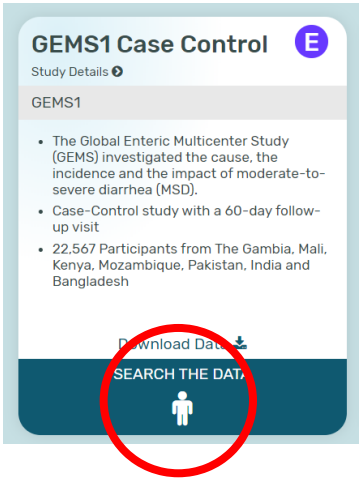
In GEMS1, how many Giardia positive cases had at least one matched control that was Giardia negative?

What is the relationship between Giardia and MSD? Does this differ for Cryptosporidium? What about by age? (30 min)

In this exercise, you will learn to do a more advanced search using our related case/control function. One of the unique assets of the GEMS1 and GEMS1A data is the matched set design and the ability to search these data easily is valuable feature of the ClinEpiDB resource.

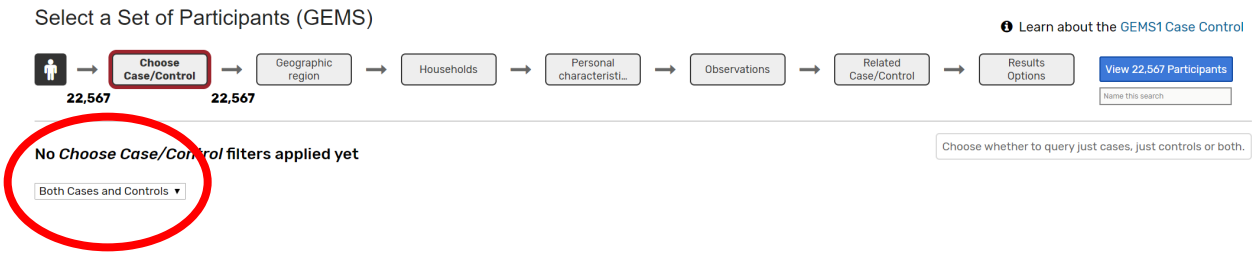
We will start with a very simple related case/control query. For this exercise, to start, we will modify as few filters as possible to keep things simple.

To explore how many Giardia positive cases had at least one matched control that was Giardia negative, your first step might be to explore the data and see how many Cases tested positive for *Giardia*. Then you would ask the *Giardia* status of those Cases matched controls. *You can do this by using the related Case/Control functionality!*

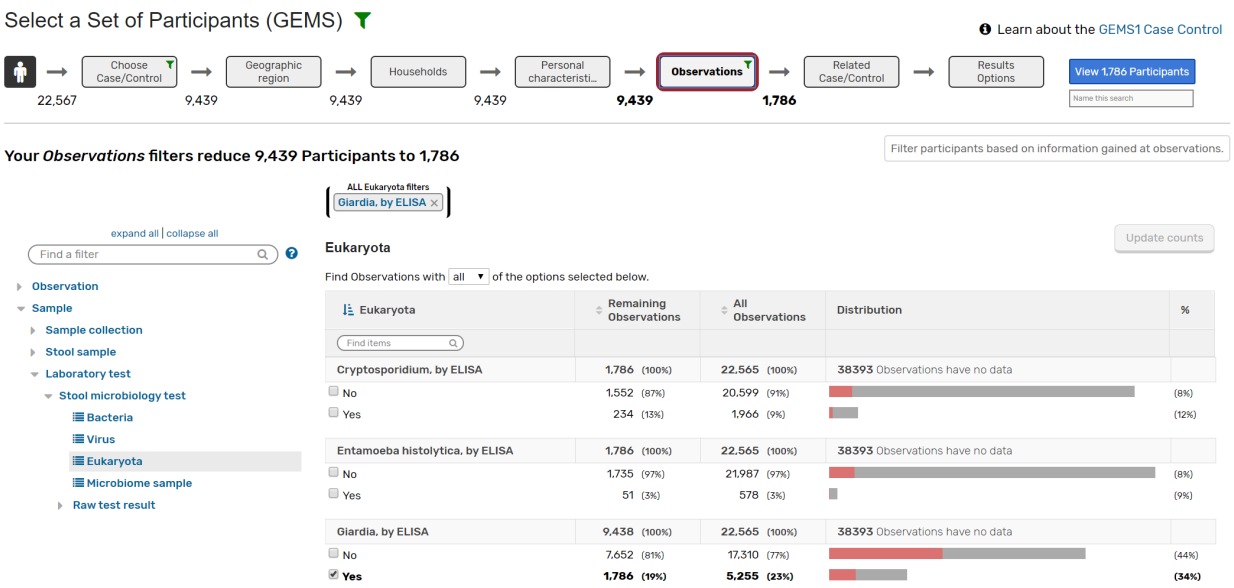


1. Start this exercise at the GEMS1 Participant level in the search wizard. We describe this functionality as a 'Related Case/Control' search because you are narrowing your dataset based on the relationship between your Case and Control participants. First select Cases who tested positive for *Giardia*.

2. Navigate to the 'Choose Case/Control' box in the Search Wizard. Select 'Cases'

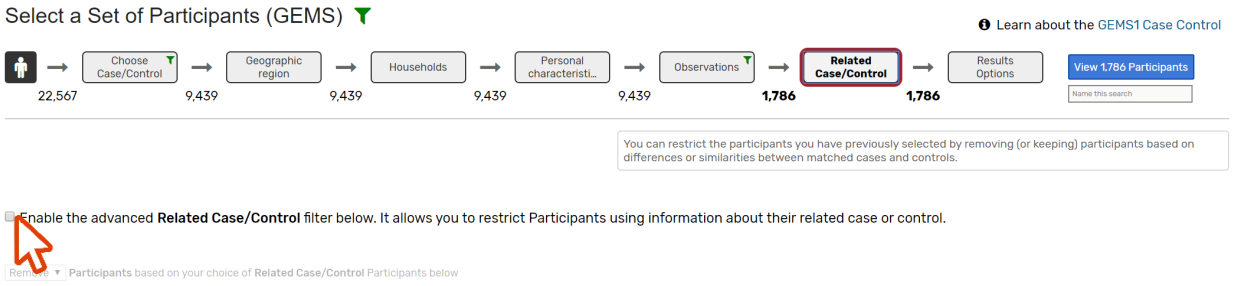


3. Next, select from Observations that *Giardia* positive. This can be found in the 'Stool microbiology tests' filter in observations. You can navigate to this through the filter hierarchy on the left or by typing "Stool microbiology" into the search box.



You should notice that the number of participants reduced to 1,786. This means that 1,786 Cases had a positive test for *Giardia*.

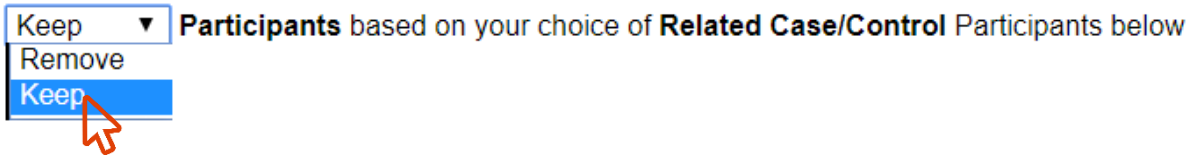
4. Now, ask that of these *Giardia* positive Cases, how many had at least one matched Control that was Giardia negative. To do this, you would use the 'Related Case/Control box in the query wizard.
5. Click on 'Related Case/Control'



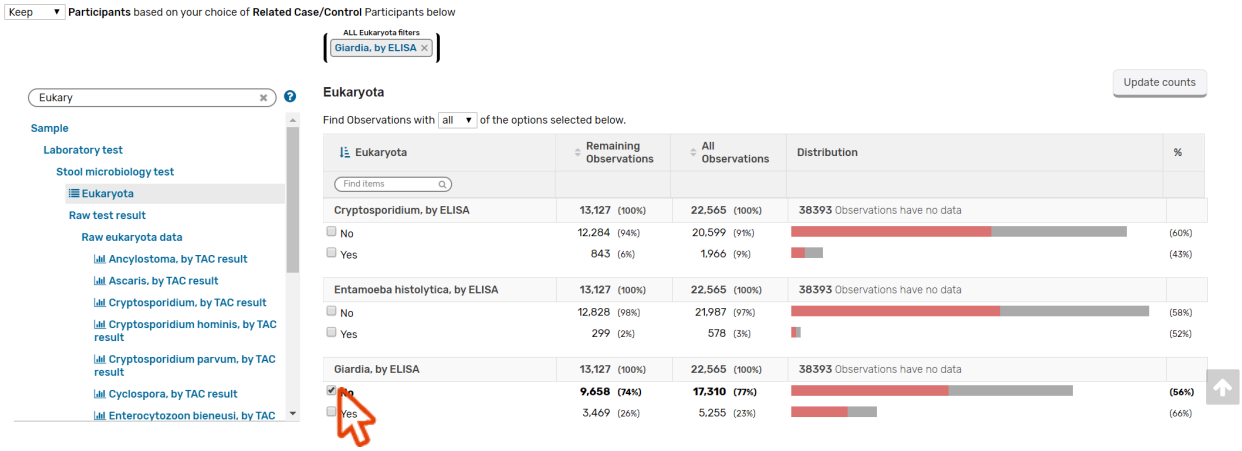
6. Then, click on the small check-box next to the sentence 'Enable the Related observations filter below. Enabling this option will allow you to restrict Observations by relating them to your choice of Related observations.'

Think carefully about what we want to do. We want to keep the selected Cases that have a Control that is Giardia negative.

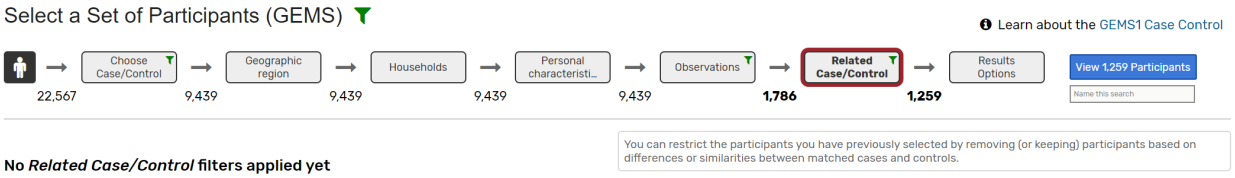
Complete the sentence like this:



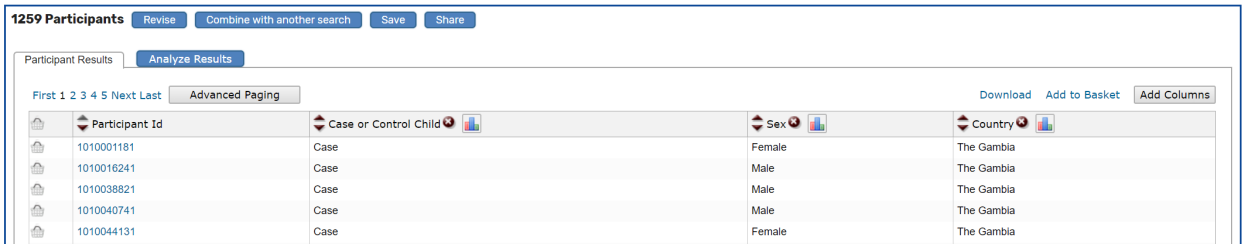
7. Then, select *Giardia* 'No' from the Stool microbiology test, Eukaryota filter.



How does this impact the number of participants that match your search? There should only be 1,259 participants remaining.

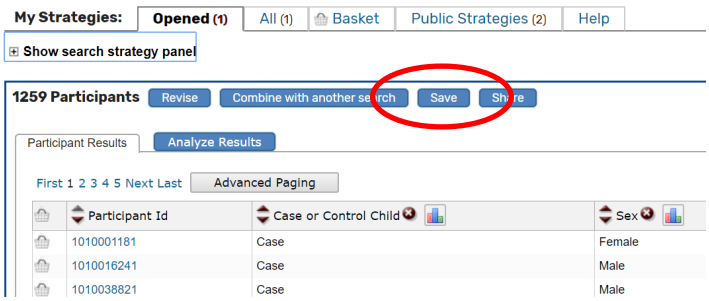


8. Click on the 'View 1,259 Participants' box to see the table listing of these Participants. Click on the small histogram icons next to the column names to see the distribution of characteristics in this sample.



To change the default columns listed in the result click on the 'Add Columns' button located at the top right corner of your Participant result tab and add the desired columns. Try adding 'Child given untreated drinking water' to the columns.

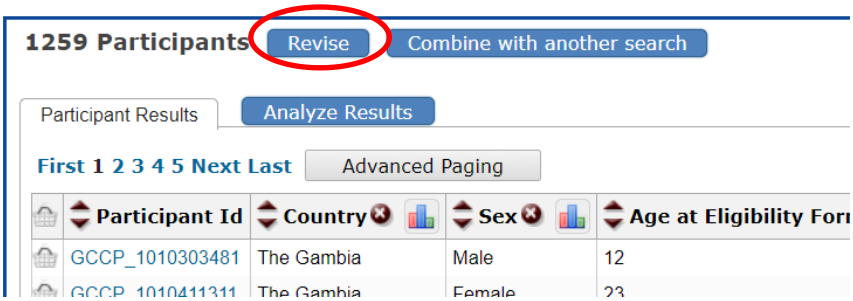
9. Save this search by clicking on the 'Save' button above the search results. You will need to login to save.



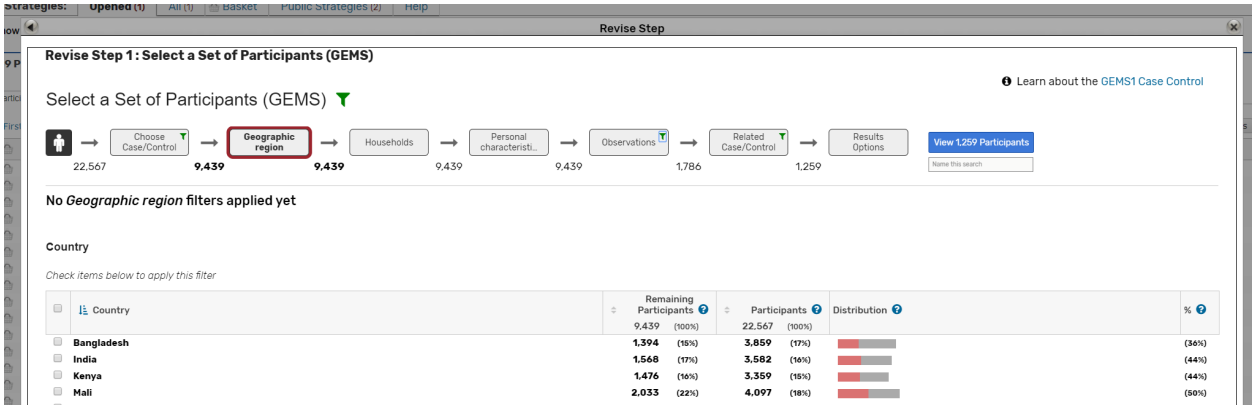
Revising Your Search

From your previous search, you now know that there were 1259 Cases in the GEMS study that were positive for *Giardia* had at least one matched control that was negative for *Giardia*. Maybe now you would like to ask, what is the relationship between Giardia and moderate-to-severe diarrhea in the first two years of life? For matched case-control studies the most robust analysis of associations is to be done using a matched pairs analysis. Our apps are for exploration and hypothesis generation purposes only and a matched pairs analysis is not available at this time. However, you can generate a simple contingency table to look at associations.

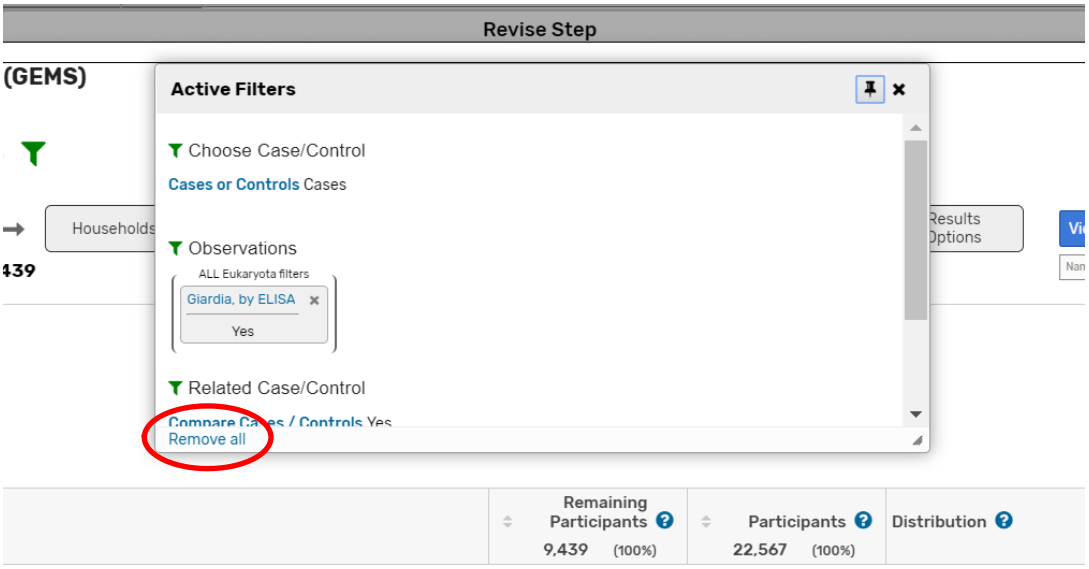
To look at the relationship between Giardia and moderate-to-severe diarrhea for your whole sample, rather than just the 1259 participants you just returned, you will want to revise your search to remove any filters you've applied. To do this click the revise button as follows:



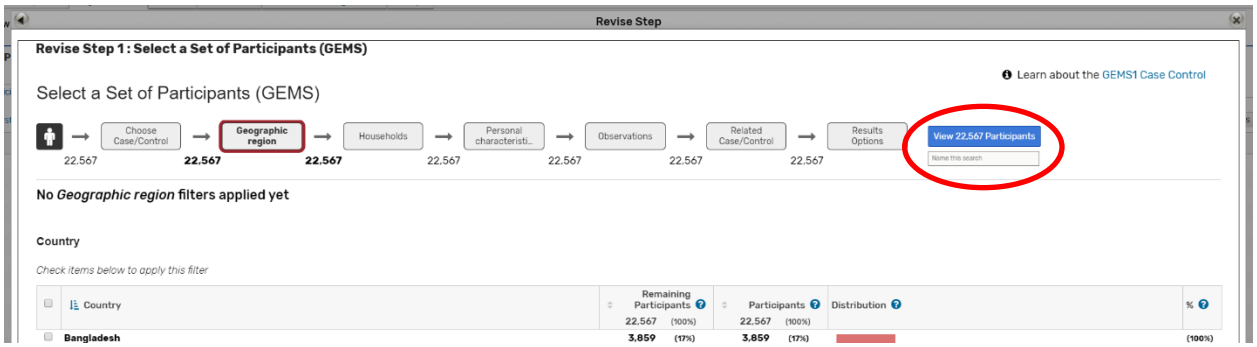
1. Click on the blue revise button above your results table.
2. Once you click on the revise button, a popup will appear that includes the search wizard with your previously selected parameters. To remove all previously applied filters, simply click on the green 'filter' icon anywhere in the pop-up. Another pop-up will open that lists all of your active filters.



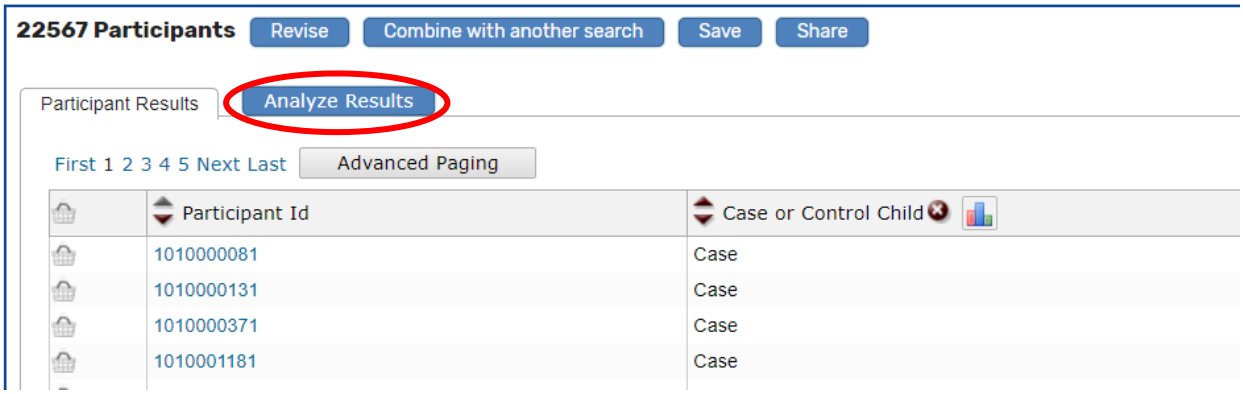
Click on ‘Remove all’ to remove all of the active filters then close the ‘Active Filters’ box with the ‘x’ in the top right corner.



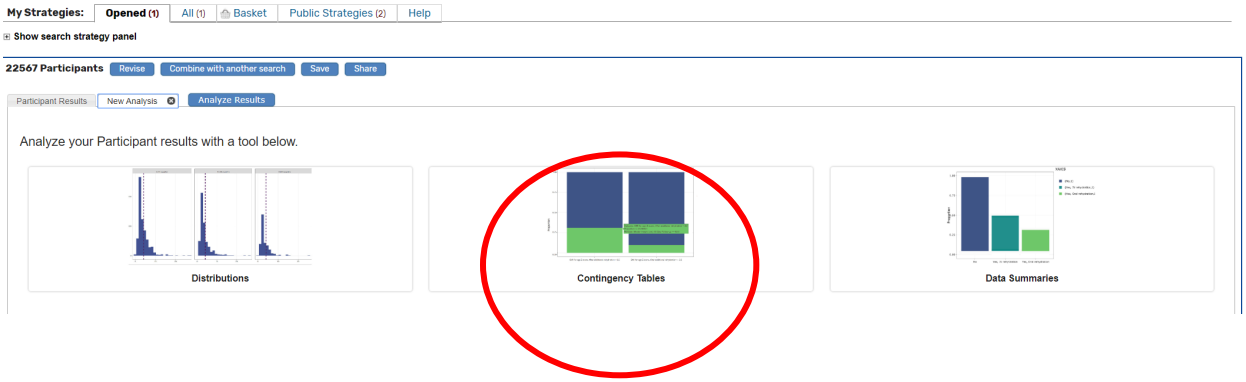
3. Finally, click ‘View 22,567 participants’ to get back to the full GEMS1 data set.



4. To analyze the association between MSD and Giardia, two binary outcomes, you want to run a contingency table. Click on the blue tab next to the Participant Results tab in your data table that says “Analyze Results”

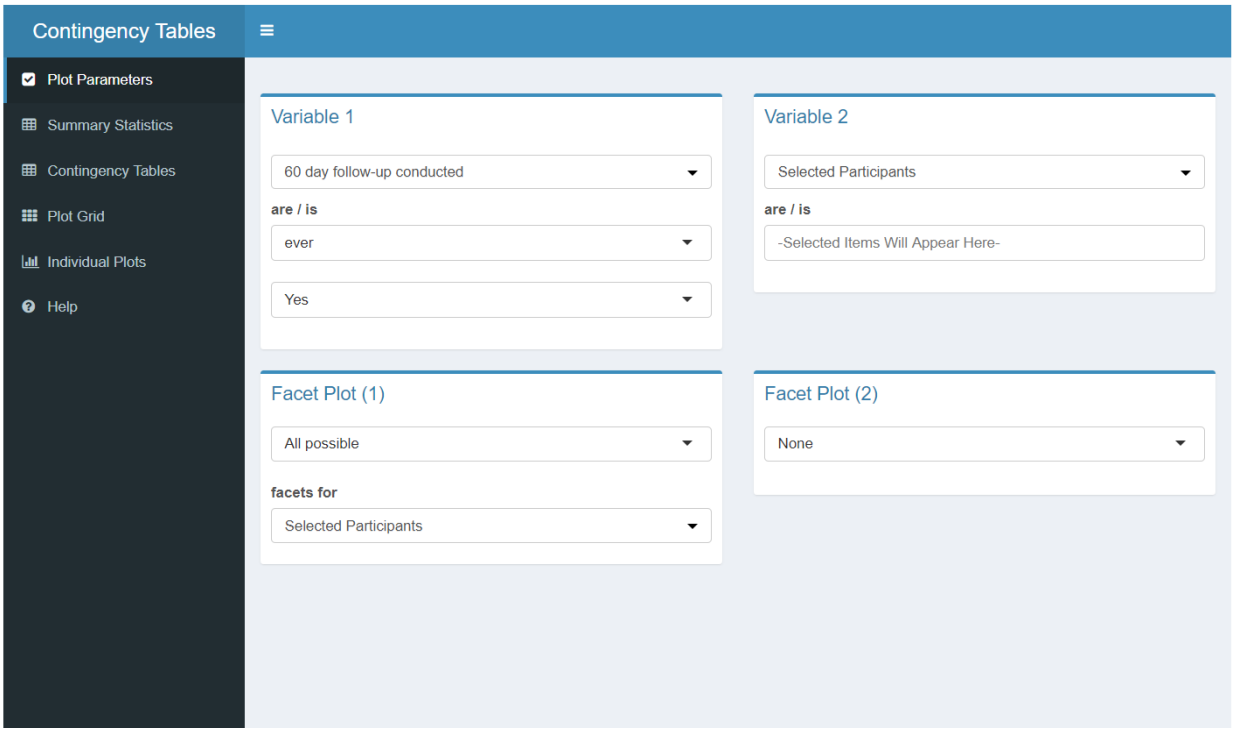


5. Initiate the app by clicking on the box labeled “Contingency Tables”



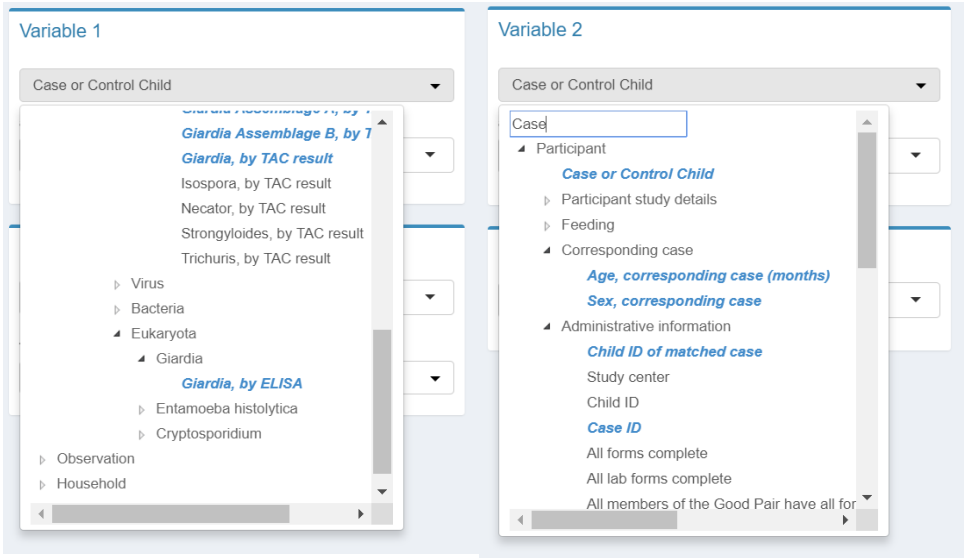
This Shiny app allows you to look at the association between any two variables that are found in the dataset. Notice that there are drop-down menus that allow you to choose what variables you are comparing. Variable 1 defines the rows that represent subsets of your population and Variable 2 defines the columns.

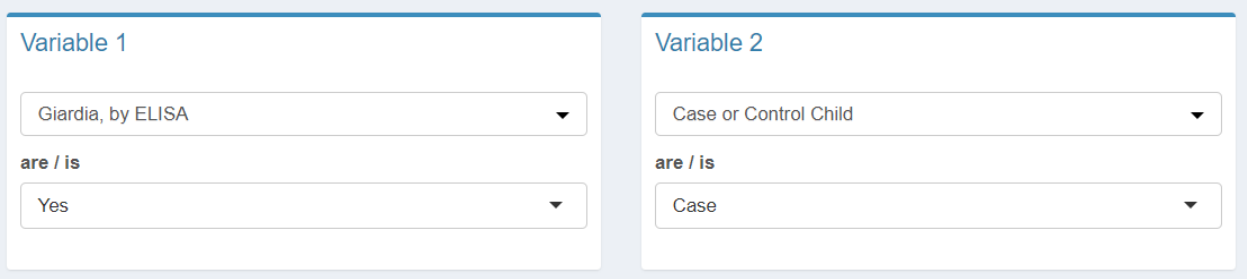
7. Examine the default characteristics of the app.



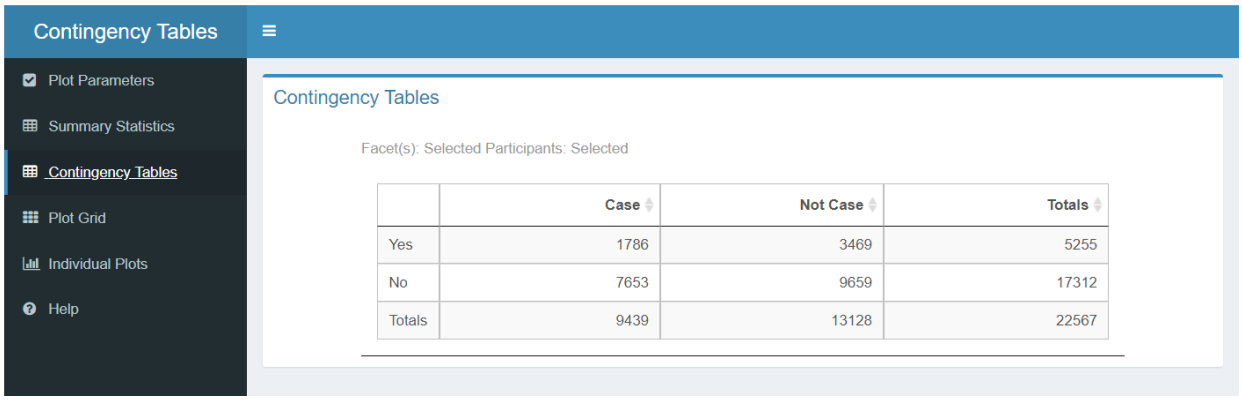
As you can see, in the large right panel you have drop down menus that allow you to select which variables you wish to compare. By default, for Variable 1, which defines the rows of the 2x2 table, 60 day follow-up conducted is selected. For variable 2, which defines the columns the participants you have selected (all of them in this case since you have removed all filters) are selected.

8. For this analysis start by selecting 'Case/Control' for Variable 1 and 'Giardia, by ELISA' for Variable 2.





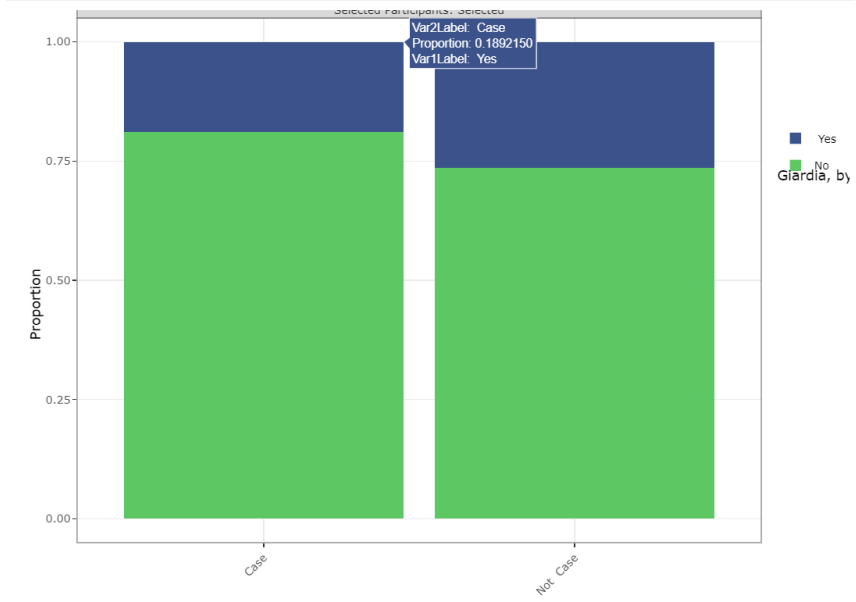
9. Next, click on the 'Contingency Tables' tab from the left panel. You can see that the numbers within the table indicate the number of participants that meet the defined criteria. For example, 1786 participants tested positive for Giardia by ELISA were cases. 3469 participants who tested positive for Giardia by ELISA were controls. The marginal and grand totals are also calculated.



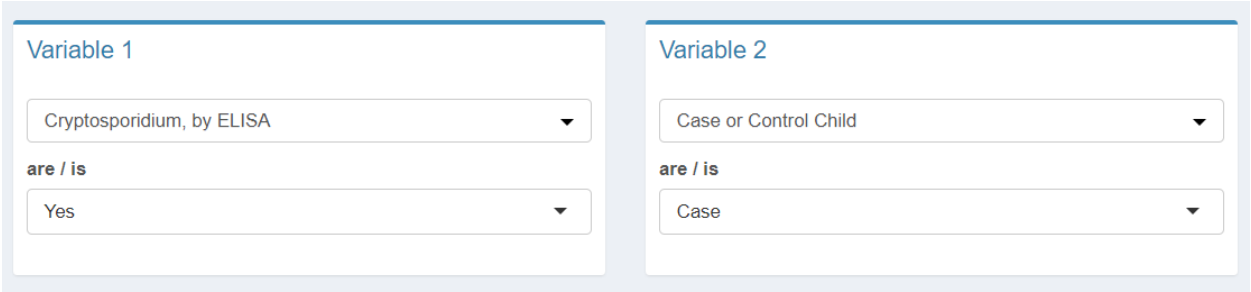
	Case	Not Case	Totals
Yes	1786	3469	5255
No	7653	9659	17312
Totals	9439	13128	22567

10. Clicking on 'Summary Statistics' shows the estimates Odds Ratios and Relative Risks and their p-values. Remember that these are just estimates and should not be over-interpreted. You'll notice that in these participants, testing positive for Giardia appears to be significantly associated with controls.

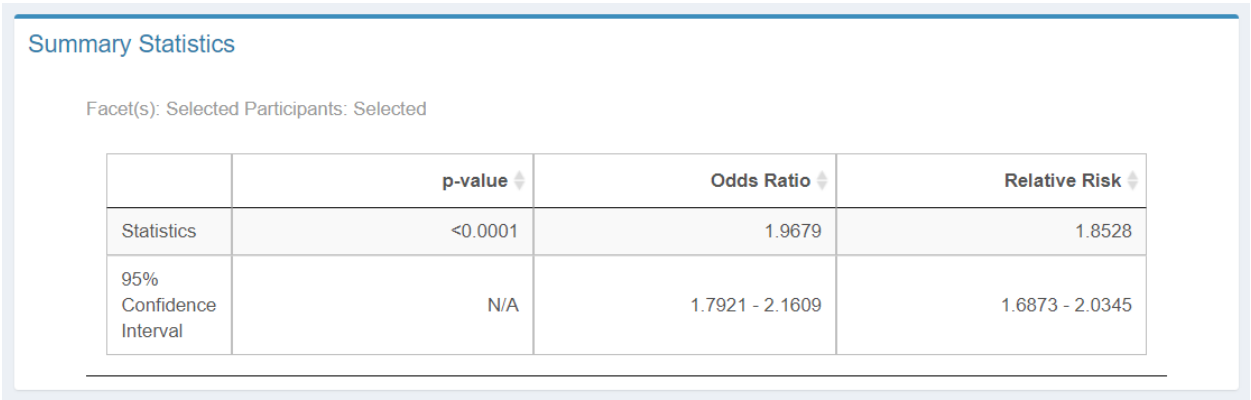
11. 'Plot Grid' will display a figure in green and blue which is a proportional physical representation of these counts. The blue box on the left hand side corresponds to the 1786 participants who tested positive for Giardia by ELISA and were cases while the green beneath it represents those that were Giardia negative cases. If you hover over any of the color shaded areas a pop-up box will tell you the attributes that you have specified and what proportion meet this group. In this case approximately about 19% of Cases are positive for Giardia.



8. Next take a look at another common pathogen. For Variable 1, select 'Cryptosporidium' from the drop down menu below. For Variable 2, keep Case.



What is the association between Cryptosporidium and MSD? Is it significant? How does this differ from Giardia?

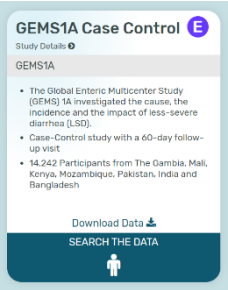


	p-value	Odds Ratio	Relative Risk
Statistics	<0.0001	1.9679	1.8528
95% Confidence Interval	N/A	1.7921 - 2.1609	1.6873 - 2.0345

Finally, you may want to look at your associations by different strata, by site or by age group for example. To do this, click back to the 'Plot Parameters' tab.

Is there a significant association between Cryptosporidium and MSD in children less than 2 years of age? How about greater than 2 years of age? You can easily select different strata and define different groups.

9. If you have time see if you can replicate this analysis quickly in the GEMS1A dataset. Are the result similar? Again, the results of these simple contingency tables are not meant to be complete analyses but rather hypothesis generating tools. You may want to see if you see a similar pattern in other pathogens? The ability to point and click and change variables easily allows you to examine your data in ways that may never had occurred to you before!



You have completed the exercise for this section! Was this query of related Cases and Controls useful? What other concordant or discordant pairs might you want to look for? Please let us know if you have any questions or submit any comments via the 'Contact Us' link located in bottom right corner of any of our web pages or by emailing help@clinepidb.org