

OBSERVATION DERIVATION RECOVER TABLE METADATA

TABLE OF CONTENTS

<i>Observation Derivation RECOVER Table Metadata</i>	1
<i>Conditions</i>	2
<i>Tests</i>	3
<i>Immunizations</i>	5
<i>Drugs</i>	6
<i>COVID+ Definitions</i>	7
<i>PASC Rules-Based Phenotype</i>	9
<i>PASC Rules-Based diagnosis clusters</i>	11

CONDITIONS

observation_concept_id	observation_concept_name	value_as_concept_id	value_as_concept_name		qualifier_concept_id	observation_type_concept_id	observation_type_concept_name	observation_source_concept_id
2000001527	COVID-19 condition status	2000001524	COVID-19 exposure	All possible values for qualifier_concept_id depend on whether diagnosis is derived from SNOMED code, ICD code or source value	- If SNOMED code, then condition_concept_id	2000001517	Structured code derived	condition_occurrence_id
2000001527	COVID-19 condition status	2000001523	COVID-19 complication		- If ICD code, then condition_source_concept_id			
2000001527	COVID-19 condition status	2000001522	History of COVID-19					
2000001527	COVID-19 condition status	2000001525	COVID-19 specific diagnosis					
2000001527	COVID-19 condition status	2000001520	Post acute sequelae of COVID-19 observation (PASC)	vocabulary_id of diagnosis term - hierarchy: SNOMED → ICD → source_value_derived → registry_derived	If source value derived, then condition_concept_id	2000001519	Source value derived	condition_occurrence_id
2000001527	COVID-19 condition status	703578	Multisystem inflammatory syndrome (MIS)					
2000001527	COVID-19 condition status	2000001533	Nonspecific post-viral disorder					

Note: Not all nonspecific post-viral disorder diagnoses (ICD10CM B94.8) are included—only those which are not attributable to a non-COVID condition as determined by the corresponding SNOMED code (condition_concept_id)

observation_c oncept_id	observation_con cept_name	value_as_co ncept_id	value_as_conc ept_name	qualifier_conc ept_id	observation_sourc e_concept_id	All possible values for observation_type_ concept_id depend on whether it is structural code derived or source value derived	observation_type _concept_id	observation_type_c oncept_name
2000001530	COVID-19 PCR test	2000001526	First positive instance	measurement_ concept_id	measurement_id		2000001517	Structured code derived
2000001530	COVID-19 PCR test	2000001415	COVID-19 Test Result Unknown	measurement_ concept_id	measurement_id		2000001519	Source value derived
2000001530	COVID-19 PCR test	9191	Positive	measurement_ concept_id	measurement_id			
2000001530	COVID-19 PCR test	9189	Negative	measurement_ concept_id	measurement_id			
2000001528	COVID-19 serology test	2000001526	First positive instance	measurement_ concept_id	measurement_id	All possible values for observation_type_ concept_id depend on whether it is structural code derived or source value derived	2000001517	Structured code derived
2000001528	COVID-19 serology test	2000001415	COVID-19 Test Result Unknown	measurement_ concept_id	measurement_id		2000001519	Source value derived
2000001528	COVID-19 serology test	9191	Positive	measurement_ concept_id	measurement_id			
2000001528	COVID-19 serology test	9189	Negative	measurement_ concept_id	measurement_id			

2000001529	COVID-19 antigen test	2000001526	First positive instance	measurement_concept_id	measurement_id	All possible values for observation_type _concept_id depend on whether it is structural code derived or source value derived	2000001517	Structured code derived
2000001529	COVID-19 antigen test	2000001415	COVID-19 Test Result Unknown	measurement_concept_id	measurement_id		2000001519	Source value derived
2000001529	COVID-19 antigen test	9191	Positive	measurement_concept_id	measurement_id			
2000001529	COVID-19 antigen test	9189	Negative	measurement_concept_id	measurement_id			

Note: Serology testing includes nucleocapsid and S-protein testing (potentially amongst other tests; see the [serology test codeset](#) for more information). Depending on test type and whether a patient was previously immunized, positive serology tests may not indicate that the patient had COVID. Note that serology testing is limited to nucleocapsid-only testing in the COVID+ definition shown below. Logic which identifies qualifying serology tests (positive nucleocapsid testing only) is implemented in the COVID ‘infection date any’ definition below (see ODR code [here](#)).

IMMUNIZATIONS

observation_ concept_id	observation_co ncept_name	value_as_c oncept_id	value_as_con cept_name	qualifier_con cept_id	observation_sour ce_concept_id	All possible values for observation_typ e_concept_id depend on whether it is structural code derived or source value derived	observation_typ e_concept_id	observation_type_ concept_name
2000001492	COVID-19 Vaccine	200000149 2	COVID-19 Vaccine	immunization _concept_id	immunization_id		2000001517	Structured code derived
							2000001519	Source value derived

DRUGS

observation_ concept_id	observation_as_c oncept_name	value_as_c oncept_id	value_as_con cept_name	qualifier_c oncept_id	observation_sour ce_concept_id	All possible values for observation_typ e_concept_id depend on whether it is structural code derived or source value derived	observation_typ e_concept_id	observation_type_ concept_name
2000001555	COVID-19 drug exposure	200000155 4	Paxlovid drug exposure	drug_conce pt_id	drug_exposure_id		2000001517	Structured code derived
2000001555	COVID-19 drug exposure	200000155 3	Remdesevir drug exposure	drug_conce pt_id	drug_exposure_id		2000001519	Source value derived

COVID+ DEFINITIONS

Note: There are three definitions of COVID+ currently in use, distinguished by the observation_source_value_field.

The "Infection date known" definition only includes COVID positive events which specify an exact date of infection: positive PCR and antigen tests, specific COVID diagnoses, and Paxlovid and Remdesivir prescriptions. The "infection date any" definition also allows for PASC (U09.9), MIS-C (M35.81), and Non-specific post-viral disorder (B94.8) diagnoses. The "def_study_full" definition has not been used in CSC queries, but is the most expansive definition by including complications of COVID diagnoses and nucleocapsid serology testing in addition to the inclusion criteria in the "infection_date_any" definition.

(In the PEDSnet data, the difference between the infection_date_any and full study definitions is small, only about 1000 patients. The infection_date_known definition is more restrictive, with a few thousand fewer patients, but the patients it misses are those with PASC diagnoses and no prior evidence of COVID and may disproportionately be of interest in studies.)

For COVID events where the infection date needs to be imputed, it is imputed as 59 days prior to the index event (e.g. U09.9 diagnosis). Within each definition, only the earliest infection date (known or, if allowed, imputed) is selected. If the observation_date is imputed, it may not correspond to an actual visit—whether the date was imputed can be identified from the value_as_concept_id.

(Note: the 'def_study_full' definition was added as of s10)

observation_concept_id	observation_concept_name	observation_source_value	value_as_concept_id	value_as_concept_name	qualifier_concept_id	observation_source_concept_id		observation_type_concept_id	observation_type_concept_name
2000001556	Query definition COVID-positive	def_infection_date_any	2000001557	Observation date known	NA	NA	All possible values for observation_type_concept_id depend on earliest COVID positive event	2000001530	COVID-19 PCR test
								2000001529	COVID-19 antigen test
			2000001558	Observation date imputed				2000001528	COVID-19 serology test (NOTE: only nucleocapsid serology included)
								2000001525	COVID-19 specific diagnosis

								2000001523	COVID-19 complication
2000001556	Query definition COVID-positive	def_study_full	2000001557	Observation date known	NA	NA			
			2000001558	Observation date imputed					
2000001556	Query definition COVID-positive	def_infection_date_known	2000001557	Observation date known	NA	NA		2000001520	PASC (U09.9) diagnosis
								703578	MIS-C diagnosis
								2000001533	Non-specific post-viral disorder
								2000001553	Remdesevir exposure
								2000001554	Paxlovid exposure

PASC RULES-BASED PHENOTYPE

The PASC rules-based phenotype classifies patients based on the level of certainty of PASC based on clinical events and characteristics. Each patient meeting the query definition COVID positive inclusion criteria is evaluated for evidence of PASC, and if a patient does meet one of the PASC definitions, one row is generated in observation_derivation_recover (at the patient-level) indicating the certainty of PASC and multiple rows are generated (per patient) with each event contributing to evidence of PASC.

The patient-level rows are coded as follows, where a patient is included at their highest level of certainty (i.e. a patient meeting the probable and conclusive criteria will only appear as conclusive):

observation_concept_id	observation_as_concept_name	value_as_concept_id	value_as_concept_name	value_source_value	observation_date	qualifier_concept_id	observation_source_concept_id	observation_type_concept_id	observation_type_concept_name	NOTES
2000001535	Rules based CP patient	2000001536	PASC rules-based CP conclusive	delineated string of all observation_type_concept_id inclusion criteria for patient	Date of first evidence of PASC	NA	NA	NA	NA	
2000001535	Rules based CP patient	2000001537	PASC rules-based CP probable	delineated string of all observation_type_concept_id inclusion criteria for patient	Date of first evidence of PASC	NA	NA	NA	NA	
2000001535	Rules based CP patient	2000001538	PASC rules-based CP possible	delineated string of all observation_type_concept_id inclusion criteria for patient	Date of first evidence of PASC	NA	NA	NA	NA	Excluded starting in s10

The inclusion criteria-level rows are coded as follows, where each patient described in the table above will have one or more rows. **NOTE** that beginning in s10, only the **Rules based CP patient** (i.e. none of the **Rules-based CP inclusion criterion** rows outlined below) rows are included, but that the observation_type_concept_id for the criterion are included in the value_source_value at the **Rules based CP patient** level. observation_type_concept_id + observation_type_concept_id that can be found in the value_source_value are described below:

observation_type_concept_id	observation_type_concept_name
2000001539	MIS-C code 2 separate visits
2000001540	PASC code 2 separate visits
2000001541	Post-viral sequelae 2 separate visits
2000001542	MIS-C code 1 visit
2000001543	PASC code 1 visit
2000001551	Post-viral sequelae 1 visit
2000001562	Evidence of COVID and cluster conditions

PASC RULES-BASED DIAGNOSIS CLUSTERS

These rows track the PASC-associated diagnosis clusters which are among the inclusion criteria for the rules-based phenotype, and are useful for assigning PASC-probable/conclusive patients to rules-based subphenotypes. For a cluster to be included as evidence of PASC for a patient, the patient must have at least two incident* diagnoses in a cluster during the 28-179 day post-acute period, at least 28 days apart.

Notes:

- As patients may present with multiple groups of PASC-associated diagnoses, there may be multiple rows per patient. For example, if a patient qualified for inclusion in the phenotype with cardiovascular signs and symptoms, fever and chills, and myositis, there would be three rows representing the patient with this observation_concept_id (and different value_as_concept_ids representing the three clusters). However, there is just one row per cluster (rather than one row per diagnosis)
- Some patients who are qualified as conclusive or probable under the rules-based phenotype may not have diagnoses in any specific cluster and may have instead been included due to a U09.9/B94.8/M35.81 diagnosis—these PASC/MIS-C diagnoses are not represented in rows below but may instead be identified via the "COVID-19 condition status" rows in ODR above.
- Diagnoses in a cluster count as incidence if they occurred in their corresponding cluster during the post-acute period and did not occur during a pre-COVID infection washout period. The length of the washout period varies by diagnosis code, and is part of the rules-based CP logic.
- Groups of clusters present below may be used to classify patient into rules-based subphenotypes. These additional rows in ODR were created to enable this use case.

observation_concept_id	observation_concept_name	value_as_concept_id	value_as_concept_name	value_source_value	observation_date	qualifier_concept_id	observation_source_concept_id	observation_type_concept_id	observation_type_concept_name
2000001563	Rules based CP cluster	2000001564	PASC CP cluster cardiovascular signs and sx	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001565	PASC CP cluster heart disease	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA

2000001563	Rules based CP cluster	2000001566	PASC CP cluster skin	NA	Date of first incident post-acute diagnosis in cluster	condition_source _concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001567	PASC CP cluster fever and chills	NA	Date of first incident post-acute diagnosis in cluster	condition_source _concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001568	PASC CP cluster myositis	NA	Date of first incident post-acute diagnosis in cluster	condition_source _concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001569	PASC CP cluster myocarditis	NA	Date of first incident post-acute diagnosis in cluster	condition_source _concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001570	PASC CP cluster cognitive function	NA	Date of first incident post-acute diagnosis in cluster	condition_source _concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA

2000001563	Rules based CP cluster	2000001571	PASC CP cluster hair loss	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001572	PASC CP cluster arrhythmias	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001573	PASC CP cluster POTS	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001574	PASC CP cluster abnormal liver enzymes	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001575	PASC CP cluster headache	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA

2000001563	Rules based CP cluster	2000001576	PASC CP cluster musculoskeletal	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001577	PASC CP cluster respiratory signs and sx	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001578	PASC CP cluster chest pain	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001579	PASC CP cluster acute kidney injury	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001580	PASC CP cluster fatigue and malaise	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA

2000001563	Rules based CP cluster	2000001581	PASC CP cluster thrombophlebitis and thromboembolism	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001582	PASC CP cluster changes in taste and smell	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001583	PASC CP cluster generalized pain	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001584	PASC CP cluster fluid and electrolyte	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
2000001563	Rules based CP cluster	2000001585	PASC CP cluster abdominal pain	NA	Date of first incident post-acute diagnosis in cluster	condition_source_concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA

2000001563	Rules based CP cluster	2000001586	PASC CP cluster acute respiratory distress syndrome	NA	Date of first incident post-acute diagnosis in cluster	condition_source _concept_id of first incident post-acute diagnosis in cluster	condition_occurrence_id of first incident post-acute diagnosis in cluster	NA	NA
------------	------------------------	------------	---	----	--	---	---	----	----