

[Logic Liaison Template][N3C Covid Enclave] Study-Specific Fact Indexing README

ASSET CREATOR(s): Evan French (VCU), Johanna Loomba (UVA), and Andrea Zhou (UVA) with design and code review support from other members of the Logic Liaison team.

README SECTIONS:

TEMPLATE STATUS

PURPOSE & USE

INPUTS

USAGE INSTRUCTIONS

DATA DICTIONARY

RELEASE NOTES (VERSION CHANGES)

TEMPLATE STATUS:

AS OF July, 2022: Initial version of the template summarizing visit-level facts around a study-specific index date. See the release notes section in this document for additional details.

PURPOSE & USE:

This template summarizes the indicators of a visit-level patient fact table with respect to whether they were present pre-, post-, or in a user-defined window surrounding an index date corresponding to a study-specific event. This template is similar to the COVID+ fact tables with the exception that it is organized around a study-specific event rather than COVID diagnosis.

INPUTS:

person_visit_facts, Table: Table containing visit-level patient facts. Must contain person_id, visit_date, and indicator columns.

indicators, Table: Table containing at least two columns: indicator_prefix and study_index_pre_window_post.

Indicator_prefix should contain the names of indicator fields in person_visit_facts. Pre_window_post should contain a comma-delimited list with some combination of pre,window,post. The values of this table configure the columns in the output dataset. It is recommended that study teams simply add a study_index_pre_window_post column to the fusion sheet which they used to build their visit-level fact table.

index_col, String: Indicator column name in person_visit_facts to used as the index for determining pre, post, window indicators.

days_pre_index, Integer: Number of days prior to the index date to start the window of interest. Should be either zero or a negative value.

days_post_index, Integer: Number of days after the index date to end the window of interest. Should be either zero or a positive value.

USAGE INSTRUCTIONS:

The template comes with "Cohort all facts table (De-identified)" dataset and fusion sheet of codesets pinned by default. These inputs are intended to be modified/replaced by study-specific datasets and codesets. For example, you can replace the inputs

with level 3 Logic Liaison fact tables or choose to use the tables from the Confirmed COVID+ Logic Liaison template instead. The primary purpose of pinning them is to give study teams an example of the input formats required.

DATA DICTIONARY:

person_id, Integer: Patient unique identifier within N3C.

<col>_date, Date: The first date on which the study-specific indicator column from the input dataset which the user selected as the index to summarize around was set to 1.

<indicator_col>_pre_<col>, Integer: A flag representing whether a given indicator column from the input dataset was set to 1 before or on <col>_date.

<indicator_col>_post_<col>, Integer: A flag representing whether a given indicator column from the input dataset was set to 1 after <col>_date.

<indicator_col>_<col>_window, Integer: A flag representing whether a given indicator column from the input dataset was set on any date within a user-defined window around <col>_date.

RELEASE NOTES:

TEMPLATE RELEASE NOTES (VERSION CHANGES):

August 23, 2022:

Initial template publication

December 9, 2022:

Updated code that searched for "visit_date" col to "date" to match change in main Logic Liaison fact tables