# Convalescent Plasma for Treatment of COVID-19 EHR Data Studies

Hospitals' Technical Guide for EHR Data Retrieval and Submission

## Background

Dear Colleagues:
Thank you for assisting in the data collection for this important study during the COVID-19 pandemic. You are part of an unprecedented collaboration of physicians, clinical investigators, informaticists and data scientists working across health systems, academia, and the EHR vendor community to bring the best evidence forward on the use of convalescent plasma for care of patients suffering with COVID-19. This EHR data study is a complement to the Expanded Access to Convalescent Plasma for the Treatment of Patients With COVID-19 study (https://clinicaltrials.gov/ct2/show/NCT04338360), managed by Mayo Clinic (referred herein as the "registry study"), which enrolls only patients receiving convalescent plasma, so there are no patients for a comparator group of patients not exposed to convalescent plasma. The EHR data will allow us to retrospectively approximate a control group using case matching. Using EHR data from a large number of hospitals will allow us to rapidly assemble evidence related to the efficacy and safety of convalescent plasma.

This guide should be used in conjunction with the "Convalescent Plasma for treatment of COVID-19 EHR Data Studies Master Protocol.", which has been included as an amendment to the registry study protocol.

## Contact Information

For general information regarding the registry study please contact uscovidplasma@mayo.edu.

For questions on this technical implementation guide please contact convplasma@mitre.org.

# Contents

## Change History

| Version | Date | Author/Owner | Section | Description of Change |
|---|---|---|---|---|
| **1.0** | 6/ 11/2020 | The MITRE Corporation | n/a | Initial version for publication |
| **1.1** | 6/16/2020 | The MITRE Corporation | Background | Added links to registry study protocol and amendment. |
| | | | 1 Frequently Asked Questions (FAQ) | Updated Cerner contact information on FAQ 1.5 How do I contact my Epic or Cerner EHR vendor for this study?<br><br>Added FAQ 1.7 What facilities should I include in the reported data?<br><br>Added FAQ 1.8 What patients should I include in the reported data?<br><br>Added FAQ 1.9 The details and data definitions included in the Study Protocol and the technical guide are not always aligned. Which definitions should I use? |
| | | | 2.5 Step 5. De-identify data including disguising dates | Updated date disguising procedure to make use of each individual patient admission date instead of the first CP patient admission to the facility. |
| | | | Appendix A – Data definitions) | Added admission epoch data element definition<br><br>Corrected typo in definition logic for severity of illness. |
| | | | Appendix C – Reporting File Structure | Corrected typo in allowable values for age in patient-level data file structure. |

| Version | Date | Author/Owner | Section | Description of Change |
|---|---|---|---|---|
| **1.2** | 6/17/2020 | The MITRE Corporation | 1 Frequently Asked Questions (FAQ) | Updated contact and instructions for connecting with Cerner and Epic on FAQ 1.5 How do I contact my Epic or Cerner EHR vendor for this study? |
| | | | 2.5 Step 5. De-identify data including disguising dates | Simplified date disguising procedure to define admission date as day 0, since the researchers will no longer need to rely on the date of admission of the first CP patient as an indicator of epoch. |
| | | | 2.4 Step 4. Extract data for and calculate phase 1 variables | Added instructions to create a deidentified study patient ID. |
| | | | Appendix A – Data definitions) | Added clarifications on the definition and use of admission epoch data element.<br><br>Added epoch as a matching variable.<br><br>Updated Table 1 title to clarify that it includes all phase 1 variables, and not just matching variables. |
| | | | Appendix C – Reporting File Structure | Added study patient ID as additional variable in the patient-level data file.<br><br>Added notes for the transmission of admission date (only required if using old date disguising method)<br><br>Added positive COVID-19 laboratory test variable to distinguish individuals who were included in the study population by testing vs. clinical diagnosis. |
| **1.3** | 6/25/2020 | The MITRE Corporation | 1 Frequently Asked Questions (FAQ) | Clarified that the same pool of patients should be used for Phase 1 and Phase 2 submission in FAQ 1.2 What is meant by “first phase” and “second phase” of data collection?<br><br>Added FAQ 1.10 What study entry date should be used to ascertain whether a patient should be included in the data set? |

| Version | Date | Author/Owner | Section | Description of Change |
|---|---|---|---|---|
| | | | 2 Technical Guidance on Data Collection and Transmission | Removed Section 2.1 Step 1. Determine and report first COVID-19 convalescent plasma use at the hospital and date of data extraction as this was intended to provide a sense of epoch, and epoch (the month of CP transfusion) is now included in the list of variables to transmit for each patient.<br><br>Added Section 2.3 Step 3: Determine study entry and exit dates, and date of data extraction, providing further details on how to determine whether a patient belongs in the data set. |
| | | | Appendix B - Phase 2 variables | Added discharge disposition to the list of Phase 2 variables. |

# Contributors

## National COVID-19 Convalescent Plasma Expanded Access Program/National COVID-19 Convalescent Plasma Project

**Mayo Clinic**

Michael Joyner, MD, Mayo Clinic
Rickey Carter, PhD, Mayo Clinic
Vitaly Herasevich, MD, PhD, Mayo Clinic

**Johns Hopkins University**

Arturo Casadevall, MD, MS, PhD, Johns Hopkins University

**Michigan State University**

Chenxi Li, PhD, Michigan State University
Mat Reeves, BVSc, PhD, Michigan State University
Nigel Paneth, MD, MPH, Michigan State University

For more information about the National COVID-10 Convalescent Plasma Expanded Access Program visit http://uscovidplasma.org. For more information about the National COVID-19 Convalescent Plasma Project visit http://ccpp19.org.

## COVID-19 Health Care Coalition

**Cerner Corporation**

Peter Smart, MS, Cerner Corporation
Rehan Waheed, MD, Cerner Corporation

**Epic**

Andrea Noel, MD, Epic
Janet Campbell, BA, Epic

**The MITRE Corporation**

Brian Anderson, MD, The MITRE Corporation
Francis X. Campion, MD, The MITRE Corporation
Rute Martins, MS, The MITRE Corporation

For more information about the COVID-19 Healthcare Coalition visit http://c19hcc.org.

# 1 Frequently Asked Questions (FAQ)

## 1.1 Which hospitals should provide data?

Every hospital participating in the Convalescent Plasma Expanded Access Program coordinated through the Mayo Clinic may consider participating. Those hospitals with a high volume of patients (>30 convalescent plasma treated patients) and are on either the Epic EHR or the Cerner EHR are preferred sites for contributing data in rapid fashion.

Details regarding the registry study can be found on the website https://www.uscovidplasma.org/.

## 1.2 What is meant by "first phase" and "second phase" of data collection?

The protocol mentions that time is of the essence to detect an initial signal regarding the safety of convalescent plasma in COVID-19 patients. For this reason, "phase 1" will focus on the outcomes of progression to mechanical ventilation or death adjusted only in the design, via matching. To more fully validate those findings, a second phase of the study will likely require a second, more detailed data pull, with additional data elements such as co-morbidities and laboratory test results, to permit further adjustment for potential confounders and biasing factors. The phase 2 data transmission will encompass the same patients included in the phase 1 data set (no newly eligible patients are expected to be included at this time).

## 1.3 How often should our hospital data set be updated and sent to Mayo Clinic?

Hospitals are asked to send their phase 1 data as soon possible during June 2020. Hospitals can then send the phase 2 data as soon as it is available, and no later than July 31, 2020. Based on the total number of cases reported, investigators will determine if additional data cuts will be needed after that time.

## 1.4 Who will perform the matching to controls?

The researchers at Mayo Clinic will perform the matching. We will seek to match at least one and up to four control patients for each patient who has received CP.

## 1.5 How do I contact my Epic or Cerner EHR vendor for this study?

Both Cerner and Epic have developed detailed queries and data definitions specific for their EHRs in support of this study that can be leveraged by participating hospitals. Hospitals seeking to leverage this assistance should contact their respective vendors per the instructions below.

**Instructions for Cerner EHR customers**:

- Please contact convplasma@mitre.org for assistance in reaching the appropriate Cerner contact.

**Instructions for Epic customers:**

You will want to work with your local IT team to extract data from your EHR (Epic) and transmit it to Mayo/MSU.

- **Step One:** Find a contact in the health IT department. In the same way that you might receive other IT support (for example, a support ticket or a help line), reach out to find "someone who works on the EHR." Remember that your organization may call your EHR something other than "Epic."
    - If you know the CMIO at your health system, that person may be able to help you as well.
    - It will be important for you to make this connection with IT directly. Your IT team may need to tweak the reports to match how you document in your EHR. They will also be responsible for actually running the reports.
- **Step Two:** When you reach someone on the IT team, tell them "I was told you could help me extract data that Mayo/MSU needs for reporting. They said I should contact you and ask you to contact Epic. They said to mention SLG 5197774."

**If you have are having trouble contacting the right individuals within your EHR vendor organization, please contact us at convplasma@mitre.org and we will facilitate an introduction.**

## 1.6 What format should be used for reporting data?

Data will be reporting using two .csv files. The first will include hospital-level information, including the names and addresses for each hospital for which data is being collectively reported, as well as high-level information about data extraction. The second file will include patient-level data. Appendix C – Reporting File Structure provides detailed information about the structure and content of these files.

## 1.7 What facilities should I include in the reported data?

If your organization has transfused plasma at more than one facility, please report the data for each facility. The researchers will be performing matching at the individual facility level, as opposed to the health system level. Appendix C has information about how to identify each facility and tie it to individual patient data.

## 1.8 What patients should I include in the reported data?

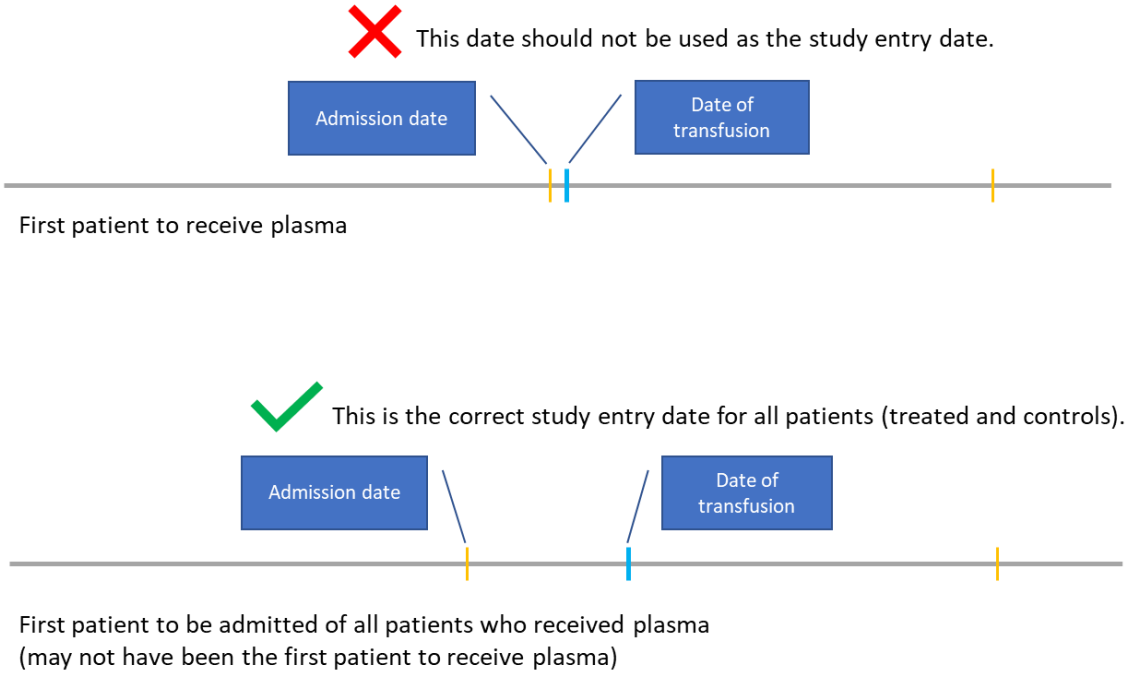
This study is a case-control study. Unlike the registry study, we're looking for data for both plasma-transfused patients and patients who did not receive convalescent plasma. Section 2 describes the steps to identify the relevant patients. Appendix A includes data definitions that can be useful to ensure consistent data extraction across hospitals. Specific variables to be reported are included in appendix C.

## 1.9 The details and data definitions included in the Study Protocol and the technical guide are not always aligned. Which definitions should I use?

The technical guide includes the most up-to-date definitions that should be used for data extraction, based on iterative discussions with researchers, EHR vendors and participating hospitals. However, the variables in the technical guide are the same as those outlined in the protocol.

## 1.10 What study entry date should be used to ascertain whether a patient should be included in the data set?

There was some ambiguity in the definition of the study start date in the protocol and earlier versions of this technical guide. The date that should be used to ascertain inclusion of *any* patient in the data set is the earliest admission of any patient who received convalescent plasma, regardless of when they received it during the hospital stay. As it turns out, that patient might not be the first to receive convalescent plasma in your hospital. You might have a patient admitted later than the patient described above, but who received convalescent plasma earlier. For example, a patient might be admitted on April 1 and receive plasma on April 7. Another patient might have been admitted on April 3 and received plasma on April 5. The first admission of someone who received plasma was April 1, even though that patient was not the first in your hospital to get plasma, and we would like that admission date – April 1 - to be the point from which you start identifying COVID-19 patients for this study. The diagram below illustrates this distinction.




Section 2.3 Step 3: Determine study entry and exit dates, and date of data extraction includes more information about study entry and exit dates.

# 2 Technical Guidance on Data Collection and Transmission

This section provides high-level information about data extraction and transmission. For complete details on data definitions, please consult Appendix A – Data definitions (for phase 1 data).

## 2.1 Step 1: Identify all COVID-19 hospitalized patients

The source population is adult patients (18 years and older) hospitalized for inpatient care associated with COVID-19 infection. This is done to ensure that we have an adequate period of follow up for all subjects. Patients with lesser periods of follow-up now will not be included in the present study but may be studied later. A COVID-19-related hospitalization includes patients who may have been diagnosed up to 14 days prior to the hospitalization, or during the hospitalization, as well as patients who may only have a clinical diagnosis of COVID-19 but no confirmatory laboratory test. We realize that lab testing may have occurred prior to admission and the lab result itself might not reside in the EHR data set from the hospital episode of care. See Appendix A – Data definitions for more details on how to identify COVID-19 hospitalized patients.

## 2.2 Step 2. Identify COVID-19 patients who have received convalescent plasma (CP) infusion

The specific implementation of blood bank software and the EHR will determine the most efficient and complete manner to detect COVID-19-convalescent plasma transfusion. Some EHRs use the International Society of Blood Transfusion "E-codes" and others use specific order-entry routines for determining CP transfusion. We are interested in those patients who received at least one unit of COVID-19 CP. You do not need to report the number of transfused units. You do not need to report antibody titers or content in the transfused unit(s).

## 2.3 Step 3: Determine study entry and exit dates, and date of data extraction

**The population will be restricted to admissions starting between the earliest date of admission for CP transfused patients in each hospital (study entry date) until 28 days before the date of data extraction (study exit date)**. This includes patients who were discharged within this timeframe, as well as patients who may still be hospitalized at on the date the data is extracted.
The date of data extraction, i.e. the date that the data set was assembled, should be reported in the Hospital Level Data file described in Appendix C – Reporting File Structure.

## 2.4 Step 4. Extract data for and calculate phase 1 variables

Table 1 identifies the variables needed for phase 1 data extraction, including matching, outcome and other variables. Detailed definitions for each variable and its supporting data are provided in Appendix A – Data definitions.

Severity of respiratory illness measured on a daily basis is needed for the first 10 days of the hospital stay. Table 2 shows the specific definitions of each of 4 severity levels used in the study. Reporting hospitals should calculate and report just one score for each day. If a patient qualifies for more than one

severity level, choose the most severe level experienced during a given calendar day. Daily severity is needed on all patients (those with and without CP infusion) since the intent is to match controls by first day of treatment in the treated population, and that matching cannot be known ahead of time for non-CP patients.

| TABLE 1 – PHASE 1 VARIABLES | |
|---|---|
| **Data Element Category** | **Data Element** |
| Matching | Hospital name |
| | Age |
| | Administrative gender |
| | Severity of respiratory illness on day of admission (day 0) |
| | Severity of respiratory illness (day 1) |
| | Severity of respiratory illness (day 2) |
| | Severity of respiratory illness (day 3) |
| | Severity of respiratory illness (day 4) |
| | Severity of respiratory illness (day 5) |
| | Severity of respiratory illness (day 6) |
| | Severity of respiratory illness (day 7) |
| | Severity of respiratory illness (day 8) |
| | Severity of respiratory illness (day 9) |
| | Admission epoch |
| Outcome/time variable | Admission date |
| | Date of convalescent plasma administration |
| | Mechanical ventilation start date |
| | Mechanical ventilation end date |
| | Discharge date |
| | Date of death |
| Other variables | COVID-19 positive laboratory test |
| | Patient study ID |

Severity of illness refers to the severity of the patient's respiratory symptoms. The goal is to measure worst severity on the day of treatment with CP infusion. It is recommended to calculate and store one daily value (midnight to midnight) of severity for the first 10 days of hospitalization. "Case matching" will require use of this value, therefore we need this value to be also calculated on "control" patients. The logic for calculating worst severity for any given day is provided in Appendix A – Data definitions.

| TABLE 2: SEVERITY OF ILLNESS CATEGORIES | | |
|---|---|---|
| **Severity of Illness Allowable Value** | **Description** | **Definition** |
| 5 | Not on supplemental oxygen | Not on supplemental oxygen, on high-flow supplemental oxygen or mechanically ventilated, per definitions below. |
| 4 | On conventional supplemental oxygen therapy | On nasal cannula or oxygen facial mask < 30L/min |
| 3 | On high-flow supplemental oxygen | - On high-flow nasal cannula (HFNC) or oxygen facial mask >= 30L/min<br>- Non-invasive positive pressure ventilation (NIPPV), including BiPAP, or CPAP between 8am and 9pm (8am to 9pm requirement on CPAP to rule-out regular home CPAP use) |
| 2 | Invasive mechanical ventilation | Mechanical ventilation (as evidenced by PEEP, vent mode change, FiO2 flowsheet documentation) or ECMO |

## 2.5 Step 5. De-identify data including disguising dates

Review data to insure there are no patient identifiers included in the data set. There should be no patient names, DOB, MRNs, insurance numbers, phone numbers, home zip codes, home addresses, email addresses, SSN's or other personal identifiers in the data set.

Additionally, we ask hospitals to report disguised dates using an integer. The patient's admission date should be assigned the number zero (0). All subsequent dates should be computed as days from admission date. For example, if the patient was admitted on March 25$^{th}$, received CP on April 2$^{nd}$ and was extubated on April 5$^{th}$, and discharged on April 10$^{th}$, those dates would be recorded as 4 (0+4), 8, 11, and 16.

Note: if your organization has already implemented the date disguising method described in the study protocol and previous versions of this guide, you don't need to implement this new method. However, you will need to report the random number used to disguise the admission date so that the researchers are able to calculate the time intervals (measured in days) needed in the study.

## 2.6 Step 6. Assign a study ID to each patient

To facilitate future data updates, including corrections and the transmission of phase 2 variables, hospitals should assign each patient should be assign a randomly generated unique identifier (cannot contain any personally identifiable information). Hospitals should keep a record of the crosswalk between the patient's medical record number and their study identifiers for future reference.

## 2.7 Step 7. Create CSV/TSV data set

Convert your data file to a CSV or TSV format ready for transmission. See Appendix C – Reporting for more information of file structure.

## 2.8 Step 8. Transmit data to Mayo Clinic

Mayo will be the data aggregation center for the study. Participating hospitals can upload their data using secure SharePoint. Mayo will share a link to the appropriate SharePoint website.

## 2.9 Step 9. Phase 2 Select co-morbidities, medications, test results data

See Appendix B - Phase 2 variables for the data elements which will be collected in Phase 2. Epic EHR and Cerner EHR will be able to provide guidance on specific tables and data definitions based on ICD-10, SNOMED, CPT, LOINC and other parameters used in their systems. Additional details on these variables and how to report them may also be added to this document as they become available. Additional variables may also be added based on the growing body of scientific knowledge and work done by organizations pursuing real world evidence around COVID-19 and convalescent plasma.

## Appendix A – Data definitions

These definitions were arrived upon through the work of the COVID-19 Healthcare Coalition partners, including a multidisciplinary group of clinical, informatics and EHR data experts. Vendor-specific implementations of these definitions may be available from your EHR vendor. If you have any questions or feedback on these definitions, please reach out to us using the contact information above.

| Data Element | Description | Logic |
|---|---|---|
| Administrative gender | The gender of the patient used for administrative purposes, typically the patient's sex as specified in their legal documents. | n/a |
| Admission date | The date the patient was admitted for inpatient care. | n/a |
| Admission epoch | The month and year the patient was admitted for inpatient care for the COVID-19 related hospitalization. Admission epoch is intended to provide an indication of when the patient was hospitalized over the course of the pandemic. | n/a |
| Age at COVID-19 positive date | The age (in years) of the patient on the date they were diagnosed with COVID-19. | IF COVID-19 positive date - date of birth> 90, “>90” ELSE COVID-19 positive date – date of birth |
| Convalescent plasma administration | Convalescent plasma administration regardless of the number of units or antibody titer. | Blood product order LIKE %COVID-19%<br>**OR**<br>Blood product order administration product code in ISBT 128 E codes for COVID-19 convalescent plasma) |
| Convalescent plasma administration date | The first date when the patient was administered convalescent plasma during the hospitalization. | Earliest Convalescent plasma administration |

| Data Element | Description | Logic |
|---|---|---|
| COVID-19 confirmed diagnosis | A clinical diagnosis (any encounter diagnosis, billing diagnosis or problem list entry) of Confirmed COVID-19 infection. See Terminology section below for inclusion and exclusion criteria for COVID-19 confirmed infection below. | Condition in Confirmed COVID-19 Infection value set AND Condition.type ~ (encounter diagnosis, discharge diagnosis, final diagnosis, primary diagnosis, billing diagnosis, problem list entry) AND Condition.type NOT ~ admitting diagnosis |
| COVID-19 positive | A patient who has been clinically diagnosed with COVID-19 or who tested positive for COVID-19. | COVID-19 confirmed diagnosis **OR** COVID-19 confirmatory laboratory test |
| COVID-19-positive date | The earliest date associated with the confirmation of the COVID-19 infection. | Earliest of (COVID-19 confirmatory laboratory test specimen collection date[1], first COVID-19 confirmed diagnosis) |
| COVID-19-positive laboratory test | A laboratory test indicating that the patient has a COVID-19 infection. See Terminology section below for inclusion and exclusion criteria for COVID-19 laboratory tests. | Laboratory test in COVID-19 Qualitative Laboratory Test value set AND (laboratory test result ~ detected **OR** ~positive) |

[1] The result date can be used when specimen collection date is not available

| Data Element | Description | Logic |
|---|---|---|
| COVID-19-related hospitalization[2] | An encounter for inpatient care that is associated with COVID-19. If a patient is diagnosed with COVID-19 prior to the hospitalization, a respiratory diagnosis is used as a proxy for symptomatic COVID-19 infection when patient is admitted for reasons unrelated to COVID-19. It is recognized that this definition may include COVID-related hospitalizations but also hospitalizations of COVID-19 patients who may be asymptomatic, and nosocomial COVID-19 infections. | Encounter class ~ INP **OR** ~ inpatient **OR** ~acute inpatient<br>**AND**<br>(<br>COVID-19-positive date <= 7 days after hospital admission<br>**OR**<br>COVID-19-positive date <= 14 days prior to hospitalization<br>**AND** respiratory diagnosis associated with hospitalization<br>) |
| Date of death | The date the patient expired. It is only expected that deaths during hospitalization are recorded. | IF discharge status LIKE %deceased% OR %expired% (or similar) then discharge date<br>**OR**<br>IF vital status LIKE %deceased% or %expired% (or similar) then date of death |
| Discharge date | The date the patient was discharged from inpatient care. The discharge date is expected to be the same as the death date for deceased patients. | n/a |
| Mechanical ventilation start date | The first date the patient was mechanically ventilated. See definition of “On invasive mechanical ventilation” for more details. | Earliest (vent mode OR vent activity) indicative of invasive mechanical ventilation. |

| Data Element | Description | Logic |
|---|---|---|
| Mechanical ventilation end date | The last date the patient was mechanically ventilated following the first intubation. See definition of "On invasive mechanical ventilation" for more details. | Most recent (vent more OR vent activity) NOT LIKE ~ CPAP% OR ~ %BiPAP% OR ~ %High Flow Nasal Canula% for 1st intubation |
| Respiratory diagnosis associated with hospitalization | Any diagnosis for a respiratory condition associated with a hospitalization | Condition code ICD-10-CM J00-J99<br>**AND**<br>Condition.type ~ (encounter diagnosis, discharge diagnosis, final diagnosis, chief complaint, primary diagnosis, billing diagnosis, problem list entry) |
| Not on supplemental oxygen | Indication that the patient did not receive any form of supplemental oxygen on a given hospital day (midnight to midnight). | **NOT** on invasive mechanical ventilation<br>**AND NOT** on high flow invasive mechanical ventilation<br>**AND NOT** on invasive mechanical ventilation |
| On conventional supplemental oxygen therapy | Evidence that the patient received conventional oxygen therapy on a given hospital day (midnight to midnight) | On nasal cannula or oxygen facial mask < 30L/min |
| On high-flow supplemental oxygen | Evidence that the patient received high-flow supplemental oxygen on a given hospital day (midnight to midnight). CPAP use between 8am and 9pm is excluded to rule out regular home CPAP use. | (On high-flow nasal cannula (HFNC) or oxygen facial mask >= 30L/min **OR**<br>Non-invasive positive pressure ventilation (NIPPV), including BiPAP **OR**<br>CPAP (between 8am and 9pm)<br>**AND** |
| On invasive mechanical ventilation | Evidence that the patient was mechanically ventilated on a given hospital day (midnight to midnight). See Terminology section below for inclusion and exclusion criteria for invasive mechanical ventilation. | n/a |

| Data Element | Description | Logic |
|---|---|---|
| Severity of respiratory illness | The worst severity of the patient's respiratory symptoms on a given hospitalization day (with the day of admission as hospital day 0), as evidenced by the patient's oxygenation requirements. A day is considered an individual date, i.e. midnight to midnight. | IF On invasive mechanical ventilation<br>THEN 2<br>ELSE IF On high-flow supplemental oxygen THEN<br>ELSE IF On conventional supplemental oxygen therapy THEN 4<br>ELSE IF Not on conventional supplemental oxygen THEN 5 |

## Terminology

**COVID-19 Confirmed Diagnosis**

| | |
|---|---|
| Includes | Conditions associated with confirmed COVID-19 infection, including laboratory-confirmed COVID-19 (symptomatic or asymptomatic).<br><br>ICD-10-CM:<br>U07.1 only available since 4/1/2020, is used for lab-confirmed cases regardless of symptom presentation<br><br>SNOMED-CT:<br>840539006 Disease caused by severe acute respiratory syndrome coronavirus 2 (disorder) |
| Excludes | ICD-10-CM and SNOMED-CT codes indicative of suspicion or exposure only<br>B97.29 (used largely before 4/1/2020) |

**SARS-CoV-2 Laboratory Tests**

| | |
|---|---|
| Includes | SARS-CoV-2-specific or SARS-like<br>PCR or NAAT<br>SARS-CoV-2 RNA in serum/plasma<br>SARS-CoV-2 panels (not recommended for results by Regenstrief but used in the field)<br>Qualitative results |
| Excludes | Human coronavirus tests (non-SARS/SARS-like tests) and MERS tests<br>Antibody and antigen tests<br>Quantitative results (e.g. cycle threshold #) |

**Invasive Mechanical Ventilation**

| | |
|---|---|
| Includes | • ICU flowsheet documentation of vent mode or vent activity, evidence of positive end-expiratory pressure (PEEP) documentation.<br>• Mechanical ventilation-associated procedure performed (CPT 94002 or 94003) – if flowsheet documentation isn't reliable<br>• Extracorporeal membrane oxygenation (ECMO) |
| Excludes | • Intubation procedure<br>• High-flow oxygen delivered through nasal cannula, BiPAP, or CPAP<br>• Sedation meds without evidence of ventilator management<br>• Conditions indicative of respiratory failure without evidence of ventilator management |

# Appendix B - Phase 2 variables

The following table provide a list of data elements currently identified as potential covariates to support logistic regression and Cox models. These elements will be considered in a second phase of the study. Specific definitions and reporting templates for these elements are actively under development and will be added to this document as they become available. Additional variables may also be added based on the growing body of scientific knowledge and work done by organizations pursuing real world evidence around COVID-19 and convalescent plasma.

| **Data Element Category** | **Data Element** |
|---|---|
| Patient characteristics | Race |
| | Ethnicity |
| | Body Mass Index |
| | Smoking status |
| Comorbidities | Pre-hospital diabetes |
| | Pre-hospital chronic lung disease |
| | Pre-hospital asthma |
| | Pre-hospital hypertension |
| | Pre-hospital cardiovascular disease |
| | Pre-hospital cerebrovascular disease |
| | Pre-hospital End Stage Renal Disease (ESRD) |
| | Pre-hospital chronic kidney disease (CKD) other than ESRD |
| | Pre-hospital malignancy |
| Hospitalization | ICU admission |
| | In-hospital renal impairment |
| | In-hospital sepsis |
| | In-hospital multi-organ failure |
| | Discharge disposition (to identify transfers to another acute care facility) |
| Laboratory test results[3] | D-dimer |
| | Full WBC |
| | pO2/FIO2 and SpO2/FIO2 ratios |
| | SPO2 |
| | CRP |
| | LDH |
| | Cardiac troponin |
| Other therapies administered during hospitalization | Antivirals (Remdesivir) |
| | Interleukin-6 agents (tocilizumab, sarilumab) |
| | H2-blockers (famotidine) |
| | Corticosteroids (prednisone, methylprednisolone, dexamethasone) |

[3] For laboratory values, we will have to determine whether we want all values downloaded or select highs or lows or other fractions of the distributions, since there are likely to be multiple tests, especially with oxygenation.

# Appendix C – Reporting File Structure

## Hospital Level Data

This file should include information about the individual facilities where convalescent plasma has been administered. Data should not be aggregated at a health system level, because matching will be done at the facility level.

| VARIABLE CODE | VARIABLE NAME | DATA TYPE | FORMAT/ALLOWABLE VALUES | OCCURS | NOTES FOR EXTRACTION |
|---|---|---|---|---|---|
| hospital | Hospital name | string | Any valid string | 1-20 | Each individual facility should be listed, as matching will occur at the facility level. |
| address | Hospital address | String | Any valid string | 1-20 | The address for each facility. |
| date_first_cp | First plasma patient | Date | YYYY-MM | 1-20 | The month associated the first CP patient treated at an individual facility was admitted. |
| date_extraction | Data extraction date | Date | YYYY-MM-DD | 1 | The date the data was extracted from the EHR. |

## Patient-level Data Phase 1

This file should include variables for Phase 1, for all patients identified in Step 1: Identify all COVID-19 hospitalized patients. This includes both patients who were treated with convalescent plasma as well as those who were not. Please consult the data definitions in Appendix A – Data definitions for detailed definitions and inputs for each reported variable.

| Variable Code | Variable Name | Data Type | Allowable Values | Occurs | Notes for Extraction |
|---|---|---|---|---|---|
| hospital | Hospital name | string | Any valid string | 1 | The name of the hospital where the COVID-19 hospitalization occurred. This should match the hospital name included in the hospital level data file. |
| id | Patient study ID | Integer | Any valid integer | 1 | Instructions on how to generate a patient study ID are available in section 2.6 Step 6. Assign a study ID to each patient. |
| age | Age | Integer OR string | 18-89, “>90” | 1 | For purposes of deidentification, if patient age is >90, populate ">90" instead of the numeric age |
| admin_gender | Administrative gender | string | M = male<br>F = female<br>U = unknown/other | 1 | |
| severity_day[x] | Severity of respiratory illness | Integer | 5<br>4<br>3<br>2 | 10 | Consult Appendix A – Data definitions for definitions of each allowable value. |
| admission_epoch | Admission epoch | Date | YYYY-MM | 1 | |
| date_admission | Admission date | Integer | any valid integer | 1 | Only needs to be reported if using the date disguising method described in the study protocol or earlier versions of this guide. |
| covid_positive_test | COVID-19-positive laboratory test | Boolean | TRUE = positive test result<br>FALSE = negative test result<br>NULL = patient was not | 0 -1 | This variable is not mandatory. |

| | | | | | |
|---|---|---|---|---|---|
| | | | tested, result not available (e.g. pending) or unknown (e.g. inconclusive, invalid, equivocal, indeterminate) | | |
| date_cp_admin | Date of convalescent plasma administration | boolean or integer | NULL = not administered convalescent plasma<br><br>any valid integer = disguised date | 1 | Follow the instructions on section 2.5 Step 5. De-identify data including disguising dates. |
| start_date_vent | Mechanical ventilation start date | boolean or integer | NULL = not mechanically ventilated at any point from admission until discharge or data extraction date (if still hospitalized)<br><br>any valid integer = disguised date | 1 | Follow the instructions on section 2.5 Step 5. De-identify data including disguising dates. |
| end_date_vent | Mechanical ventilation end date | boolean or integer | NULL = never mechanically ventilated or still mechanically ventilated<br><br>any valid integer = disguised date | 1 | If start_date_vent = integer AND end_date_vent = FALSE, then the patient will be considered to be mechanically ventilated at the time of data extraction.<br>If start_date_vent = FALSE and end_date_vent = FALSE, then the patient will be considered to never have been mechanically ventilated from admission to the date of data extraction. |
| date_death | Date of death | boolean or integer | NULL = not deceased as of date of data extraction<br><br>any valid integer = disguised date | 1 | Follow the instructions on section 2.5 Step 5. De-identify data including disguising dates |

| | | | | | |
|---|---|---|---|---|---|
| date_discharge | Date of discharge | boolean or integer | NULL = patient is currently hospitalized<br><br>any valid integer = discguised date | 1 | Note this variable subsumes the "hospitalization at the time of data extraction" variable as allowable value FALSE. |