OHDSI Oncology Working Group

Meeting Agendas and Meeting Decisions

**April 28th, 2022 Meeting**

Decisions that were made final:

1. For AJCC-associated staging variables in the cancer registry dataset (source dataset), the decision has been made to concurrently engage in two different mechanisms for bringing these pre-selected fields from the registry into the OMOP CDM (target dataset).
   1. 1st approach: The first method will be to hardcode mappings between post-coordinated NAACCR data and the newly pre-coordinated Cancer Modifier vocabulary (target vocabulary) into the ETL itself. This approach requires the latest version of Cancer Modifier ontology to be released by Odysseus.
   2. 2nd approach: The second method will be to utilize tableized mappings from the NAACCR source vocabulary, which is being de-standardized, to the Cancer Modifier, which is becoming the standard vocabulary and replacing the NAACCR vocabulary as the target vocabulary for NAACCR (among other) source data. The nature of this tableized mapping approach (whether using a new wide mapping table or the existing CONCEPT\_RELATIONSHIP table from newly created synthetic pre-coordinated concepts within the NAACCR source vocabulary has not yet been decided). This 2nd approach also requires the release of the Cancer Modifier ontology.
2. For preselected fields in the NAACCR data (source data) that refer to Grade of cancer cells, the decision has been to make ‘Map to’/’Maps From’ entries in the CONCEPT\_RELATIONSHP table from the lowest level in the NAACCR source vocabulary, Concept Class=’NAACCR Values’, to concepts in the new Cancer Modifier vocabulary. This approach of mapping from ‘NAACCR Value’ concepts to the new Cancer Modifier vocabulary will be the default mapping approach for mapping NAACCR to the Cancer Modifier vocabulary. This work also relies on the release of the newest version of the Cancer Modifier vocabulary.
3. Item number 1a will be fully completed before item number 2 commences.
4. No other decisions regarding any other fields of source dataset were made.

Written by: Thomas Falconer

Date: 5/5/2022

Reviewed by: Mike Gurley

Date: 5/17/2022

**May 12th, 2022 Meeting** – CANCELED as we await the release of the Cancer Modifier vocabulary

**May 26th, 2022 Meeting**

Agenda:

1. Discuss release of Cancer Modifier vocabulary to Athena
   1. Discuss ETL modifications
      1. Comment out all EPISODE table inserts
      2. Hardcode mappings between staging values and Cancer Modifier concepts into the MEASUREMENT table (again this hardcoding is only done for staging variables)
         1. This will use as the source dataset not the original NAACCR dataset format but a transposed dataset (long dataset instead of wide dataset) that is already created by the ETL.
2. For the remaining variables in NAACCR, we have decided to leverage a tableized process
   1. Who will be responsible for creating this table?
3. Programmatic implementation of past decision:
   1. It was decided by this group to add a concept to condition occurrence table that indicates initial cancer (initial diseases). This will require an ‘initial disease’ concept in the Cancer Modifier vocabulary to be created by the vocab team.