**PREPARATION STEPS**

1. Install oncology module CDM objects:
   1. Update MEASUREMENT table.

DDL is here: <https://github.com/OHDSI/OncologyWG/blob/master/ddl/PostgreSQL/OMOP%20CDM%20postgresql%20ddl%20Oncology%20Module.txt>

<https://github.com/OHDSI/OncologyWG/blob/master/ddl/Sql%20Server/OMOP%20CDM%20sql%20server%20ddl%20Oncology%20Module.txt>

* 1. Instantiate EPISODE and EPISODE\_EVENT tables.

DDLs are here:

<https://github.com/OHDSI/OncologyWG/blob/master/ddl/PostgreSQL/OMOP%20CDM%20postgresql%20ddl%20Oncology%20Module.txt>

<https://github.com/OHDSI/OncologyWG/blob/master/ddl/Sql%20Server/OMOP%20CDM%20sql%20server%20ddl%20Oncology%20Module.txt>

* 1. Instantiate and populate CONCEPT\_NUMERIC table

DDL is here:

<https://github.com/OHDSI/OncologyWG/blob/master/ddl/PostgreSQL/OMOP%20CDM%20postgresql%20ddl%20Oncology%20Module.txt>

<https://github.com/OHDSI/OncologyWG/blob/master/ddl/Sql%20Server/OMOP%20CDM%20sql%20server%20ddl%20Oncology%20Module.txt>

* + 1. Content is here:

<https://github.com/OHDSI/OncologyWG/tree/master/ddl/PostgreSQL/VocabImport>

https://github.com/OHDSI/OncologyWG/tree/master/ddl/Sql%20Server/VocabImport

* + 1. Script to load data is here:

<https://github.com/OHDSI/OncologyWG/tree/master/ddl/PostgreSQL/VocabImport>

<https://github.com/OHDSI/OncologyWG/tree/master/ddl/Sql%20Server/VocabImport>

1. Install the latest version of Athena
2. Create NAACCR\_DATA\_SOURCE table

DDLs is available here: <https://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr_etl_input_format_ddl.txt>

1. Convert Tumor Registry source data into the NAACCR\_DATA\_SOURCE table format
   1. Use these scripts (to obtain):
      1. NAACCR flat file format: email rmiller1@tuftsmedicalcenter.org.
      2. METRIQ (script cannot be publicly shared): email m-gurley@northwestern.edu
   2. Make sure that the format of dates, histology and topography are in the right format:

YYYY-MM-DD

8010/3-C50.9

* 1. Make sure to generate unique id for each primary tumor diagnosis

**ETL STEPS**

1. ETL of diagnosis and diagnostic modifiers into CONDITION\_OCCURRENCE& MEASUREMENT.

The SQL is here: <https://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr_etl_sql.docx>

or

https://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr\_etl\_temporary\_tables.sql

* 1. Ingest condition occurrence
  2. Ingest first disease occurrence modifier
  3. Ingest diagnostic modifiers with values as concepts
  4. Ingest diagnostic modifiers with values as number

1. ETL of treatments and treatment modifier into PROCEDURE\_OCCURRENCE & MEASUREMENT and DRUG\_EXPOSURE. The SQL Is here:

https://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr\_etl\_temporary\_tables.sql

* 1. Ingest surgeries and radiotherapy into PROCEDURE\_OCCURRENCE
  2. Ingest PROCEDURE\_OCCURRENCE treatment modifiers into MEASUREMENT
  3. Ingest systemic treatments into DRUG\_EXPOSURE

1. ETL of disease and treatment episodes into EPISODE, EPISODE\_EVENT, and MEASUREMENT
   1. Ingest first disease occurrence

The SQL is <https://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr_etl_sql.docx>

or

https://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr\_etl\_temporary\_tables.sql

* 1. Ingest treatments

The SQL is here: herehttps://github.com/OHDSI/OncologyWG/blob/master/documentation/naaccr\_etl\_temporary\_tables.sql