Frequently Asked Questions – COVID-19 Datamarts

1. *Who is listed in the datamarts?*

All patients who have been tested for COVID-19 at UCMC, including curbside testing and employees. The datamarts includes patients who test both positive and negative.

2. *Where can I see what data is included in each datamart?*

We have a table saved in [Box](#) and [online](#) where you can see what data fields are listed in each type of datamart.

3. *What is the difference between the de-identified, limited, and identified datamarts?*

The primary difference between the de-identified and limited datamarts is that the limited datamart has real dates and zip codes available (zip codes are only available to projects that specifically request them). The identified datamart is available on a case by-case basis. Please contact Julie Johnson at cri-covid@lists.uchicago.edu with more questions about the identified dataset. The IRB does require consenting for prospective, identified data collections.

4. *I do not see the data I need listed in the datamarts. Can you send me the data I am missing?*

We are adding new data fields to the datamarts as researchers request them. New data will be added based on prioritized need.

5. *Do I need to have IRB review in order to access the de-identified datamart?*

Yes, you will need a non-human subjects determination from the IRB in order to receive access to the de-identified datamart. You must include Break the Glass employees in your IRB submission. Specify that you would like access to notes IF you will need notes for your project. Please send us the IRB number once the determination is made so that we can verify.

6. *Is there a data dictionary for the COVID-19 datamarts?*

The CRI has been developing a data dictionary for the datamarts, saved in Box. The folder will be shared with study staff at the same time they receive access to the datamart.

7. *Can I access real dates for my project?*

Yes, real dates are available in the limited dataset. You will need to specifically request real dates in your IRB submission. You will also need to specifically request access to notes, and break-the-glass employees in order to access the limited dataset.

8. *Why are there encounters which occurred in 2019 in the dataset?*

If you are using the de-identified datamart, because the dates have all shifted by a different integer for each patient, the dates listed are not the actual dates when the encounters took place. Also, in all datasets, 1 year of medical history is included for each patient, which may have taken place in 2019.

9. *Are images available through the datamarts?*

The CRI and HIRO partnered to offer chest radiographs (aka chest x-rays, including standard and portable exams) and chest CT scans in the limited datamart.

10. *Why does the* C19_Patient_ID *assigned to the chest radiograph not correspond with a patient in the data sheet* C19_PATIENT_DEMO*?*

Occasionally a patient will receive care from UCMC using the incorrect MRN. This happens for a variety of reasons, particularly in the ED, a patient may get admitted and generate data (i.e. radiograph) with one MRN, before it is discovered that patient already exists in the system as another MRN. Patient IDs in the COVID-19 datamarts are associated with a corresponding MRN, and so the incorrect MRN/Patient ID is sometimes assigned to the chest radiograph. Other data for this patient would be associated with the correct MRN/Patient ID. Search for the HAR associated with the image in the other data sheets to find the other encounters and correct Patient ID.

11. *Is there a fee associated with accessing the COVID-19 datamarts?*

As of September 08, 2020, there are no fees charged by the CRI in order to access the limited or de-identified COVID-19 datamarts.