**DIAGNOSIS**

**PART I. Conventions for ETL of NAACCR items into low level tables related to the creation of episodes**

**NAACCR INITIAL DIAGNOSIS**

1. The NAACCR ETL **WILL NOT** enter a MEASUREMENT modifier of ‘'Entire Disease Course’ to a CONDITION\_OCCURRENCE entry.

NAACCR information about the most recent follow-up or death lands in the Observation\_Period and Person tables respectively. Derivation of the overarching episode based on these events will take place post ETL.

1. The NAACCR ETL **WILL** create for each record of a primary tumor an entry in CONDITION\_OCCURRENCE and **WILL** enter a MEASUREMENT modifier of ‘Initial diagnosis’:

Measurement\_Concept\_ID = 32528

Current concept in the Episode domain is invalidated. It should be moved to Cancer Modifiers domain and reinstated as standard:

32528 OMOP4822259 **Disease First Occurrence** Disease Extent Non-standard Invalid Episode Episode

If a patient has 2 BC at different times, there will be 2 entries in the CONDITION\_OCCURRENCE with 2 separate MEASUREMENT modifiers of ‘Initial diagnosis’, each one corresponding to the each primary.

1. All the diagnosis modifiers will hang off of the same CONDITION\_OCCURRENCE entry that has the initial diagnosis.

THE NAACCR ETL **WILL** enter all disease modifiers associated with each NAACCR items (for example, Clinical TNM staging, Pathological TNM Staging, Grade Clinical, Grade Pathological, SEER Summary Staging, Tumor Size Clinical, Tumor Size Pathological and even Recurrence Type--1st, Recurrence Date--1st) as MEASUERMENT modifiers to the same CONDITION\_OCCURRENCE entry having the ‘Initial diagnosis’ MEASUREMENT modifier.

All NAACCR variables representing extent of disease should be mapped/grouped into one of the following concepts. This is done to avoid complicated rules of mapping multiple modifiers representing these concepts to episodes:

* In situ disease – create new
* Local disease – create new
* Regional disease – create new
* Metastatic disease – create new

Even those NAACCR items that may have date different from the initial diagnosis date (e.g. lymph node assessment) will be modifiers of the initial diagnosis record. The actual date will be reflected in the measurement\_date of these modifiers.

**NAACCR RECURRENCE**

1. The NAACCR ETL **WILL** create records of recurrence or remission as a modifier of the initial diagnosis in the MEASUREMENT table corresponding to the NAACCR Recurrence Type--1st variable (item 1880). Depending on the value of the NAACCR Recurrence Type--1st variable, one or more records may be recorded using the following concepts as measurement\_concept\_id:

Disease dynamic modifiers

* **Progression** – create new
* **Remission** – create new

Disease extent modifiers (same as for the Initial disease above)

* **In situ disease** – create new
* **Local disease** – create new
* **Regional disease** – create new
* **Metastatic disease** – create new

Additional measurement records may be created based on the information provided in the NAACCR Recurrence Type—1 value (e.g. ).

A complete mapping table between the values of NAACCR Recurrence Type—1 and measurement\_concept\_id are provided in Mapping NAACCR to Episode.xlsx, NAACCR to Measurement sheet.

The dates of all modifier records should be the same as the date of recurrence in NAACCR.

The date of recurrence should also be reflected in the most recent OBSERVATION\_PERIOD record.

1. The NAACCR ETL **WILL** also create modifier records for the initial diagnosis in the MEASUREMENT table describing extent of the original tumor if this information is available in to the NAACCR Recurrence Type--1st variable (item 1880). Depending on the value of the NAACCR Recurrence Type--1st variable, one or more records may be recorded using the following concepts as measurement\_concept\_id:

The only available extent modifiers is

* In situ disease – create new

The dates of these modifier records should be the same as the date of the Initial Diagnosis modifier.

A complete mapping table between the values of NAACCR Recurrence Type—1 and measurement\_concept\_id are provided in Mapping NAACCR to Episode.xlsx, NAACCR to Measurement sheet.

1. The NAACCR ETL **WILL NOT** enter a MEASUREMENT modifier of ‘Dynamic’ or ‘Extent’ to any CONDITION\_OCCURRENCE entry.  Further, the NAACCR ETL **WILL NOT** create any ‘Dynamic’ or ‘Extent’  modifiers based on the Recurrence Type--1st/ Recurrence Date--1st NAACCR variables.

**GENERALISIBILITY**

These above conventions are fully applicable to the data sourced from EHR or any other sources that have explicit information about initial diagnosis, recurrence, remission, and disease extent. The respective modifiers have to be attached to the initial diagnosis condition\_occurrence record. The measurement\_date\_time for the first occurrence modifiers should be the same as condition\_date\_time; for recurrence and remission it should have their respective dates. Measurement\_concept\_type\_id should reflect the source of data.

**PART II. Conventions for post-ETL of low-level events sourced from NAACCR into EPISODES**

**DIAGNOSIS EPISODES**

1. Overarching episode

Post-ETL **WILL** createan EPISODE record corresponding to the CONDITION\_OCCURRENCE record with the modifier of ‘Initial diagnosis’. The fields in the EPISODE will be populated as follows:

* episode\_concept\_id = 32533 (‘Disease episode’)
* episode\_object\_concept\_id = CONDITION\_CONCEPT\_ID
* episode\_start\_datetime = CONDITION\_START\_DATE\_TIME
* episode\_end\_datetime = PERSON.death\_date if present or OBSERVATION\_PERIOD. observation\_period\_end\_date where OBSERVATION\_PERIOD.period\_type\_concept\_id = 32535 (‘Tumor Registry’)
* episode\_type\_concept\_id = 32535 (‘Tumor Registry’)

It **WILL** also create an EPISODE\_EVENT record linking the EPISODE and CONDITION\_OCCURRENCE records.

1. Extent and dynamic episode rules are described in the attached “Mapping NAACCR to Episode.xlsx”, Measurement to Episode sheet.