

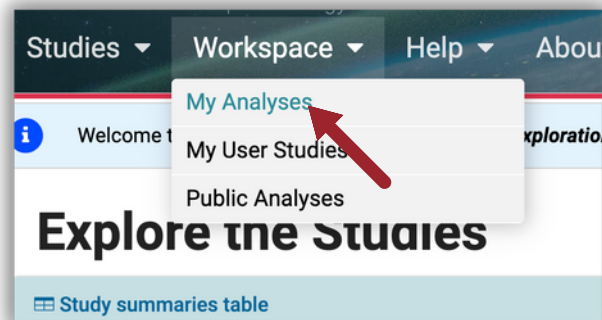


## Exercise: Downloading data from ClinEpiDB



**ClinEpiDB** allows you to download data, subject to data access requirements put in place by the study team. In this exercise, you will open the **PRISM ICEMR Cohort** analysis from the previous exercise (exploratory data analysis on malaria in Uganda), and download files from **your subset** and the **full dataset**.

On the ClinEpiDB home page, click on **My Analyses** under the **Workspace** menu in the header. This opens a table with your analyses. Open the **PRISM ICEMR Cohort analysis** that you made before.



1. Go to the **Download** tab to explore options for downloading data.
2. The data accessibility level for the **PRISM ICEMR Cohort** is (check one)  
Controlled      Protected      Public
3. Downloading files from **My Subset**:
  - a. This option allows you to customize the rows and columns that you wish to download
    - i. Rows are chosen by filtering data in the **Browse and Subset** tab
    - ii. Columns are the variables that you specify in the **Download** tab
  - b. In your analysis, you have already added two filters to create a subset:
    - i. **Observation type** = Enrollment/Scheduled visit
    - ii. **Age** <= 10
  - c. Let's imagine we are looking for data on **Signs and Symptoms**. Based on the variable tree in the **Browse and Subset** tab, which file do you think will let you download these variables? \_\_\_\_\_
  - d. Now go to the **Download** tab.
  - e. How many **Participant repeated measures** are in our chosen subset? \_\_\_\_\_

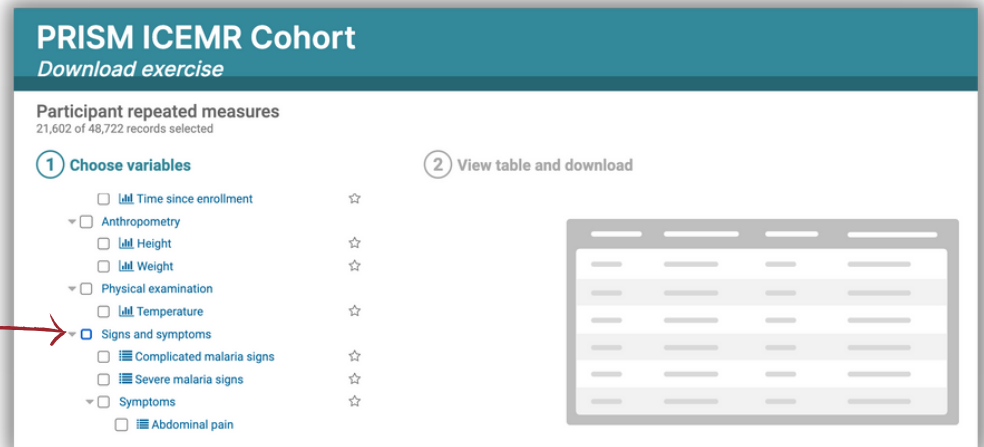
f. Click on the **Participant repeated measures** in **My Subset**. This opens a pop-up window which lets you choose which variables to include in your subset download file.

 21,602 of 48,722 Participant repeated measures

g. Scroll down the variable tree on the left, find the **Signs and Symptoms** variables and add them to the table.

h. Examine the table and look at the **Participant**

**repeated measure ID o\_239520054** (second row). Scroll through the row. Did this participant have **Abdominal pain**? \_\_\_\_\_



#### 4. Downloading files from the **Full Dataset**:

a. On the **Download** tab, scroll down to the **Full Dataset** section and take a look at the files listed.

b. Download and open the **README** file. This file contains helpful notes and tips for working with the data files, such as key identifiers, which are useful for merging files.

 [rls0021\\_PRISM\\_cohort\\_Download\\_Files\\_README](#)

c. How many key identifiers are listed in this README file? \_\_\_\_\_

d. The **OntologyMetadata** file contains important metadata about the study (i.e. information about the data in the study). Download the file and open it using a program such as Excel.

e. Look for the variable label **Febrile**. What definition is listed for this variable?

\_\_\_\_\_

Turn to the next page for answers to this exercise!



1. The data accessibility level for the PRISM ICEMR Cohort is **Public**.
- 3c. Let's imagine we are looking for data on Signs and Symptoms. Which subset file will let you download these variables? **Participant repeated measures**
- 3e. How many Participant repeated measures are in our chosen subset? **21,602**
- 3h. Examine the table and look at the Participant repeated measure ID o\_239520054 (second row). Scroll through the row. Did this participant have Abdominal pain? **No**
- 4c. How many key identifiers are listed in this README file? **5** (*Household\_Id, Household\_repeated\_measure\_Id, Participant\_Id, Participant\_repeated\_measure\_Id, Sample\_Id*)
- 4e. Look for the variable label Febrile. What definition is listed for this variable? **Elevated temperature or subjective fever**

Thank you for completing this exercise on downloading data!