[Logic Liaison Template][N3C Covid Enclave] Covid-19 Cases and Controls README

ASSET CREATOR(s): Andrea Zhou (UVA) and Johanna Loomba (UVA) with ideation and concept set contributions from the Logic Liaisons, Data Liaisons, and N3C Domain Team leads.

README SECTIONS:

TEMPLATE STATUS:

AS OF APR, 2022: The underlying concept sets used for conditions in this template were built using key concepts in SNOMED-CT plus their descendants. As per direction from the N3C PIs, we plan to transition to externally validated concept sets that are being curated by the data liaisons. That change will be noted here once made. The template also accepts input from a fusion sheet/manual entry dataset to allow researchers to more easily adapt the template to read in their desired concept sets. See the release notes section in this document for additional details.

AS OF SEPT, 2024: Concept sets can now be added to the fusion sheet/manual entry dataset using the corresponding codeset_id. Whether the "_of_interest" node identifies a concept set using most recent version of a named concept set or using codeset_id has been parameterized for the user.

PURPOSE & USE:

This set of templates (Level 2 and Level 3 data versions) provides sample code and summary datasets including a day-level and a patient-level table. The day-level "all facts" table has a single row for each patient and each day (where any of the facts searched are found to be present). The patient-level table contains one row for each patient and a number of commonly referenced facts and indicators derived from the N3C datasets. The template outputs can be used in a variety of ways:

1) **ANALYZE IN CONTOUR:** You can analyze the output datasets directly in Palantir's Contour App. Simply open the LDS (Level 3 projects) or the De-ID (Level 2 projects) folder (links located just underneath where the link for this READ ME was located), select the dataset you want to use (person-level or visit-day level "all facts") then select "Analyze" in the upper-right to open the corresponding dataset in Contour. **The datasets in the knowledge object are updated with each refresh of the OMOP data in N3C.**

2) **IMPORT TEMPLATE OUTPUTS:** You can import the day-level and/or patient-level datasets that are produced by these templates directly into your code workbook or a new code workbook. **The datasets in the knowledge object are updated with each refresh of the OMOP data in N3C.**

To import into your existing code workbook, simply open the LDS (L3 projects) or De-ID (L2 projects) dataset to copy the file path next to “Location” in the middle left properties panel using the clipboard symbol that appears when hovering over the path. Then, open your code workbook, select “Import dataset” in the upper-left, select "All" in the breadcrumbs, and paste the file path into the search bar.

To analyze in a new code workbook, go to the "Actions" menu in upper-right after opening the desired dataset and select "Analyze in Code Workbook". Once in the correct workbook, the dataset can be used directly for analysis or joined with other tables you have created (for example, a cohort table you have created under your DUR).

3) **IMPORT AND CUSTOMIZE TEMPLATE:** Alternatively, you can import the appropriate code template and modify the code so that the resulting tables reflect the specific cohort and derived facts that your DUR requires for its analyses. Detailed instructions are provided below for this process.

Notes About Variable Creation

SDOH (Social Determinants of Health) DATA:

SDOH variables can be joined based on person_id using the output dataset of this template which provides variables at the ZCTA and County level as recommended by the SDOH domain team.

DEFAULT Concept Sets Used:

This set of comorbidities includes conditions listed in the Charlson Comorbidity Index. However, instead of using Quan et. al list of ICD codes only, we instead created these concept sets using the primary conditions listed in SNOMED-CT hierarchy and including all descendants.

We also created concept sets for each of the comorbidity sections listed on this CDC website as of Dec 2021: CDC "People of any age with the conditions listed below are more likely to get severely ill from COVID-19". Notes regarding each concept set and validation can be found in the Concept Set Browser. Search the comorbidity as named in the variable, but include spaces between words. For example, to look up the concept set used for the "HEARTFAILURE_indicator", search "HEART FAILURE" in the concept set browser. Johanna Loomba is the creator of the concept sets used, so you can also use this for your search.

DATA DICTIONARY (for variables in final patient-level table):

All concept sets in the below descriptions are marked by italics. You can swap out concept names in the code template’s fusion sheet/manual entry dataset to use your preferred source concept sets for any of the below variables.

person_id, string: Unique identifier for the N3C patient

total_visits, long: The total number of visit days the patient was recorded having, with hospitalizations collapsed to a single visit

observation_period, integer: The number of days between the patient's earliest recorded visit and the patient's latest recorded visit

sex, string: gender_concept_name from main OMOP Person Table. Per OMOP ETL conventions defined here, "use the gender or sex value present in the data under the assumption that it is the biological sex at birth. This field should not be used to study gender identity issues. If the source data captures gender identity it should be stored in the OBSERVATION table." Per OHDSI documentation here, "the Gender domain captures all concepts about the sex of a person, denoting the biological and physiological characteristics."

city, string: City from main OMOP location table

state, string: State from main OMOP location table

postal_code, string: Zip from main OMOP location table

county, string: County from main OMOP location table

age, long: Calculated age using the year of birth (for L2) or date of birth (for L3) and the current date. As of 7/15/22, July 1st used as placeholder month and day of birth when there are 0s or nulls in the OMOP person table to avoid biasing towards older age.

race, string: "Hispanic or Latino" for anyone with this race, otherwise "Asian", "Black or African American", “Native Hawaiian or Other Pacific Islander”, "American Indian or Alaska Native", "White", or "Other". "Unknown" is assigned to all others. NOTE: A number of patients with a source value of Asian or Other Pacific Islander are categorized as “Unknown” until the site is able to differentiate and correct their source values since the current value does not reasonably fall into only one of the 6 minimum race/ethnicity categories defined by the NIH.

race_ethnicity, string: "Hispanic or Latino Any Race" for anyone with this ethnicity, otherwise "Asian Non-Hispanic", "Black or African American Non-Hispanic", “Native Hawaiian or Other Pacific Islander Non-Hispanic”, "American Indian or Alaska Native Non-Hispanic", "White Non-Hispanic", or "Other Non-Hispanic". "Unknown" is assigned to all others. **NOTE**: A number of patients with a source value of Asian or Other Pacific Islander are categorized as “Unknown” until the site is able to differentiate and correct their source values since the current value does not reasonably fall into only one of the 6 minimum race/ethnicity categories defined by the NIH.

data_partner_id, integer: The anonymized institution contributing this patient's record to N3C

data_extraction_date, date: The date that the patient's site last ran data extraction scripts for N3C

cdm_name, string: The Clinical Data Model (CDM) associated with this patient's contributing institution (OMOP, TriNetX, ACT, or PCORNet). The different CDMs have different strengths and weaknesses and may result in a pattern of missingness.

cdm_version, string: The version of the Clinical Data Model used by this patient's contributing institution

shift_date_yn, string: Y if the patient's site shifts dates, otherwise N (note that for De-Id L2 data, all the records are shifted, so this field has been uniformly changed to reflect Y for all sites)

max_num_shift_days, string: The maximum number of days that the patient's site shifts dates, otherwise N (note that for De-Id L2 data, all records are shifted, so this field has been uniformly changed to reflect +180 days for all sites)

BMI_max_observed_or_calculated, double: The maximum Body Mass Index for a patient. Default parameters of reasonability are set for height, weight, and BMI, but can be controlled by template user as input parameters. The max reasonable BMI measure "*body mass index*" is reported. Both calculated BMI (using a weight "*Body weight (LG34372-9 and SNOMED)*" and temporally associated height "*Height (LG34373-7 + SNOMED)*" for calculation) and reported BMI as a measure are used to identify the closest measure. If both are available on the same date, the EMR reported BMI is used instead of calculated BMI.

OBESITY_indicator, integer: Value of 1 if patient has any record of condition *"OBESITY"* or calculated BMI > 30

LL_Long_COVID_clinic_visit_indicator, integer: Value of 1 if the patient had any known *"Long COVID Clinic Visit"* [note that not all sites are transmitting this data]

LL_Long_COVID_diagnosis_indicator, integer: Value of 1 if the patient has Long COVID Diagnosis U09.9 (concept set *"Long-COVID (PASC)"*)

LL_COVID_diagnosis_indicator, integer: Value of 1 if the patient had a COVID-19 diagnosis *"N3C Covid Diagnosis"* charted.

LL_MISC_indicator, integer: Value of 1 if the patient has any record of *"MULTISYSTEM INFLAMMATORY SYNDROME - CHILDREN"*

LL_PNEUMONIADUETOCOVID_indicator, integer: Value of 1 if the patient has any record of *"PNEUMONIA DUE TO COVID-19"*

LL_SUSPECTEDCOVID19_indicator, integer: Value of 1 if the patient has any record of *“SUSPECTED COVID-19”*

SYSTEMICCORTICOSTEROIDS_indicator, integer: Value of 1 if COVID-19 patient has any record (order or administration) of *“N3C CORTICOSTEROIDS FOR SYSTEMIC USE”* drug. See the "Logic Liaison Tips" section for more information about medications.

PCR_AG_Pos_indicator, integer: Value of 1 if the patient has any record of *"ATLAS SARS-CoV-2 rt-PCR and AG"* with *"ResultPos"* value.

PCR_AG_Neg_indicator, integer: Value of 1 if the patient has any record of *"ATLAS SARS-CoV-2 rt-PCR and AG"* with *"ResultNeg"* value.

Antibody_Pos_indicator, integer: Value of 1 if the patient has any record of *"Atlas #818 [N3C] CovidAntibody retry"* with *"ResultPos"* value. Please note that some positive Antibody tests are the result of prior vaccine and therefore non-specific to prior infection.

Antibody_Neg_indicator, integer: Value of 1 if the patient has any record of *"Atlas #818 [N3C] CovidAntibody retry"* with *"ResultNeg"* value.

total_number_of_COVID_vaccine_doses, long: Count of COVID-19 vaccine doses received by the patient. See community note regarding N3C vaccination data considerations.

confirmed_covid_patient, integer: Value of 1 if patient has a value of 1 for either COVID_diagnosis_indicator or PCR_AG_Pos_indicator

possible_covid_patient, integer: Value of 1 if patient has a value of 0 for confirmed_covid_patient but has a value of 1 for any of the following: Antibody_Pos_indicator, LL_Long_COVID_diagnosis_indicator, LL_Long_COVID_clinic_visit_indicator, LL_PNEUMONIADUETOCOVID_indicator, LL_MISC_indicator, LL_SUSPECTEDCOVID19_indicator

patient_death_indicator, integer: Value of 1 if the patient had death or discharge to hospice recorded in N3C

COMORBIDITY indicators:

The_indicator columns have a value of 1 if a patient has any record of given comorbidity.

PREGNANCY_indicator, integer: Value of 1 if patient has any record of condition *"PREGNANT"*

SOLIDORGANORBLOODSTEMCELLTRANSPLANT_indicator, integer: Value of 1 if patient has any record of condition *"TRANSPLANT OF SOLID ORGAN OR BLOOD STEM CELL"*

TOBACCOSMOKER_indicator, integer: Value of 1 if patient has any record of condition *"TOBACCO SMOKER"*

CARDIOMYOPATHIES_indicator, integer: Value of 1 if patient has any record of condition *"CARDIOMYOPATHIES"*

CEREBROVASCULARDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"CEREBROVASCULAR DISEASE"*

CHRONICLUNGDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"CHRONIC LUNG DISEASE"*

CONGESTIVEHEARTFAILURE_indicator, integer: Value of 1 if patient has any record of condition *"CONGESTIVE HEART FAILURE"*

CORONARYARTERYDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"CORONARY ARTERY DISEASE"*

DEMENTIA_indicator, integer: Value of 1 if patient has any record of condition *"DEMENTIA"*

DEPRESSION_indicator, integer: Value of 1 if patient has any record of condition *"DEPRESSION"*

DIABETESCOMPLICATED_indicator, integer: Value of 1 if patient has any record of condition *"DIABETES COMPLICATED"*

DIABETESUNCOMPLICATED_indicator, integer: Value of 1 if patient has any record of condition *"DIABETES UNCOMPLICATED"*

DOWNSYNDROME_indicator, integer: Value of 1 if patient has any record of condition *"DOWN SYNDROME"*

HEARTFAILURE_indicator, integer: Value of 1 if patient has any record of condition *"HEART FAILURE"*

HEMIPLEGIAORPARAPLEGIA_indicator, integer: Value of 1 if patient has any record of condition *"HEMIPLEGIA or PARAPLEGIA"*

HIVINFECTION_indicator, integer: Value of 1 if patient has any record of condition *"HIV INFECTION"*

HYPERTENSION_indicator, integer: Value of 1 if patient has any record of condition *"HYPERTENSION"*

KIDNEYDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"KIDNEY DISEASE"*

MALIGNANTCANCER_indicator, integer: Value of 1 if patient has any record of condition *"MALIGNANT CANCER"*

METASTATICSOLIDTUMORCANCERS_indicator, integer: Value of 1 if patient has any record of condition *"METASTATIC SOLID TUMOR CANCERS"*

MILDLIVERDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"MILD LIVER DISEASE"*

MODERATESEVERELIVERDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"MODERATE OR SEVERE LIVER DISEASE"*

MYOCARDIALINFARCTION_indicator, integer: Value of 1 if patient has any record of condition *"MYOCARDIAL INFARCTION"*

OTHERIMMUNOCOMPROMISED_indicator, integer: Value of 1 if patient has any record of condition *"IMMUNODEFICIENCY"*

PEPTICULCER_indicator, integer: Value of 1 if patient has any record of condition *"PEPTIC ULCER"*

PERIPHERALVASCULARDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"PERIPHERAL VASCULAR DISEASE"*

PSYCHOSIS_indicator, integer: Value of 1 if patient has any record of condition *"PSYCHOSIS"*

PULMONARYEMBOLISM_indicator, integer: Value of 1 if patient has any record of condition *"PULMONARY EMBOLISM"*

RHEUMATOLOGICDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"RHEUMATOLOGIC DISEASE"*

SICKLECELLDISEASE_indicator, integer: Value of 1 if patient has any record of condition *"SICKLE CELL DISEASE"*

SUBSTANCEUSEDISORDER_indicator, integer: Value of 1 if patient has any record of condition *"SUBSTANCE USE DISORDER"*

THALASSEMIA_indicator, integer: Value of 1 if patient has any record of condition *"THALASSEMIA"*

TUBERCULOSIS_indicator, integer: Value of 1 if patient has any record of condition *"TUBERCULOSIS"*

TEMPLATE LIMITATIONS

Because our goal is to provide a widely usable set of variables with a well-documented and easy-to-understand approach, we have not customized the underlying concept sets to meet the needs of any specific project. Instructions for modifying the template are provided below should you want to replace the concept sets used with your own versions. Other limitations of the template reflect common limitations of this data, some of which are described in the tips below.

INTRODUCTORY LOGIC LIAISON TIPS

(see N3C Community Note here)

HOW TO MODIFY THIS TEMPLATE FOR A SPECIFIC N3C PROJECT

You may wish to modify input parameters (such as acceptable BMI range, for example), underlying concept sets, or time frames used (say to define hospitalization). You may also want to add additional variables or join the template output tables with a cohort or other data frame you have produced. The below section describes how to import the code template and convert to code transform wherever code customization is required.

Import Template

- Open a new code workbook and select "Skip this Step" under "Import Dataset".
- Make sure you are using the "*profile-high-executor-cores-and-memory*" environment (selected under the environment menu towards the top middle of the window: Configure environment → profile-high-executor-cores-and-memory on the left panel of pop-up → Update environment).
- Click "New Transform", select "Templates", search for the "[LOGIC LIAISON TEMPLATE] L2 and L3 Fact Tables: Covid-19 Cases and Controls" and import into the workbook. When you press the 'apply transformation' button, a box pops up that says that there are resources within the template that are not within the scope of the project. You can agree to add these resources.
- Once the template nodes appear, select "zoom to fit" icon from lower-left corner so that the transforms all fit on your screen.
- *Y*ou will need to review and edit each transformation as needed in each step below.
- **To just view the code**, select a transformation node and click the "Toggle View" button in upper-right of the control window that appears at the bottom half of the screen
- **To change the code**, select **"**Actions" button to the right of "Toggle View" **i**n the transformation node, and click **"Convert to code transform"**. **Note: once converted to code transform, you will not be able to update your imported template with the most recent version of the Knowledge Store template while keeping your changes made within the same nodes.*

Import OMOP Data

- The template selected will import the domain tables for the level of data chosen. The L2 template will automatically import deidentified domain tables and the L3 template will automatically import the limited dataset domain tables.

These tables will appear on the left side. Notice, we use the microvisit_to_macrovisit table for identifying hospitalizations (see enclave documentation here) as opposed to the visit_occurrence table.

Import Fusion Sheet/Manual Entry Dataset

- The template will default to importing the dataset generated by the "SNOMED" tab of a fusion sheet. This set of concept sets are ones developed using a small selection of parent SNOMED concepts and including all SNOMED-CT descendants.
- To customize this list of concept sets used in the fact table, you should copy and paste the default LL_concept_sets_fusion dataset into a fusion sheet or manual entry dataset to edit.
- **A note on uploading concept sets:** *Determine if you will use codeset_id to identify any of your concept sets*; this will allow you to use a concept_set that is not the most recent version. To do so, toggle the option to TRUE using "use_codeset_id" for each domain where using a version that is not the most recent is desired. When this option is on, you may identify a concept set using either the codeset_id or the concept_set_name. *If you will only be using concept_set_name to identify all concept sets,* and therefore will be using the most recent version for all concept sets, then keep the option off by ensuring the "use_codeset_id" is toggled to FALSE which is the template default.
- Depending on how you will be identifying concept sets, ensure you have the necessary column(s) specified:
- If you will be using **concept_set_name** to identify **ALL** of your concept sets:
    - **concept_set_name** should be exactly as it is found in the concept set browser. The code will default to the most recent version of the specified concept set.
- If you will be using **codeset_id** to identify **ANY** of your concept sets:
    - **codeset_id** should be exactly as it is found in the concept_set_members table. The codeset_id can be for any version of the concept set. If you add a concept_set_name and a codeset_id for the same concept set, the codeset_id will be prioritized.
    - **concept_set_name** should be exactly as it is found in the concept set browser. The code will default to the most recent version of the specified concept set. While you have the option to use codeset_id, you do not have to, and any concept set can be identified using concept_set_name. However, if you enter both a concept_set_name and a codeset_id for the same concept set, the codeset_id is prioritized.

Ensure you specify all two other necessary columns using the appropriate information:

- **indicator_prefix** field is whatever you wish to name the variable
- **domain** is the OMOP table(s)/domain(s) in which the concepts of the concept set specified in column A can be found
- **lower_bound** and **upper_bound** is used to provide bounds for measurements of interest. Harmonized measurements will have built in bounds, however if you wish to have tighter bounds, or use non-harmonized measures, you can define your own. **Special Note:** If the measurement concept set being added is a qualitative measure you must enter **qual** into the lower_bound to ensure it is handled as a qualitative measure.
- Once you have a new input table with your chosen concept sets, connect your new table as input to the **customize_concept_sets** node by clicking the white box with "LL_concept_sets_fusion_everyone" displayed in the template view of the node. Once the box has turned a light orange, select your new table to set it as the input (no manual changes needed to the code!).
- Running the pipeline will now create event flags to the visit-level and patient-level tables for all the concept sets you defined in the customized input.

Everyone_Cohort Node

- This node gathers some commonly used facts about these patients from the "person" and "location" tables, as well as some facts about the patient's institution (from the "manifest" table). Available age, race, ethnicity, and locations data is gathered at this node. The patient's total number of visits as well as the number of days in their observation period is calculated from the "microvisits_to_macrovisits" table in this node after filtering for plausible visit dates based on a site's latest data_extraction_date. These facts will eventually be joined with the final patient-level table in the final node.
- **IMPORTANT NOTE FOR ANALYSTS:** If you are modifying code or adding to this pipeline, you might consider setting the value of the 'proportion_of_patients_to_use' variable to something <1.0. This allows you to work with a small sample of the patients while testing. You can change this value back to 1.0 once done debugging your code.

Find Facts Related to the Patients in Cohort

Note that the below nodes all filter rows to those patients that are identified in the COHORT node.

- The **conditions_of_interest** node filters the condition_occurrence table for rows that have a condition_concept_id associated with one of the concept sets described in the data dictionary in the README through the use of a fusion sheet. Indicator names for these conditions are assigned, and the indicators are collapsed to unique instances on the basis of patient and date.
- The **observations_of_interest, drugs_of_interest, procedures_of_interest, and devices_of_interest** nodes filter the source OMOP tables for rows that have a standard concept id associated with one of the concept sets described in the data dictionary through the use of a fusion sheet/manual entry dataset. Indicator names for these variables are assigned, and the indicators are collapsed to unique instances on the basis of patient and date.
- The **quant_measurements_of_interest** node filters the measurements table for rows that have a measurement_concept_id associated with one of the concept sets described in the above data dictionary. It finds the harmonized value as a number for the quantitative measurements and collapses these values to unique instances on the basis of patient and date.
- The **quali_measurements_of_interest** node also filters the measurements table for rows that have a measurement_concept_id associated with one of the concept sets described in the above data dictionary. Then it finds the value as concept id for the qualitative measurements (positive and negative lab results) and collapses these to unique instances on the basis of patient and date.
- Measurement BMI cutoffs included are intended for adults. Analyses focused on pediatric measurements should use different bounds for BMI measurements. Note that the code looks for both BMI that is directly recorded as well as BMI that is calculated. If both are in a reasonable range and have the same date associated, the recorded BMI is used. Please note that default "reasonable" high low limits are as follows, but can be controlled by the user as a template parameter. Just click on this template node to edit the input parameters and rerun the workbook.
- reasonable BMI: 10 - 100
- reasonable Height: 0.6 - 2.43 meters
- reasonable Weight: 5-300 kilograms
- The **death_persons** node code filters the visits table for rows that have a discharge_to_concept_id that corresponds with the DECEASED or HOSPICE concept sets and combines these records with the patients in the deaths table. Death dates are taken from the deaths table and from the visits table if the patient has a discharge_to_concept_id that corresponds with the DECEASED concept set. Death dates are prioritized such that we filter to only plausible death dates from the OMOP death table and take a min prior to adding in any additional plausible death dates from the OMOP visits table with concept id belonging to the DECEASED concept set and take a max for patients who do not already have a death date from the OMOP death table. No date is retained for patients who were discharged to hospice.

Assemble Cohort All Facts Table (visit day table)

- All facts collected in the previous steps are combined in the **all_patients_visit_day_facts_table** on the basis of unique patient visit days and/or start dates from OMOP tables. Indicators are created for the presence or absence of events, medications, conditions, measurements, device exposures, observations, procedures, and outcomes. It also creates an indicator for whether the visit date where a fact was noted occurred during any hospitalization.
- This table is useful if the analyst needs to use actual dates of events as it provides more detail than the final patient-level table. Use the max and min functions to find the first and last occurrences of any events.

Create Final Patient Level Table

- The final step is to aggregate information to create a data frame that contains a single row of data for each patient in the cohort.
- The right-most node in the code workbook (**all_patients_summary_facts_table**) aggregates all information from the cohort_all_facts_table and summarizes each patient's facts in a single row.

UPDATING TO LATEST TEMPLATE VERSION:

- See N3C Community Note titled LL Technical Tip: Updating multi-node template versions

RELEASE NOTES:

May 6, 2022:

Initial publication of template.

May 18, 2022:

Modifications:

- Updated joins in the cohort_all_facts node to automatically join on any shared columns between the two datasets being joined. This will allow researchers to utilize input fusion sheet without need for manually accounting for shared columns from the _of_interest nodes when bringing all facts together for the final fact tables.

June 17th, 2022:

Modifications:

- Retain PCR_AG_Pos, PCR_AG_Neg, Antibody_Pos, and Antibody_Neg flags from cohort_all_facts node when collapsing to all_patients_summary_facts_table.
- Added PREGNANCY, OTHERIMMUNOCOMPROMISED, SOLIDORGANORBLOODSTEMCELLTRANSPLANT, PNEUMONIADUETOCOVID and MISC comorbidity condition indicators via fusion sheet input.

June 29th, 2022:

- Added confirmed_covid_patient_indicator and possible_covid_patient_indicator columns to summary table to flag whether each person id represents a confirmed COVID or possible COVID patient

July 15th, 2022:

Additions:

- Created flag for vaccination event in visit-level table and created count of vaccinations in person-level table, using the output of the Vaccine Fact template
- Flag patients who were who were expired at the time of discharge or discharged to hospice from the visits domain as deceased

Modifications:

- Switched from using January 1st to July 1st as placeholder month and day of birth when there are 0s or nulls in the OMOP person table to avoid biasing towards older age.
- Separated the fusion input dataframe into two, one that can be completely customized and one that the required input rows for the template to run and create the proper derived variables in the patient summary node (LL_MISC, LL_PNEUMONIADUETOCOVID, LL_Long_COVID_clinic_visit, LL_Long_COVID_diagnosis, and LL_COVID_diagnosis)
- Retain all possible visit dates from the visit_occurrence table for each patient as opposed to the visit dates for each patient where a fact was found

September 7th, 2022:

- "race_ethnicity" variable has been updated to include new values of race_concept_name from latest data release on 9/1/22, including American Indian or Alaska Native Non-Hispanic

October 28th, 2022:

- Added PULMONARYEMBOLISM comorbidity condition indicator via fusion sheet input

December 7th, 2022:

Additions:

- Added "race" variable using only race_concept_name from the OMOP person table. Created for researchers who do not necessarily want to use the combined "race_ethnicity" variable from the template.

Modifications:

- Updated column names to more accurately reflect the data that is represented in those fields. See Data Dictionary for more details.
    - "gender_concept_name" (now named "sex") OHDSI definitions and explanations are linked in the data dictionary entry.
    - "visit_date" (now named "date")
- County level SDoH variables removed from L3 version of template due to new release of SDoH template that provides ZCTA level social determinants of health along with a select few county level variables not covered by the ZCTA level columns.
- Added Suspected Covid-19 to concept set fusion sheet and to the possible_covid_patient flag definition.

January 4th, 2023:

- Switched to using SUBSTANCE USE DISORDER concept set with appropriate renaming of indicator from prior concept set of SUBSTANCE ABUSE via fusion sheet input

February 2nd, 2023:

- Switched to using AUTOIMMUNE DISEASE/IMMUNODEFICIENCY concept set with appropriate renaming of indicator from prior concept set of IMMUNODEFICIENCY via fusion sheet input

August 9, 2023:

- Updated filter for unreasonable COVID patient death dates to be prior to 05-01-2017 and after data_extraction_date plus two years for visit level facts table in Level 2 or De-Identified template to account for date shifting

December 6, 2023:

- patient_death_indicator updated to prioritize death dates from the death domain over death dates from the visits domain
- Removed future visits from observation period and visit count calculations
- Added macrovisit start and end date columns to all facts table rows

August 2024:

- Reviewed and updated default concept sets for domain changes after OHDSI OMOP vocabulary refresh

September 2024:

- LL_DO_NOT_DELETE_REQUIRED_concept_sets_all and LL_concept_sets_fusion_everyone fusion sheets have an added column "codeset_id" which allows users to enter a codeset_id to identify concept sets. This option can be toggled on using the "use_codeset_id" parameter in each domain. While using concept_set_name automatically filters to check for the most recent version, using codeset_id allows you to choose any version.

November 2024:

- Added the concept sets body mass index, Body weight (LG34372-9 and SNOMED), Height (LG34373-7 + SNOMED), ATLAS SARS-CoV-2 rt-PCR and AG, Atlas #818 [N3C] CovidAntibody retry, ResultPos, and ResultNeg to LL_DO_NOT_DELETE_REQUIRED_concept_sets_confirmed. These are the required measurement concepts. ResultPos, and ResultNeg are value concept sets for qualitative measurements.
- Added canonical_units_of_measure used to confirm the harmonized unit for quantitative measurements.
- customize_concept_sets will now check that included measurement concepts have a harmonized unit associated with them. If one or more measurements does not have a harmonized unit, this node will now produce a warning.
- quali_measurements_of_interest uses Pos_Value and Neg_Value from LL_DO_NOT_DELETE_REQUIRED_concept_sets_confirmed to find concepts that correspond to a positive or negative test value. Tests with a positive or negative result are the only qualitative measurements accepted at the moment
- quant_measurements_of_interest will provide the harmonized value and unit for any quantitative measurement. If no harmonized value is available a list of non-harmonized units with be provided with the value.