

A Data-Driven Analysis of Pediatric Organ Dysfunction Patterns to Discover Sepsis Phenotypes

Study Database

Project Title: A Data-Driven Analysis of Pediatric Organ Dysfunction Patterns to Discover Sepsis Phenotypes (Grant number: 1R21HD096402-01A1; PI: Sanchez-Pinto)

Description: The study database is a structured, de-identified, multi-center, electronic health record-based database of pediatric intensive care unit (PICU) patients with a confirmed or suspected infection on admission that is designed to support the NIH-funded project “A Data-Driven Analysis of Pediatric Organ Dysfunction Patterns To Discover Sepsis Phenotypes”, and subsequent secondary analyses. The study database consists of a data model with a relational schema and a set of structured data elements necessary to carry out the project and related analyses. This document contains the information necessary to identify the cohort of patient encounters that meet criteria for inclusion in the database, define the data elements required, and structure these data elements in series of tables in a common data model.

De-identification: Patients and patient encounters will be assigned unique, randomly-generated study codes and will be fully de-identified at each site by eliminating all 18 Safe Harbor personal health information (PHI) elements, as per section 164.514(b)(2) of the HIPAA Privacy Rule. This will include elimination of names, medical record numbers, billing numbers, and dates (which will be transformed to time from PICU admission in minutes). The tables mapping the coded data with PHI at each site will not be shared across sites and hence these data will be considered completely de-identified for the purposes of the study.

Document Author: L. Nelson Sanchez-Pinto

Document Version: 4-22-2019

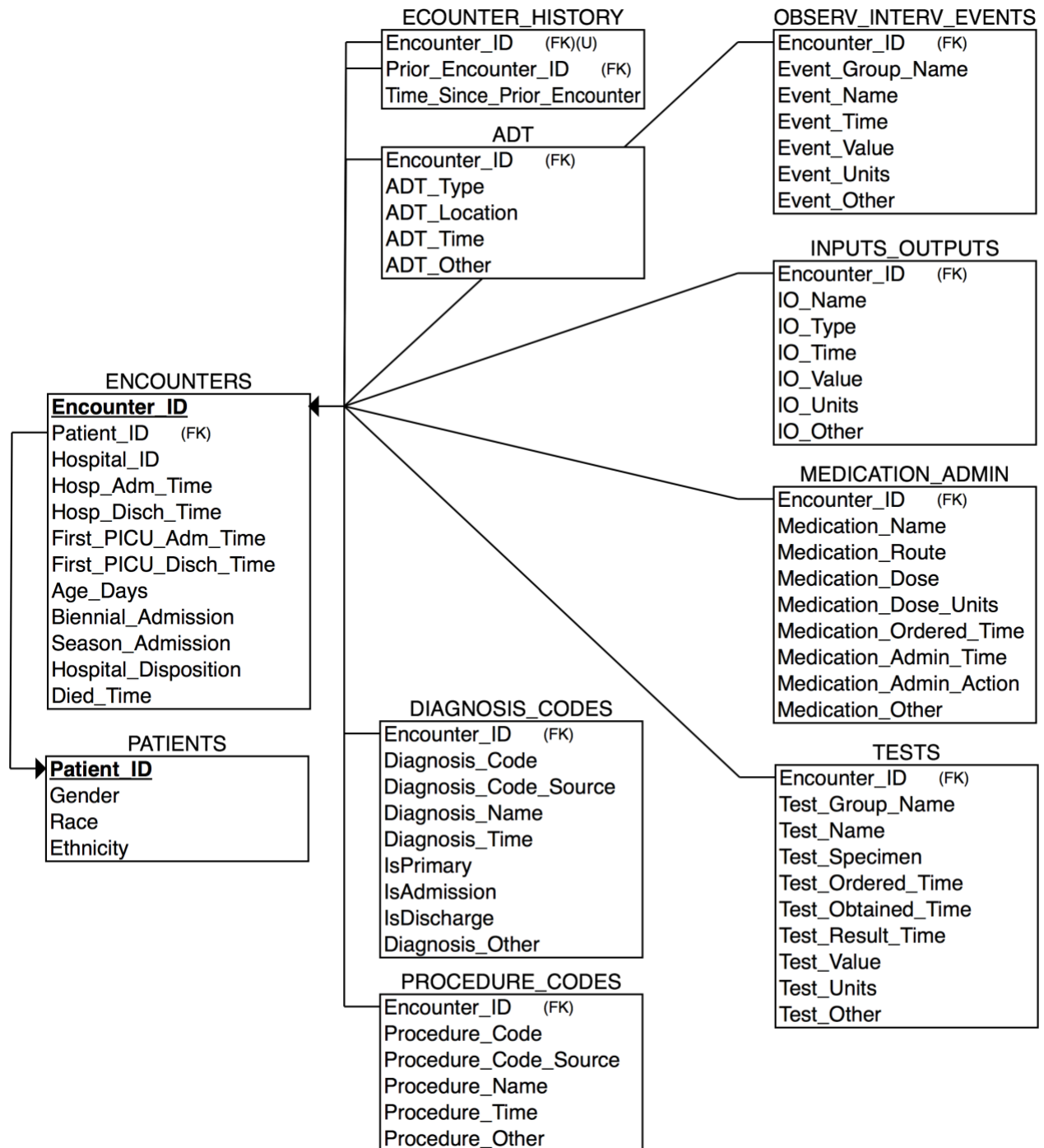
Questions?

Please contact the study Principal Investigator with any questions related to this document (and, please, CC the principal investigator from your site):

L. Nelson Sanchez-Pinto, MD, MBI:
lsanchezpinto@luriechildrens.org

COMMON DATA MODEL

Relational schema



Tables

Description: The common data model consists of 10 tables in a relational schema.

Data to be included in each table: For the first 6 tables (ENCOUNTERS, PATIENTS, ENCOUNTER_HISTORY, ADT, DIAGNOSIS_CODES, and PROCEDURE_CODES) all information associated with the patient encounters meeting inclusion criteria should be included. For the other 4 tables (OBSERV_INTERV_EVENTS, INPUTS_OUTPUTS, MEDICATION_ADMIN and TESTS) only the **data elements listed** (see next section) that were recorded during the **study time-window** should be included. The study time-window **starts** 24 hours (or 1,440 minutes) **before** the *First_PICU_Adm_Time* (if any information available before the PICU admission; see below for definition) and **ends** 28 days (or 40,320 minutes) **after** the *First_PICU_Adm_Time* or upon hospital discharge or death, whichever comes first.

Times: Since this is a de-identified database, times will be recorded in minutes from *First_PICU_Adm_Time*, as either a negative integer (for event occurring before the PICU admission), a positive integer (for events after the PICU admission) or zero if the event occurred at the same time of the PICU admission (unlikely). If the most granular timestamp for an event is a date without hours and minutes, this date should be assumed to happen at midnight when the dates starts, e.g. an event occurring on 1/1/2012, would be assumed to have happened on 1/1/2012 at 00:00 to calculate the minutes from the *First_PICU_Adm_Time*.

Data types: To ensure data consistency during query process, we recommend all data types to be **VARCHAR(40)** except for the “_Time” columns that should be **INT**, and the “_Value” and “_Other” columns that should be **VARCHAR(255)** format. Numeric data types should be avoided as they can coerce real values like “< 5” to “Null”.

Output files: Output files should be in **Comma Separated Value (.csv)** format with **UTF-8 character encoding**. Each site will provide a total of 10 files, one for each table in the schema. Files should include the name of the originating site after a “_” separator, for example: “Encounters_Lurie.csv” or “Tests_Lurie.csv”. Ensure that none of the row values contain commas, particularly the “_Value” and “_Other” columns, and if they do, that these commas are erased from the text in which they appear prior to creating the output file.

Data transfer: The 10 files with de-identified data will be uploaded to a site-specific subfolder in a study-specific folder within the secured, HIPPA-compliant File Sharing System administered by Information Management at Ann & Robert H. Lurie Children’s Hospital of Chicago. Access to each site-specific subfolder will only be granted to the site data analyst, site PI, study PI and study lead data architect at Lurie.

Table descriptions:

1. ENCOUNTERS

This is a wide format table, i.e. each row contains the information of one unique patient encounter. For example, if there are 3,000 encounters meeting criteria in a given site, this table should contain 3,000 rows. It has the following columns:

- 1.1. *Encounter_ID* (Primary Key): This is a randomly-generated, unique study encounter identifier. This identifies a hospitalization in which a patient had at least one PICU admission. These random IDs are generated by each site.
- 1.2. *Patient_ID* (Foreign Key): See PATIENTS table for definition. One patient may have multiple encounters (i.e. multiple hospitalizations with at least one PICU admission).
- 1.3. *Hospital_ID*: This is a randomly-generated, unique hospital identifier. Hospital identifiers will be used to adjust for site as a confounder in mixed models. Information relating to the identity of any specific hospital in association with one or more specific patients and their information will not be disclosed nor published. Please request random Hospital_ID number for your site directly by email from the study PI (lsanchezpinto@luriechildrens.org).
- 1.4. *Hosp_Adm_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the hospital admission date/time. This time will be either a negative integer or zero if the patient was admitted directly to the PICU.
- 1.5. *Hosp_Disch_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the hospital discharge date/time. This time will be a positive integer.
- 1.6. *First_PICU_Adm_Time*: This is the start of the first PICU admission in the hospital encounter. This time will always be zero and is used as a checksum.
- 1.7. *First_PICU_Disch_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the first PICU discharge date/time. A patient being transferred to an inpatient ward or being discharged home, for example, are treated the same for this time stamp. This time will be a positive integer.
- 1.8. *Age_Days*: This is the age in days of the patient when the first PICU admission in the hospital encounter started. This number should be rounded *down* to the nearest day.
- 1.9. *Biennial_Admission*: This is the two-year time-window of admission when the first PICU admission in the hospital encounter started. There are 3 options:
 - 1.9.1. "2012-2013"
 - 1.9.2. "2014-2015"
 - 1.9.3. "2016-2017"
- 1.10. *Season_Admission*: This is the season corresponding to the month when the first PICU admission in the hospital encounter started, based on the meteorological definition. There are 4 options:
 - 1.10.1. "Spring" - if PICU admission started in March, April, or May
 - 1.10.2. "Summer" - June, July, or August
 - 1.10.3. "Fall" - September, October, or November

- 1.10.4. "Winter" - December, January, or February
- 1.11. *Hospital_Disposition*: This is the destination of the patient after discharge from the hospitalization. There are 5 options:
 - 1.11.1. "Discharged to Home"
 - 1.11.2. "Morgue/Expired"
 - 1.11.3. "Transfer to another hospital"
 - 1.11.4. "Transfer to nursing home/rehab"
 - 1.11.5. "Discharge to Hospice"
- 1.12. *Died_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date and time of death. This time will be a positive integer or "Null" if patient didn't die during the hospitalization.

2. PATIENTS

This is a wide format table, i.e. each row contains the information of one unique patient. For example, if there are 2,500 patients meeting criteria in a given site, this table should contain 2,500 rows. It has the following columns:

- 2.1. *Patient_ID* (Primary Key): This is a randomly-generated, unique study patient identifier. These random IDs are generated by each site.
- 2.2. *Gender*: This is the patient's gender. There are 2 options:
 - 2.2.1. "Male"
 - 2.2.2. "Female"
- 2.3. *Race*: This is the patient's race. There are 4 options:
 - 2.3.1. "White"
 - 2.3.2. "Black"
 - 2.3.3. "Asian"
 - 2.3.4. "Other/Unknown"
- 2.4. *Ethnicity*: This is the patient's ethnicity. There are 3 options:
 - 2.4.1. "Hispanic"
 - 2.4.2. "Non-Hispanic"
 - 2.4.3. "Unknown"

3. ENCOUNTER_HISTORY

This is a wide format table, i.e. each row contains the information of one unique patient encounter. For example, if there are 3,000 encounters meeting criteria in a given site, this table should contain 3,000 rows. It has the following columns:

- 3.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 3.2. *Prior_Encounter_ID* (Foreign Key): This is the *Encounter_ID* of the one hospital encounter with a PICU admission meeting criteria for the study that the patient had immediately prior to this one. A large portion of patients will have a "Null" here since no prior admissions meeting criteria will be identified.
- 3.3. *Time_Since_Prior_Admission*: This is the time in minutes from the date/time used to derive the *Hosp_Disch_Time* of the prior encounter to the date/time used to derive the *First_PICU_Adm_Time*. If the *Prior_Encounter_ID* was "Null" this should also be "Null".

4. ADT

ADT stands for “Admission, Transfer, Discharge”. It should include any movement of the patient within the hospital, including any stay in the Emergency Department prior to admission to a ward or PICU. This is a long-format table, i.e. each row contains the information of one ADT event for one patient encounter, and one patient encounter may have multiple ADT events. For example, if there are 3,000 encounters and each encounter has exactly 5 ADT events, this table should contain 15,000 rows. It has the following columns:

- 4.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 4.2. *ADT_Type*: This is the type of ADT event. There are 3 options:
 - 4.2.1. “Admitted to”
 - 4.2.2. “Transfer to”
 - 4.2.3. “Discharged”
- 4.3. *ADT_Location*: This is the new location of the patients after the event of the type *ADT_Type* takes place. There are 7 options:
 - 4.3.1. “Emergency Department”
 - 4.3.2. “Operating Room” – Including catheterization, interventional radiology.
 - 4.3.3. “Inpatient” – This is any type of inpatient ward.
 - 4.3.4. “Inpatient Intermediate Care” – This is any type of intermediate care unit.
 - 4.3.5. “Other ICU” – This includes neonatal ICU, cardiac ICU, etc.
 - 4.3.6. “PICU”
 - 4.3.7. “Other” – Such as a day infusion center, imaging center, etc.
- 4.4. *ADT_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time of the ADT event. This time will be a positive integer or zero if the patient was admitted directly to the PICU.
- 4.5. *ADT_Other*: This column is used to include any additional information deemed necessary to analyze or understand the ADT data in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table from the original table separated by a “_” is also accepted as a value in this column if necessary.

5. DIAGNOSIS_CODES

This is a long-format table, i.e. each row contains the information for each diagnosis code contained in a patient encounter, and one patient encounter may have multiple diagnosis codes. For example, if there are 3,000 encounters and each encounter has exactly 10 diagnosis codes associated, this table should contain 30,000 rows. It has the following columns:

- 5.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 5.2. *Diagnosis_Code*: The diagnosis code, usually a numeric or alpha-numeric string.
- 5.3. *Diagnosis_Code_Source*: The source, vocabulary, or ontology the code belongs to. For example this might be “ICD-9-CM Diagnosis Code”, “ICD-10”, “SNOMED_CT”,...
- 5.4. *Diagnosis_Name*: This is the description of the diagnosis associated with the code, if available. For example, “Septic pulmonary embolism”.

- 5.5. *Diagnosis_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the diagnosis was recorded, if available. This might be a negative integer, positive integer, a zero, or “Null” if unknown.
- 5.6. *IsPrimary*: This is a “Yes” or “No” if the diagnosis is or is not denoted as the primary diagnosis, or “Null” if unknown.
- 5.7. *IsAdmission*: This is a “Yes” or “No” if the diagnosis is or is not denoted as the admission diagnosis, or “Null” if unknown.
- 5.8. *IsDischarge*: This is a “Yes” or “No” if the diagnosis is or is not denoted as the discharge diagnosis, or “Null” if unknown.
- 5.9. *Diagnosis_Other*: This column is used to include any additional information deemed necessary to analyze or understand the diagnosis data in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table separated by a “_” is also accepted as a value in this column if necessary.

6. PROCEDURE_CODES

This is a long-format table, i.e. each row contains the information for each procedure code contained in a patient encounter, and one patient encounter may have multiple procedure codes. For example, if there are 3,000 encounters and each encounter has exactly 20 procedure codes associated, this table should contain 60,000 rows. It has the following columns:

- 6.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 6.2. *Procedure_Code*: The procedure code, usually a numeric or alpha-numeric string.
- 6.3. *Procedure_Code_Source*: The source, vocabulary, or ontology the code belongs to. For example this might be “ICD-9-CM Procedure Code”, “ICD-10”, “CPT”,...
- 6.4. *Procedure_Name*: This is the description of the procedure associated with the code, if available. For example, “Excision procedure on the appendix”.
- 6.5. *Procedure_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the procedure was recorded, if available. This might be a negative integer, positive integer, a zero, or “Null” if unknown.
- 6.6. *Procedure_Other*: This column is used to include any additional information deemed necessary to analyze or understand the procedure data in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table separated by a “_” is also accepted as a value in this column if necessary.

7. OBSERV_INTERV_EVENTS

This table contains the observation and intervention events that a patient may have during a hospitalization. This will largely contain nursing-charted assessments, including vital signs and status of lines and devices (i.e. observations) and nursing- and respiratory therapy-charted organ-support interventions including oxygen therapy and ventilation, extra-corporeal life support, dialysis, and other interventions. See the Data Elements section for

details regarding the data to be included in this table. In many EHR implementations and mirror database versions in enterprise data warehouses this information will be contained in the so-called “Flowsheets”. This table excludes inputs/outputs and medication administration (see below). This is a long-format table, i.e. each row contains the information for each observation or intervention event meeting inclusion criteria contained in a patient encounter, and one patient encounter may have multiple observation or intervention events. For example, if there are 3,000 encounters and each encounter has exactly 100 observations or interventions associated, this table should contain 300,000 rows. It has the following columns:

- 7.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 7.2. *Event_Group_Name*: This the grouper of the *Event_Names*, if available. For example, this might be “Pupillary Response Assessment” or “Ventilator Settings”, which groups the components of the pupillary exam or the different ventilator settings, denoted in the *Event_Name* column.
- 7.3. *Event_Name*: This is the description of the most granular observation or intervention, for example “Left Pupil Response” for a pupillary exam, or “PEEP” for a ventilator setting.
- 7.4. *Event_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the observation or intervention was recorded. This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.
- 7.5. *Event_Value*: This is the actual value of the observation or intervention in string/varchar format. For example, “Brisk” for Left Pupil Response, or “8” for PEEP.
- 7.6. *Event_Units*: This the unit of measurement corresponding to the value recorded, if any. For example, “cmH2O” for PEEP.
- 7.7. *Event_Other*: This column is used to include any additional information deemed necessary to analyze or understand the observation or intervention data in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table separated by a “_” is also accepted as a value in this column if necessary. **For example, including the original name of the event in the site’s database may help clarify some specifics of the event being recorded.**

8. INPUTS_OUTPUTS

This table contains the inputs (such as fluid, blood products, TPN, enteral nutrition, medication volumes, etc.) and outputs (such as urine, drain outputs, emesis, stool, etc.) that a patient may have during a hospitalization. This table excludes actual medication administration (see below), but may contain the associated volume of fluid the medication represented. This is a long-format table, i.e. each row contains the information for each input or output event meeting inclusion criteria contained in a patient encounter, and one patient encounter may have multiple input and output events. For example, if there are 3,000 encounters and each encounter has exactly 100 input and output events associated, this table should contain 300,000 rows. It has the following columns:

- 8.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.

- 8.2. *IO_Name*: This is the description of the input or output event, for example “D5 0.9%NS” or “Urine -Foley”.
- 8.3. *IO_Type*: This is the type of event (either “Intake” or “Output”) and route (e.g. IV, PO, GT, etc. for intake, or Foley, Drain, etc. for output), if recorded. A concatenation of two or more columns separated by “_” is acceptable as a value here, if needed.
- 8.4. *IO_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the input or output event was recorded. This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.
- 8.5. *IO_Value*: This is the actual value of the input or output event in string/varchar format. For example, “10” or “Large emesis”.
- 8.6. *IO_Units*: This the unit of measurement corresponding to the value recorded, if any. For example, “ml”.
- 8.7. *IO_Other*: This column is used to include any additional information deemed necessary to analyze or understand the input or output event data in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table separated by a “_” is also accepted as a value in this column if necessary.

9. MEDICATION_ADMIN

This table contains the medications given to the patient during the hospitalization. This information is usually recorded in the Medication Administration Record (MAR) or similar table. This does not include information on prescriptions preceding or following the hospitalization. This is a long-format table, i.e. each row contains the information for each medication administration event meeting inclusion criteria contained in a patient encounter, and one patient encounter may have multiple medication administration events. For example, if there are 3,000 encounters and each encounter has exactly 500 medication administration events associated, this table should contain 1,500,000 rows. It has the following columns:

- 9.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 9.2. *Medication_Name*: This is the name of the medication being administered, for example, “PIPERACILLIN-TAZOBACTAM 80-10 MG/ML SYRINGE PUMP”. Of note, this should be the column with a medication name that has the largest number of unique values (usually medication administration tables have many columns containing different versions of a medication name, and many of these columns will have many “Null” values, for example in the case of medications mixed by the hospital pharmacy, which will only show up in one column and as “Null” in the other columns).
- 9.3. *Medication_Route*: This is the route of administration, if recorded, for example: oral, IV/intravenous, IH/inhaled, GT/G-Tube, etc.
- 9.4. *Medication_Dose*: This is the actual dose of the medication, for example “10”.
- 9.5. *Medication_Dose_Units*: This is the unit of measurement of the medication, for example “mg” or “mg/kg/hr” for some continuous medications

- 9.6. *Medication_Ordered_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the medication was ordered, if available. This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.
- 9.7. *Medication_Admin_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the medications administration event was recorded. This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.
- 9.8. *Medication_Admin_Action*: This the action that happened during the administration event, for example: "Given", "Dose Change", "New Bag", "Not given", etc.
- 9.9. *Medication_Other*: This column is used to include any additional information deemed necessary to analyze or understand the medicatrion administration event in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table separated by a "_" is also accepted as a value in this column if necessary.

10. TESTS

This table contains the tests obtained in the patient during the hospitalization, including laboratory and microbiological tests. This information is usually recorded in a laboratory information system and in the EHR as Laboratory Results or something similar. This is a long-format table, i.e. each row contains the information for each laboratory or microbiological test meeting inclusion criteria contained in a patient encounter, and one patient encounter may have multiple tests. For example, if there are 3,000 encounters and each encounter has exactly 200 tests associated, this table should contain 600,000 rows. It has the following columns:

- 10.1. *Encounter_ID* (Foreign Key): See ENCOUNTERS table for definition.
- 10.2. *Test_Group_Name*: This the grouper of the *Test_Name*, if available. This is sometimes referred to as Procedure Name. For example, this might be "Complete Blood Count" or "Respiratory Viral Panel", which groups the components of the CBC or the RVP, denoted in the *Test_Name* column.
- 10.3. *Test_Name*: This is the description of the most granular test, sometimes referred to as Component Name. For example, "WBC" or "SODIUM".
- 10.4. *Test_Specimen*: This is the type of specimen in which the test was done, if available, for example: "Blood" or "Nasal swab"
- 10.5. *Test_Ordered_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the test order was placed, if available. This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.
- 10.6. *Test_Obtained_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the test specimen was obtained or taken from the patient. **This is the most important time to record for Tests.** This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.

- 10.7. *Test_Result_Time*: This is the time in minutes from the start of the first PICU admission in the hospital encounter to the date/time when the test result was released to the clinicians, if available. This might be a negative integer, positive integer, or a zero if it occurred right at the time of PICU admission.
- 10.8. *Test_Value*: This is the actual value of the test in string/varchar format. For example, "10.4" or "Positive".
- 10.9. *Test_Units*: This the unit of measurement for the value, if available, for example: "K/uL".
- 10.10. *Test_Other*: This column is used to include any additional information deemed necessary to analyze or understand the test result data in the corresponding row by the extracting data analyst or site PI. Concatenation of several column values from the original table separated by a "_" is also accepted as a value in this column if necessary. **For example, including the original name of the test** in the site's database may help clarify source of the sample and type of test performed.

DATA ELEMENTS

Data to be included in each table: For the first 6 tables (ENCOUNTERS, PATIENTS, ENCOUNTER_HISTORY, ADT, DIAGNOSIS_CODES, and PROCEDURE_CODES) all information associated with the patient encounters meeting inclusion criteria should be included. For the other 4 tables (OBSERV_INTERV_EVENTS, INPUTS_OUTPUTS, MEDICATION_ADMIN and TESTS) only the **data elements listed** (see below) that were recorded during the **study time-window** should be included. The study time-window **starts** 24 hours (or 1,440 minutes) **before** the *First_PICU_Adm_Time* (if any information available before the PICU admission; see below for definition) and **ends** 28 days (or 40,320 minutes) **after** the *First_PICU_Adm_Time* or upon hospital discharge or death, whichever comes first.

Data elements in each table:

1. ENCOUNTERS:
 - 1.1. All data elements relating to the ENCOUNTERS table for the **duration of the hospital encounter** should be included.
2. PATIENTS:
 - 2.1. All data elements relating to the PATIENTS table for the duration of the hospital encounter should be included.
3. ENCOUNTER_HISTORY:
 - 3.1. All data elements relating to the ENCOUNTER_HISTORY table for **any hospital encounters** should be included.
4. ADT:
 - 4.1. All data elements relating to the ADT table for the **duration of the hospital encounter** should be included.
5. DIAGNOSIS_CODES:
 - 5.1. All data elements relating to the DIAGNOSIS_CODES table for the **duration of the hospital encounter** should be included.
6. PROCEDURE_CODES:
 - 6.1. All data elements relating to the PROCEDURE_CODES table for the **duration of the hospital encounter** should be included.
7. OBSERV_INTERV_EVENTS:
 - 7.1. All data elements relating to the observations and intervention events **listed below** that were recorded **during the study time-window** (-24 hours to +28 days/hospital discharge/death from PICU start time) should be included.
 - 7.2. The list below includes some general estimates of expected rates:

EVENT_NAME	UNITS	COMMENTS	IN % ENC. WITH	COUNT /ENC. WITH
PULSE	BPM	Heart rate, pulse rate.	>99%	150-250
RESP_RATE	BPM	Respiratory rate, RR, AWRR. If AWRR is used, please include this in the Events_Other column.	>99%	100-250
PULSE_OX	%	Pulse oximetry, SpO2.	>99%	100-250
TEMP	Deg. C	Temperature. If recorded in degrees Fahrenheit, may use this equation to transform: $T(^{\circ}\text{C}) = (T(^{\circ}\text{F}) - 32) \times (5/9)$. If transformin, please include this in Event_Other	>99%	80-120
SBP_CUFF	mmHg	Systolic blood pressure from non-invasive or cuff device. SBP and DBP may be recorded under "BP" or "Blood Pressure" in XXX/XX format. If so, record as is here.	>99%	100-200
DBP_CUFF	mmHg	Diastolic blood pressure from non-invasive or cuff device	>99%	100-200
MAP_CUFF	mmHg	Mean arterial pressure calculated from non-invasive or cuff device	>99%	100-200
SBP_ART	mmHg	Systolic blood pressure from arterial line. SBP and DBP may be recorded under BP or Blood Pressure as "XXX/XX". If so, record as is here.	30-40%	100-250
DBP_ART	mmHg	Diastolic blood pressure from arterial line.	30-40%	100-250
MAP_ART	mmHg	Mean arterial pressure from arterial line.	30-40%	100-250
HEIGHT	Cm	Height	40-60%	1-2
WEIGHT	Kg	Weight, daily weight	>90%	2-10
ETCO2	mmHg	EtCO2, end tidal carbon dioxide.	40-60%	100-200
CVP	mmHg	Central Venous Pressure	15-30%	50-150
ICP	mmHg	Intracranial pressure	2-5%	50-120
GCS_EYE		Eye response of Glasgow Coma Scale (GCS)	>99%	50-120
GCS_MOTOR		Motor response of GCS	>99%	50-120
GCS_VERBAL		Verbal response of GCS	>99%	50-120
GCS_TOTAL		Total GCS score		
PUPIL_RESP_R		Right pupillary response or reaction, qualitative	>99%	50-120
PUPIL_SIZE_R	mm	Right pupillary size	>99%	50-120
PUPIL_RESP_L		Left pupillary response or reaction, qualitative	>99%	50-120
PUPIL_SIZE_L	mm	Left pupillary size	>99%	50-120
NEURO_LOC		Neurologic assessment of level of consciousness, qualitative. Recorded in lieu of GCS in some sites	>99% (alt.)	50-120 (alt.)
NEURO_ACTIVITY		Neurologic assessment of level of activity, qualitative. Recorded in lieu of GCS in some sites	>99% (alt.)	50-120 (alt.)
AVPU		Neurologic assessment of Alert Verbal Pain Unresponsive level. Recorded in lieu of GCS in some sites	>99% (alt.)	50-120 (alt.)
O2_MODE		Oxygen mode, oxygen device	60-100%	50-150
O2_FLOW	LPM	O2 Flow, oxygen flow	60-90%	50-150
FIO2	%	FiO2, fraction of inspired oxygen. If source of oxygen specified as ventilator vs. non-	40-60%	40-100

		ventilator/NIV, please include in following 2 elements.		
FIO2_VENT	%	FiO2 from ventilator	40-60%	40-100
FIO2_NON_INV	%	FiO2 from non-invasive device	60-90%	50-150
VENT_MODE		Mechanical ventilation mode	40-60%	40-100
PEEP_VENT	cmH2O	Positive end-expiratory pressure set on the ventilator.	40-60%	40-100
PIP_VENT	cmH2O	Peak inspiratory pressure set on the ventilator.	40-60%	30-50
PS_VENT	cmH2O	Pressure support set on the ventilator.	40-60%	40-100
TV_EXP	ml	Expiratory tidal volume observed	30-40%	20-50
TV_SET	ml	Tidal volume set on the ventilator.	30-40%	40-80
VENT_RATE		Ventilator rate set on the ventilator.	40-60%	40-100
IT_VENT	seconds	Inspiratory time set on the ventilator.	40-60%	30-50
MAP_VENT	cmH2O	Mean airway pressure set on the ventilator.	40-60%	30-50
AMP_HFOV		Amplitude set on high frequency oscillatory ventilation, Delta P	1-4%	50-120
MAP_HFOV	cmH2O	Mean airway pressure set on high frequency oscillatory ventilation, might be recorded as MAP_VENT	1-4%	50-120
FREQ_VENT	Hz	Frequency set on high frequency oscillatory ventilation.	1-4%	50-120
NIV_MODE		Non-invasive mode of ventilation, might be recorded under O2_MODE	10-20%	15-20
IPAP_NIV	cmH2O	Inspiratory positive airway pressure set on non-invasive ventilator mode.	10-20%	15-20
EPAP_NIV	cmH2O	Expiratory positive airway pressure set on non-invasive ventilator mode.	10-20%	15-20
NITRIC_OX	PPM	Nitric oxide level administered inhaled.	1-5%	50-120
AIRWAY_TYPE		Type of airway (e.g. natural, tracheostomy, oral ETT, nasal ETT)	40-60%	40-100
FOLEY_STATUS		Foley catheter, check, assessment, status	20-40%	50-80
CVC_TYPE		Central venous catheter type (e.g. lumens, size, PICC)	30-40%	100-250
CVC_LOC		CVC location	30-40%	100-250
CVC_STATUS		CVC check, assessment, status, dressing	30-40%	100-250
CHEST_TUBE_LOC		Chest tube location	5-10%	80-100
CHEST_TUBE_STATUS		Chest tube check, assessment, status, dressing		
ART_LINE_LOC		Arterial line location	30-40%	150-200
ART_LINE_STATUS		Arterial line check, assessment, status, dressing	30-40%	150-200
ECMO_TYPE*		Extra corporeal membrane oxygenation (ECMO) type (e.g. VV, VA)	<2%	100-250
ECMO_FLOW	ml/min	ECMO flow	<2%	100-250
CRRT_TYPE*		Continuous renal replacement therapy (e.g. CVVH, CVVHD)	2-5%	50-120
CRRT_NET_UF	ml/h	CRRT net ultrafiltration rate	2-5%	50-120
PLEX_TYPE*		Plasma exchange type	<2%	10-20
PLEX_VOLUME	ml/kg	Plasma exchange volume exchange performed	<2%	10-20

*For ECMO, CRRT, and PLEX the primary goal is to determine the type of intervention performed and the duration. The second data element for each of these three types of interventions will help determine its duration by including a time stamp for each nursing or specialist-based assessment of the corresponding value. **IN % ENC.** **WITH** refers to the % of patient encounters in the cohort of patients meeting inclusion criteria for the study who

have at least one recording of the event (i.e. count of patient encounters that had at least one value recorded divided by the total number of patient encounters meeting inclusion criteria multiplied by 100). **COUNT /ENC. WITH** refers to the number of events recorded within the study time window in patient encounters who had at least one value of the event recorded (i.e. count of all individual values of the event divided by count of patient encounters that had at least one value recorded). The last 2 columns of this table may be used as a quality check to ensure that the right data elements are being extracted.

8. INPUTS_OUTPUTS:

- 8.1. All data elements relating to the inputs and outputs events that were recorded **during the study time-window** (-24 hours to +28 days/hospital discharge/death from PICU start time) should be included.
- 8.2. This includes inputs (such as fluids, blood products, TPN, enteral nutrition, medication volumes, etc.) and outputs (such as urine, drain outputs, emesis, stool, etc.) recorded in volumes (using **milliliters** as the unit of measurement) or occurrences (e.g. “once”). Completeness is necessary to calculate daily fluid balances.

9. MEDICATION_ADMIN:

- 9.1. All data elements relating to medication administration events that were recorded **during the study time-window** (-24 hours to +28 days/hospital discharge/death from PICU start time) should be included.
- 9.2. This includes all medications administered during the study time window. Completeness is necessary to adjust for medication confounding in analyses.

10. TESTS:

- 10.1. All data elements relating to the test events **listed below** that were recorded **during the study time-window** (-24 hours to +28 days/hospital discharge/death from PICU start time) should be included.
- 10.2. These tests are the laboratory and microbiological tests performed during the time-window.
- 10.3. The list below includes some general estimates of expected values and rates:

TEST_NAME	UNITS	COMMENTS	MEDIAN (IQR)	IN % ENC. WITH	COUNT PER PT. WITH
WBC	K/uL	White blood cell count from a blood specimen.	9 (4.5-14)	>90%	6-10
HGB	g/dL	Blood hemoglobin concentration. May include values from point-of-care or blood gas measurements but please specify this in the Test_Other column.	9.8 (8.5-11.3)	>90%	6-10
PLTS	K/uL	Platelet count.	160 (65-300)	>90%	6-10
RDW	%	Red blood cell distribution width.	15.5 (14.2-17)	>90%	6-10
MPV	fL	Mean platelet volume.	9 (8-10)	>90%	6-10
MCH	pg	Mean corpuscular hemoglobin.	29 (27.6-30.2)	>90%	6-10
MCV	fL	Mean corpuscular volume.	86 (82-90)	>90%	6-10

ANC	K/uL	Absolute neutrophil count from blood sample. If count not available, "neutrophils (%)" is acceptable but please specify this in Test_Other.	5.8 (2.5-10)	>90%	5-10
ALC	K/uL	Absolute lymphocyte count from blood sample. If count not available, "lymphocyte (%)" is acceptable but please specify this in Test_Other.	1.5 (0.7-3)	>90%	5-10
AMC	K/uL	Absolute monocyte count from blood sample. If count not available, "monocyte (%)" is acceptable but please specify this in Test_Other.	0.6 (0.3-1)	>90%	5-10
AES	K/uL	Absolute eosinophil count from blood sample. If count not available, "eosinophil (%)" is acceptable but please specify this in Test_Other.	0.1 (0-0.3)	65-95%	3-7
BANDS	%	Percent neutrophil bands in blood sample. If percentage not available, a bands absolute count is acceptable but please specify this in Test_Other.	5 (0-15)	10-60%	2-6
RETIC	%	Percent reticulocyte in blood sample. If percentage not available, a reticulocyte absolute count is acceptable but please specify this in Test_Other.	2.4 (1-5.5)	10-20%	2-3
ABO_RH		Blood type including ABO and RH, e.g. "A+", "AB-"	Other	40-60%	1-2
SODIUM	mEq/L	Sodium, Na, or NA+ from serum* sample. May include values from point-of-care or blood gas measurements but please specify this in Test_Other.	139 (136-144)	>90%	10-30
POTASSIUM	mEq/L	Potassium, K, or K+ from serum sample. May include values from point-of-care or blood gas measurements but please specify this in Test_Other.	3.7 (3.2-4.2)	>90%	10-30
CHLORIDE	mEq/L	Chloride from serum sample	104 (100-109)	>90%	8-12
BICARBONATE	mEq/L	Carbon dioxide, CO2, Total CO2, HCO3- from serum sample. May include values from point-of-care or blood gas measurements but please specify this in Test_Other.	25 (22-29)	>90%	10-30
BUN	mg/dL	Blood urea nitrogen from serum	11 (6-21)	>90%	8-12
CREATININE	mg/dL	Creatinine level from serum	0.4 (0.2-0.6)	>90%	8-12
GLUCOSE	mg/dL	Glucose level from serum or plasmacolo. May include values from point-of-care or blood gas measurements but please specify this in Test_Other.	118 (97-153)	>90%	15-30
CALCIUM_TOT	mg/dL	Total calcium level from serum.	8.7 (8.2-9.3)	>90%	8-12
CALCIUM_ION	mmol/L	Ionized calcium, ical, Ca2+, Ca++ from serum. May include values from point-of-care or blood gas measurements but please specify this in Test_Other.	1.2 (1.1-1.3)	50-70%	10-30

MAGNESIUM	mg/dL	Magnesium level in serum.	1.9 (1.7-2.2)	55-85%	8-10
PHOSPHORUS	mg/dL	Phosphorus level in serum.	4 (3.1-4.9)	55-85%	8-10
ALT	IU/L	Alanine transaminase (ALT) or serum glutamic-pyruvic transaminase (SGPT)	35 (18-90)	55-75%	2-6
AST	IU/L	Aspartate aminotransferase (AST) or serum glutamic oxaloacetic transaminase (SGOT) from serum.	44 (26-95)	55-75%	2-6
BILIRUBIN_TOT	mg/dL	Total bilirubin level in serum.	0.6 (0.3-2)	55-75%	2-6
BILIRUBIN_DIR	mg/dL	Direct bilirubin or conjugated bilirubin level in serum.	0.5 (0.1-3)	20-60%	3-5
ALBUMIN	g/dL	Albumin level is serum.	3 (2.6-3.5)	55-75%	5-10
PROTEIN	g/dL	Total protein level in serum.	5.5 (4.8-6.4)	55-75%	2-6
ALP	IU/L	Alkaline phosphatase level in serum.	135 (90-215)	55-75%	2-6
GGT	IU/L	Gamma-glutamyl transferase level in serum.	125 (53-300)	5-15%	4-6
LDH	IU/L	Lactate dehydrogenase or lactic dehydrogenase level in serum.	390 (258-730)	5-15%	1-2
INR		International normalized ratio of PT, or PT/INR in plasma.	1.2 (1.1-1.6)	40-60%	4-8
PT	Seconds	Prothrombin time in plasma	15 (13-17)	40-60%	4-8
PTT	Seconds	Partial thromboplastin time in plasma	37 (30-50)	40-60%	4-8
D_DIMER	mcg/mL FEU	D dimer level in plasma.	5.5 (2.5-13)	5-40%	3-5
FIBRINOGEN	mg/dL	Fibrinogen level in plasma	270 (190-415)	10-40%	3-5
LACTATE	mmol/L	Lactate or lactic acid level in serum. If recorded in mg/dL divide by 9.01 to convert to mmol/L. May include values from point-of-care or blood gas measurements but please specify this in Test_Other.	1.3 (0.8-2.2)	40-70%	10-30
CRP	mg/L	C-reactive protein level in serum. If recorded in nmol/L multiply by 9.524 to convert to mg/L.	4.2 (1-13)	10-60%	1-5
ESR	mm/h	Erythrocyte sedimentation rate	40 (17-75)	5-20%	1-2
FERRITIN	ng/mL	Ferritin level in serum	680 (170-3500)	5-15%	1-2
PROCALCITONIN	ng/ml	Procalcitonin level in serum	0.5 (0.1-10)	0-50%	1-2
AMYLASE	IU/L	Amylase level in serum	60 (30-140)	10-20%	1-3
LIPASE	IU/L	Lipase level in serum	40 (17-130)	10-20%	1-3
CORTISOL	mcg/dL	Cortisol level in serum.	19 (8.4-35)	8-12%	1-2
PH_ART		pH in arterial blood gas sample	7.4 (7.3-7.4)	30-50%	20-30
PO2_ART	mmHg	pO2 in arterial blood gas sample or PaO2	90 (70-130)	30-50%	20-30
PCO2_ART	mmHg	pCO2 in arterial blood gas sample or PaCO2	45 (38.5-53)	30-50%	20-30
PH_VEN		pH in venous blood gas sample	7.3 (7.3-7.4)	30-60%	5-10
PO2_VEN	mmHg	pO2 in venous blood gas sample or PvO2	42 (36-50)	30-60%	5-10
PCO2_VEN	mmHg	pCO2 in venous blood gas sample or PvCO2	49 (41-57)	30-60%	5-10
PH_CAP		pH in capillary blood gas sample	7.4 (7.3-7.4)	<2%	2-3

PO2_CAP	mmHg	pO2 in capillary blood gas sample or PcO2	44 (38-52)	<2%	2-3
PCO2_CAP	mmHg	pCO2 in capillary blood gas sample or PcCO2	49 (44-63)	<2%	2-3
BASE_EXC	mEq/L	Base excess or BE or SBE. Please ensure blood gas sample type (arterial, venous, etc.) is specified in Test_Other or Test_Specimen.	5 (2 -10)	30-70%	15-30
BASE_DEF	mEq/L	Base deficit or BD in blood gas sample. Please ensure blood gas sample type (arterial, venous, etc.) is specified in Test_Other or Test_Specimen. Base deficit is sometimes recorded as a part of BASE_EXC with a negative sign, e.g. "-5", which is equivalent of BASE_EXC=5	4 (2 -8)	30-70%	15-30
CULTURE		Bacterial or fungal culture. Ensure Test_Specimen specifies the source of sample. Some institutions may instead use one of the 5 categories that follow	Other	100%	4-6
CULT_BLOOD		Bacterial or fungal culture of the blood	Other	80-90%	2-4
CULT_RESP		Bacterial or fungal culture of the respiratory system. If specimen type is available (e.g. tracheal aspirate, BAL, etc.) please include.	Other	20-40%	1-2
CULT_URINE		Bacterial or fungal culture of the urine. If specimen type is available (e.g. Foley, clean catch, etc.) please include.	Other	50-70%	2-4
CULT_CSF		Bacterial or fungal culture of the cerebral spinal fluid.	Other	15-30%	2-5
CULT_OTHER		Bacterial or fungal culture of other sites, including body fluid, wound, etc. If specimen type is available (e.g. pleural, peritoneal, etc.) please include.	Other	20-30%	1-2
ADENOVIRUS		Adenovirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR) or other serum.	Other	50-70%	1-2
CORONAVIRUS		Coronavirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR). If multiple subtypes (e.g. 229E, HKU1, etc) please include in Test_Other.	Other	50-70%	1-2
HMPV		Human metapneumovirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR).	Other	50-70%	1-2
RHINO_ENTERO		Human rhinovirus/enterovirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR).	Other	50-70%	1-2
INFLUENZA_A		Human rhinovirus/enterovirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR).	Other	50-70%	1-2

		If from Rapid PCR, please specify in Test_Other			
INFLUENZA_B		Human rhinovirus/enterovirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR). If from Rapid PCR, please specify in Test_Other	Other	50-70%	1-2
RSV		Human rhinovirus/enterovirus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR). If from Rapid PCR, please specify in Test_Other	Other	50-70%	1-2
PARAINFLU		Parainfluenza virus from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR). If multiple subtypes (e.g. Type 1, Type 2, etc.) please include in Test_Other.	Other	50-70%	1-2
B_PERTUSSIS		Bordetella pertussis PCR from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR).	Other	50-70%	1-2
MYCO_PNEUM		Mycoplasma pneumoniae PCR from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR).	Other	50-70%	1-2
CHLAM_PNEUM		Chlamydia pneumoniae PCR from nasopharyngeal or tracheal sample (usually part of respiratory panel PCR).	Other	50-70%	1-2
CMV		Cytomegalovirus from serum or other sample. Please specify source in Test_Specimen. Also please specify type of analysis (e.g. IgM, PCR, quant.) in Test_Other.	Other	5-15%	1-2
EBV		Epstein Barr Virus from serum or other sample. Please specify source in Test_Specimen. Also please specify type of analysis (e.g. IgM, PCR, quant.) in Test_Other.	Other	2-15%	1-2
HSV_1		Herpes Simplex Virus Type 1 from serum or other sample. Please specify source in Test_Specimen. Also please specify type of analysis (e.g. IgM, PCR, quant.) in Test_Other.	Other	5-15%	1-2
HSV_2		Herpes Simplex Virus Type 2 from serum or other sample. Please specify source in Test_Specimen. Also please specify type of analysis (e.g. IgM, PCR, quant.) in Test_Other.	Other	5-15%	1-2
C.DIFF_STOOL		Clostridium difficile toxin test in stool	Other	10-15%	1-2
MRSA_NASAL		MRSA screen test, usually from nasal swab	Other	20-90%	1-2
CSF_GLUC	mg/dL	Glucose level in the cerebral spinal fluid (CSF)	64 (50-75)	15-30%	2-5
CSF_PROT	mg/dL	Protein level in CSF	42 (18-150)	15-30%	2-5

CSF_RBC	/cu.mm	Count of red blood cells in CSF	90 (3-1500)	15-30%	2-5
CSF_WBC	/cu.mm	Count of white blood cells in CSF	5 (1-50)	15-30%	2-5
CSF_LYMPH	%	Lymphocyte % in CSF	42 (18-70)	15-30%	2-5
CSF_MONO	%	Monocyte % in CSF	15 (6-32)	15-30%	2-5
CSF_NEUTR	%	Neutrophil % in CSF	10 (1-46)	15-30%	2-5
URINE_BLOOD		Blood present in urine sample	Other	50-70%	2-4
URINE_GLUC	mg/dL	Glucose level in urine sample, include qualitative results	Other	50-70%	2-4
URINE_KETO	mg/dL	Ketones level in urine sample, include qualitative results	Other	50-70%	2-4
URINE_LEUK		Leukocytes or leukocyte esterase in urine sample, qualitative.	Other	50-70%	2-4
URINE_NITR		Nitrites in urine sample, qualitative.	Other	50-70%	2-4
URINE_PH		pH of urine sample	Other	50-70%	2-4
URINE_PROT	mg/dL	Protein level in urine sample, include qualitative results	Other	50-70%	2-4
URINE_SPEC		Specific gravity of urine sample	Other	50-70%	2-4
URINE_RBC	/hpf	Count of red blood cells in urine sample	Other	40-60%	2-4
URINE_WBC	/hpf	Count of white blood cells in urine sample	Other	40-60%	2-4

*most samples denoted as obtained from serum in this table may also be obtained from plasma. The most common convention is used for simplicity. **MEDIAN (IQR)** refers to the median and interquartile range of values expected for each of the tests in the units of measurement specified. **IN % ENC. WITH** refers to the % of patient encounters in the cohort of patients meeting inclusion criteria for the study who have at least one recording of the test (i.e. count of patient encounters that had at least one value recorded divided by the total number of patient encounters meeting inclusion criteria multiplied by 100). **COUNT PER PT. WITH** refers to the number of results recorded within the study time window in patient encounters who had at least one value of the test recorded (i.e. count of all individual values of the test divided by count of patient encounters that had at least one value recorded). The last 3 columns of this table may be used as a quality check to ensure that the right data elements are being extracted. WARNING: for the Median (IQR) check a numeric transformation of Test_Value is needed, but for the actual table for data transfer, the Test_Value value should be in the original value in VARCHAR(255) format.