered non-reportable (e.g., "history of…"), or it might be determined that the CDA should have linked to a CTC and the linkage would be performed.
- Create a new record for reportable cases. The mechanism for performing this task needs to be determined. The WG will need to decide whether registries should go back to the facility to obtain an abstract, create a case with information in the CDA, or create an AFL to indicate that the registry is waiting on the case from the facility (to avoid reviewing it a second time).
- Process non-reportable cases. A method might be needed to mark non-reportable cases so that they do not appear on the casefinding list. The WG will need to determine whether subsequent CDAs received from the same facility and patient with the same code should be automatically coded as non-reportable.
- Force matching of cases that initially should have matched with an existing CTC and ensure that subsequent CDAs for the same case are linked. A mechanism is needed to perform this task.

IMS continually matches to the pre-record table. The CDA is received; the case is created later, and matched automatically. Linda recommended not worrying about the AFL, which is a technical issue that

would be resolved later. She noted that it would be more important to determine if registries want to wait 18 months to identify unlinked CDAs. The waiting period could vary by registry.

Discussion

With regard to automated coding of non-reportable cases, participants asked if these cases would be identified by primary site and histology. CDAs from dermatology offices can change even when the tumor is located at the same primary site on the same patient. Histology therefore should be part of the algorithm to identify subsequent CDAs for the same case. Physicians tend to code histology generically, however, which leads to failures to match because of variable histology coding. Most facilities are not using ICD-0-3 coding for histology. Emark has rules that can derive the correct histology code for some cases. The physician TNM coding also tends to be inconsistent.

Participants asked about a mechanism for linking to the correct CTC. Currently, many CDAs do not link to a CTC for the same case. Linda recommended that the workflow include a mechanism that allows the CTR to identify CDAs that should have linked but did not, as well as those that linked incorrectly. CTRs then should have a mechanism for manually linking or unlinking CDAs as needed.

The WG should consider how many and what kinds of CDAs merit manual review. If a CDA does not link to a patient set in SEER*DMS, it is likely to be a new case and, therefore, a priority. CDAs that link to patient sets but not to CTCs also should be examined for new treatment information. Participants suggested examining CDAs that link at the patient but not the CTC level for new treatment information. These CDAs might have poor histology or primary site information but still match to a CTC in the registry database. These could be filtered by whether they contain treatment information, which would give them higher priority. Skin NOS also could be a filter for CDAs that can be considered a match with a range of histology codes. In addition, participants wanted to examine CDAs that link to a patient set but contain information on a site that does not come close to matching a site on the registry patient set.

Participants generally agreed that manual review of CDAs that did not match at some level was worthwhile to identify patterns of codes that create matching problems. As these patterns are identified, algorithms can be refined to improve matching for casefinding purposes. Participants also agreed that a prioritization system was needed to identify CDAs that merit manual review at different points in the workflow. Participants added that it would be useful to be able to track CDA treatment information that has been added to the database and the source of the CDA.

Brent confirmed that 18 months was sufficient time for pathology reports to be received and included in the Louisiana registry database. This timeframe, however, might vary by registry. In addition, registries might want to review certain types of CDAs sooner (e.g., from facilities that do not consistently send abstracts). Linda recommended that Brent add the option for earlier review of certain types of CDAs to his draft workflow.