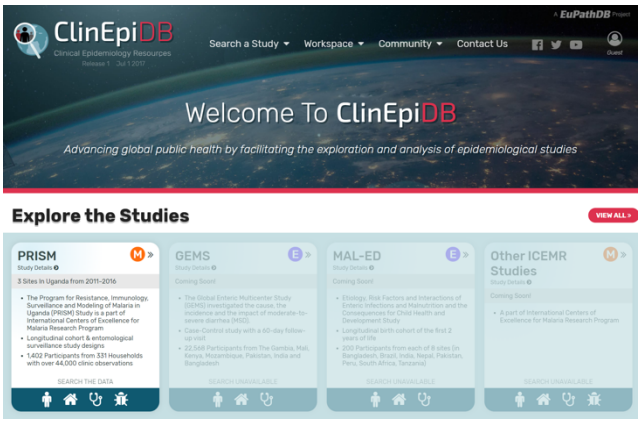
Tutorial: Accessing ClinEpiDB data and exploring via the filters

This tutorial will explain how to access data in ClinEpiDB, including:

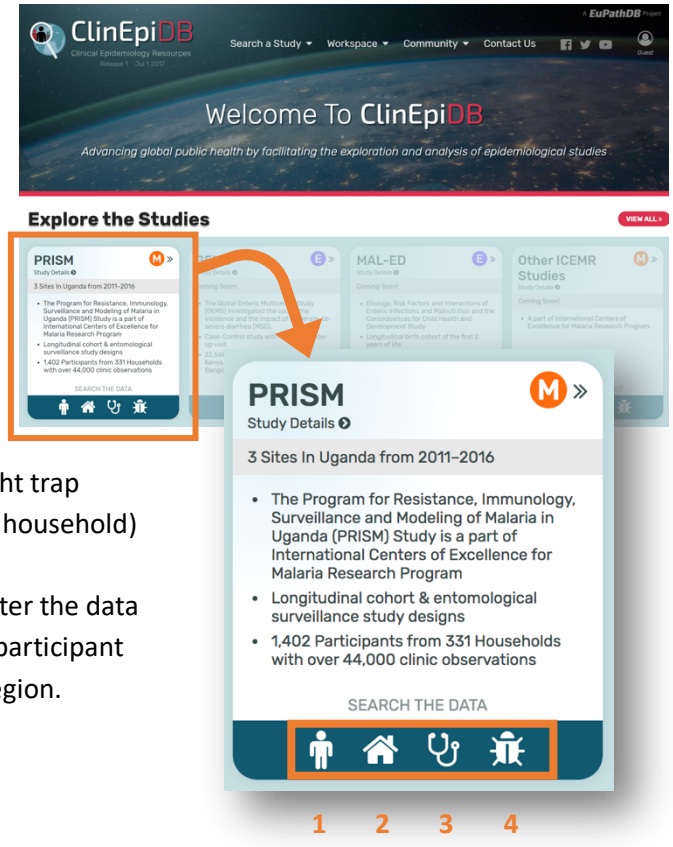
1) What types of data you can query
2) How to use search filters and corresponding histograms and bar charts



Types of Searches: Participants, Households, Observations, and Light traps

From the ClinEpiDB homepage, look at the small icons that are located along the bottom of the PRISM study card. The PRISM study has four types of searches you can execute on the data:

1) Participants – returns one row of data per participant on personal characteristics
2) Households – returns one row of data per household
3) Observations – returns one row of data per observation (can have multiple observations per participant)
4) Light traps – returns one row of data per light trap collection (can have multiple light traps per household)

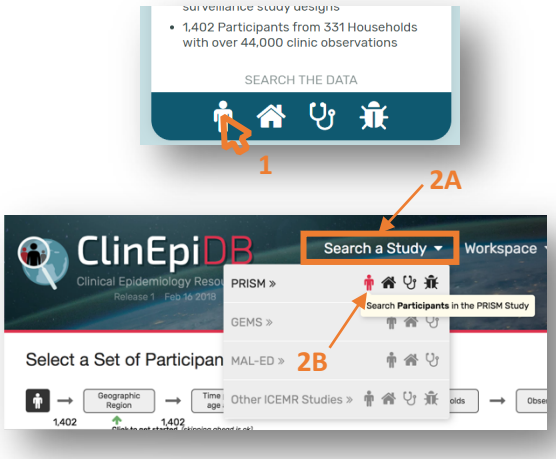


Within each type of search, you can examine and filter the data based on characteristics or observations about the participant and/or data about the household and geographic region.

There are two ways to select a search:

1) Click on the icon of interest on the PRISM study card

2) Through the navigation bar at the top of the page, hover over (A) “Search a Study” and then (B) select the icon that corresponds to the study and search type of interest

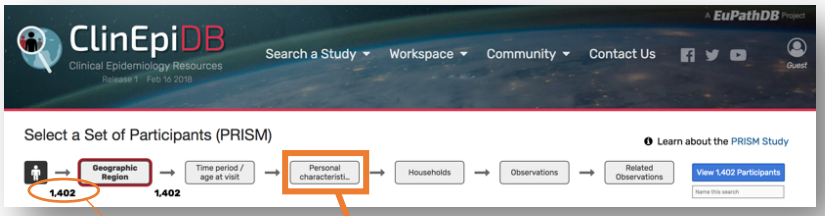


We’ll do a Participant search for this tutorial.

Explore Participant Search Filters

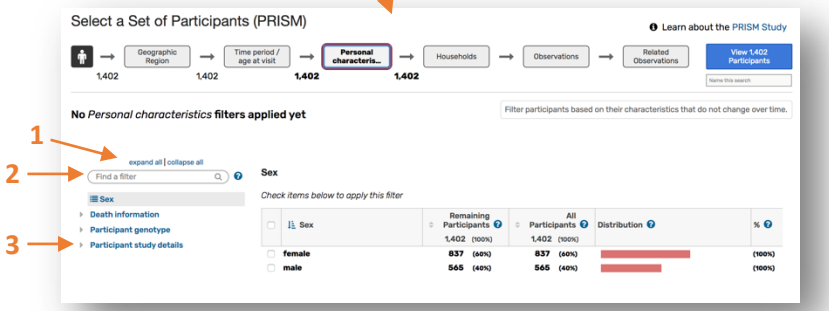
After clicking on the type of study, the search wizard appears. The search wizard categorizes components of the data, allowing for a step-wise approach to building searches and updating your selections along the way. By clicking through each major category, you can explore the dataset, filter it, and see the distribution of variables of interest.

For this tutorial, click on the “Personal characteristics” category.



In the left panel you can see a list of variable subcategories and variables in blue. You can:

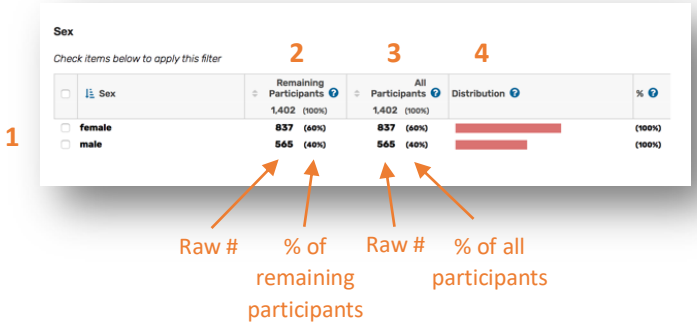
1) Click “expand all” to see all variables in the current category
2) Type the name of the variable you’re interested in directly into the search box to pull it up
3) Click a specific subcategory like “Participant study details” to reveal the variables within



When you click on a variable, you will see either a table or histogram on the right-hand side depending on the type of variable. For binary or categorical data you will see a table. For numerical data, you will see a table if there are ≤10 different values and a histogram if there are >10 different values.

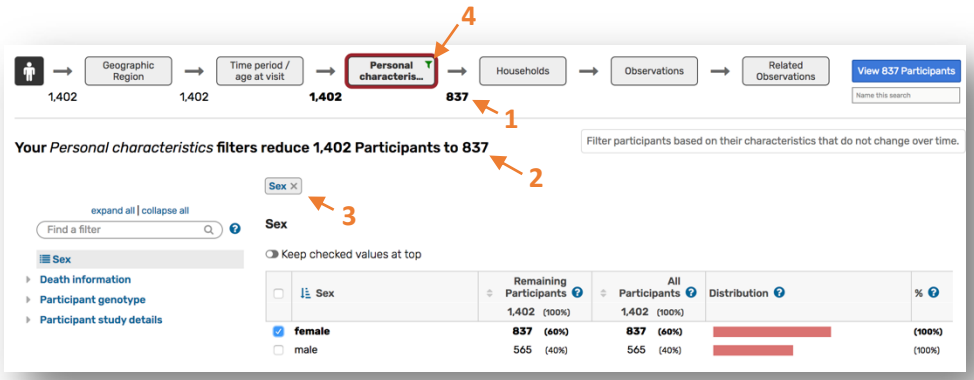
Let's explore the Sex variable, which is a categorical variable. Clicking the variable name in the left panel reveals a table that includes:

1) Possible values for the variable
2) Remaining participants – number of participants with that value after filters have been applied
3) All participants – the total number of participants with that value
4) Distribution – length of red bar indicates the number of participants with that value

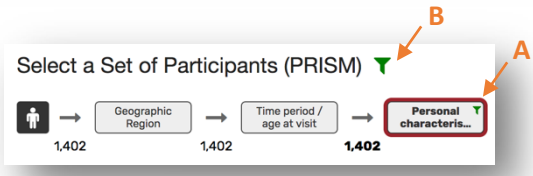


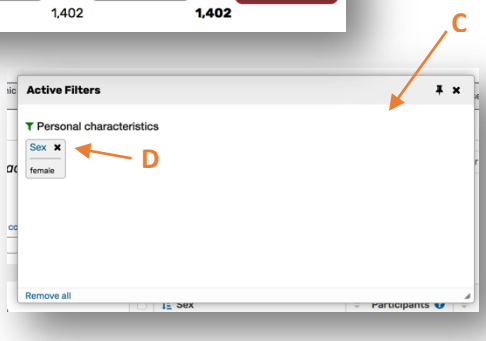
To apply a filter and select participants with specific characteristics from a table, tick the checkbox next to the characteristics you are interested in, in this case female participants. Note that if you do not tick any checkboxes, no filter is applied and all of the data is available in the next step of the wizard. There are multiple ways to tell if you have applied a filter:

1) The number of participants in the search wizard
2) The statement under the search wizard
3) The presence of a box for each filtered variable (To delete a filter, click the grey "x")
4) The appearance of the green filter icon



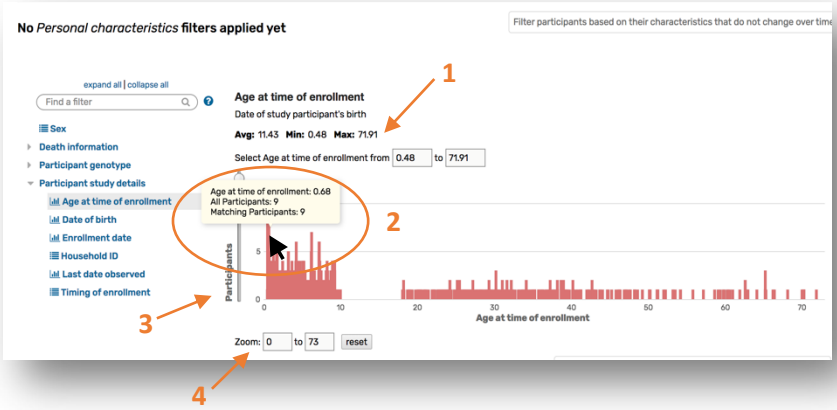
Clicking on the green filter icon either in the search wizard (A) or at the top of the page (B) generates a pop-up window listing all the filters that have been applied (C). Delete the filter for Sex by clicking the 'x' (D).





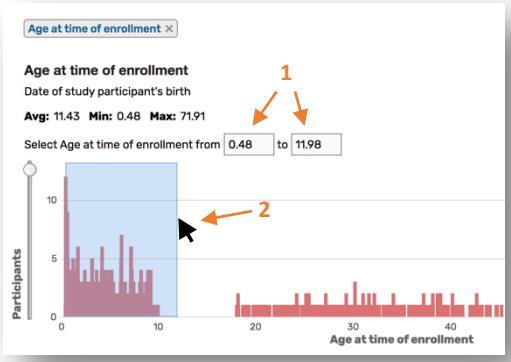
Now let's explore a numerical variable. Select the variable 'Age at time of enrollment' from the 'Participant study details' subcategory in the left panel to reveal its associated values. In the histogram you can:

1) See summary statistics
2) Mouse over individual bars to learn more about the number of participants with that x-value
3) Use the y-axis scroll bar to zoom in on data with lower numbers of participants
4) Use zoom to look at a specific part of the x-axis



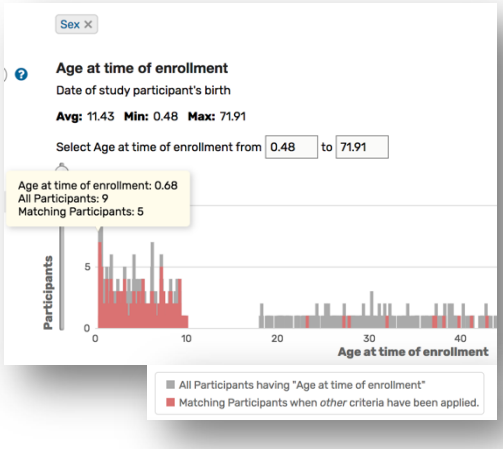
To apply a filter and select participants that fall within a range of values you can:

1) Type the exact range you are interested in and hit enter
2) Select the range you are interested in with your mouse by clicking and dragging from one end of the range to the other



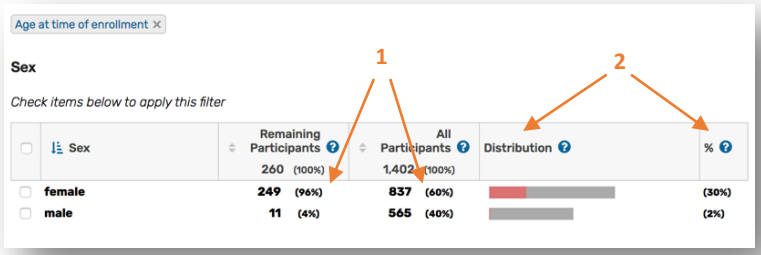
Once you have filtered the data, you can look at how the filters you applied affect the distribution of other characteristics of interest. For instance, if we select 'male' for Sex and then look at Age at time of enrollment, we can see that most adults who were enrolled were not male by:

1) Comparing the height of the red bars, which indicate the number of participants that match the filter (in this case "male") to the grey bars, which indicate the number of all participants
2) Moving the mouse over each bar to see the exact number of participants that match the selection criteria (Matching Participants) vs. the total number of participants (All Participants)



Conversely, if we clear the filter, select an age range from 20-50 years for Age at time of enrollment, and look at Sex we can see:

1) 96% of participants remaining after we applied the filter are female, but only 60% of all participants in the study are female.
2) The last two columns (Distribution and %) show what percent of participants with the characteristic you are currently looking at (Sex) also have the filtered characteristics (age 20-50)
    a. In the bar graph, total bar length indicates the number of all participants, red indicates number of participants that match the filter criteria, while grey indicates the number of participants that did not
    b. 30% of all female participants were 20-50 years old, while only 2% of all male participants were 20-50 years old



Thank you for using this tutorial. Please contact us with any questions or suggestions.