

Exercise 3.2: PROVIDE: perform a complex search

This exercise will give you an opportunity to practice utilizing the search wizard to perform a complex search. You will gain experience using “Related observations” and “Adding a step” to your search strategy.

Within the group of participants who were vaccinated/not vaccinated with Rotarix as specified in the study protocol and who were followed in the study until they reached 2+ years of age, determine:

- A. How many times was *Rotavirus* detected by qPCR AFTER the participant was at least 17 weeks of age (the age at which the second dose of Rotarix was given)?
- B. How many of the *Rotavirus* detections in (A) occurred in participants randomized to receive Rotarix?
- C. In the set of samples that corresponds to the *Rotavirus* detections in (A), how many samples were also positive for *Norovirus*?
 1. What type of search should we execute to determine how many times *Rotavirus* was detected within your population of interest?¹ On the ClinEpiDB.org homepage, click on the appropriate icon on the PROVIDE study card to initiate this search type.
 2. Next, we need to start defining our study population of interest. Use “**Rotavirus vaccination per protocol**” to determine which participants received the Rotarix vaccine as specified in the study protocol. Note that the population treated “Per protocol” includes individuals assigned to receive Rotarix oral rotavirus at both 10 and 17 weeks of age and also control individuals who were assigned to NOT receive Rotarix.

PROVIDE Randomized Controlled Trial D




[Study Details](#)

1 Site in Bangladesh, 2011-2014

- 2 x 2 randomized controlled trial with 2 years follow-up
- 700 participants with >450,000 observations
- The Performance of Rotavirus and Oral Polio Vaccines in Developing Countries (PROVIDE) study evaluated the efficacy of delayed-dose oral rotavirus vaccine and the benefit of injectable polio vaccine replacing one dose of oral polio vaccine

[Download Data](#)

EXPLORE THE DATA

Remember that you can type text into the “Find a variable” search box ABOVE the search wizard to quickly locate a variable, regardless of which box in the search wizard it has been organized under.

How many participants were treated “Per protocol”?² How many observations were there in these participants?³


¹ Since we are interested in the total number of times something occurred, we should start an Observation-level search. Remember that since PROVIDE data was collected longitudinally, each participant may have had more than one instance of *Rotavirus* detection.

² 678 participants were treated as indicated in the study protocol, meaning that they were each either given the Rotarix oral rotavirus vaccine in accordance with the treatment schedule (if assigned to the treatment arm) or were not vaccinated (if assigned to the control arm).

³ There were 454,540 observations of these 678 participants.

Select a Set of Observations (PROVIDE) [Study Details >>](#)

Summary: The Performance of Rotavirus and Oral Polio Vaccines in Developing Countries (PROVIDE) study was a randomized controlled trial evaluating oral vaccine efficacy in Bangladesh. Infants aged 0-7 days old with no congenital abnormalities or birth defects and no abnormal stools since birth were enrolled with their mothers and assigned to one of 4 intervention arms in a 2 x 2 factorial trial design. Participants were followed for two years. A visual depiction of collection timepoints for various data types can be found on the PROVIDE Study Page. Use an observation search for accurate observation counts and a sample search for accurate sample counts.



→

Study Details

→

Household

→

Participant 

→

Observation

→

Related Observation

View 454,540 Observations







Name this search

467,317
467,317
467,317
454,540
454,540

Your Participant filters reduce 467,317 Observations to 454,540 Filter based on participant variables that generally cannot change over time.

expand all | collapse all

Find a variable


-  Years living in dwelling
- Administrative information
 -  Protocol version at enrollment
 -  Polio vaccination per protocol
 -  Polio vaccination dose per protocol
 -  Received extra OPV
 -  Rotavirus vaccination per protocol
- Aggregated illness data
- Aggregated immunization and prevention data
- Clinical history

Rotavirus vaccination per protocol
(Provider label: rotaprotd)

☐ Keep checked values at top 700 (100%) of 700 Participants have data for this variable

	Remaining Participants	Participants	Distribution	%
	700 (100%)	700 (100%)		
<input type="checkbox"/> No vaccination week 10	1 (< 1%)	1 (< 1%)		(100%)
<input type="checkbox"/> No vaccination week 10 or 17	1 (< 1%)	1 (< 1%)		(100%)
<input type="checkbox"/> No vaccination week 17	4 (1%)	4 (1%)		(100%)
<input checked="" type="checkbox"/> Per protocol	678 (97%)	678 (97%)		(100%)
<input type="checkbox"/> Week 10 vaccination outside window	9 (1%)	9 (1%)		(100%)
<input type="checkbox"/> Week 17 vaccination outside window	7 (1%)	7 (1%)		(100%)

- Next, further refine your study population of interest, limiting participants to those who were followed in the study until they reached 2+ years of age. Use **“Withdrawal before 731 days old”** to identify the set of participants who were followed until they were at least 2 years of age. Amongst your selected participants who received (or did not receive) the Rotarix vaccine according to the protocol, how many withdrew before 2 years of age (and should be removed from your subset of selected participants)?⁴



→

Study Details

→

Household

→

Participant 

→

Observation

→

Related Observation

View 440,545 Observations

Name this search




467,317
467,317
467,317
440,545
440,545

Your Participant filters reduce 467,317 Observations to 440,545 Filter based on participant variables that generally cannot change over time.

Rotavirus vaccination per protocol
Withdrawal before 731 days old

withdr

Participant study details

-  Withdrawal before 731 days old
-  Reason for withdrawal
-  Withdrawal date

Withdrawal before 731 days old
Participant was dropped/terminated at age < 731.
(Provider label: drop730)

☐ Keep checked values at top 700 (100%) of 700 Participants have data for this variable

	Remaining Participants	Participants	Distribution	%
	678 (100%)	700 (100%)		
<input checked="" type="checkbox"/> No	582 (86%)	597 (85%)		(97%)
<input type="checkbox"/> Yes	96 (14%)	103 (15%)		(93%)

⁴Of the 678 participants who received “Rotavirus vaccination per protocol”, 96 withdrew before 2 years of age.

- Now, let's look at Rotavirus using “Rotavirus, by qPCR” to limit our analysis to instances when *Rotavirus* was detected. How many times was *Rotavirus* EVER detected in our population of interest?⁵

467,317 → 467,317 → 467,317 → 440,545 → 457 → View 457 Observations

Filter based on observation variables (anthropometric measurements, clinical symptoms, survey results, lab results, etc.). Data was collected at multiple timepoints so may require a multi-step strategy to return all data of interest.

Your Observation filters reduce 440,545 Observations to 457

Rotavirus, by qPCR

rotavirus by q

Sample

Laboratory test

Stool microbiology test

Virus in stool

Rotavirus in stool

Rotavirus, by qPCR

Raw test result for stool

Raw virus data for stool

Rotavirus Ct value, by qPCR result

Keep checked values at top

3,214 (1%) of 467,317 Observations have data for this variable

	Rotavirus, by qPCR	Remaining Observatio...	Observatio...	Distribution	%
		3,050 (100%)	3,214 (100%)		
<input type="checkbox"/> Negative	2,593 (85%)	2,732 (85%)			(95%)
<input checked="" type="checkbox"/> Positive	457 (15%)	482 (15%)			(95%)

- Finally, to determine the number of times *Rotavirus* was detected by qPCR AFTER the participant was 17 weeks of age (the age at which the second dose of Rotarix was given in the treatment arm), we will need to use “Related observations”. Click to enable the advanced “Related observation” filter.

467,317 → 467,317 → 467,317 → 440,545 → 457 → View 0 Observations

Filter selected observations based on a time-dependent relationship to a second set of observations defined below. For example, keep only diarrheal observations that were followed by a shigella-positive lab test 0-5 days later.

Your Related Observation filters reduce 457 Observations to 0

☒ Enable the advanced Related Observation filter below. It allows you to restrict Observation by relating them to your choice of Related Observation.

- Remember that the “Related observations” step will require us to (i) complete the sentence indicating which observations to keep or remove and (ii) define the related observations of interest. Let's define the related observations of interest now. We will determine how to complete the sentence next.

There were 457 observations where *Rotavirus* was EVER detected in our population of interest (Search 1). observations. For the related observations search, we want to find all of the observations that occurred in our population of interest AFTER the participants were 17 weeks old. Use “Observation type” for this search (Search 2).

⁵ *Rotavirus* was detected in our population of interest 457 times.

type		Observation type			
		(Provider label: ext_1, f104vn, f10vn, f14vn, f17vn, f18vn, f39vn, f40vn, f52vn, f53vn, f6vn, fanvn, flmvn, sbvn, vnum, vstnum)			
		Keep checked values at top			
		467,317 (100%) of 467,317 Related Observations have data for this variable			
	Observation type	Remaining Related Observatio...	Related Observatio...	Distribution	%
	Find items	467,317 (100%)	467,317 (100%)		
<input type="checkbox"/>	Adverse event/Severe adverse event	230 (< 1%)	230 (< 1%)		(100%)
<input type="checkbox"/>	Daily surveillance/diarrheal episode	455,541 (97%)	455,541 (97%)		(100%)
<input type="checkbox"/>	Enrollment	700 (< 1%)	700 (< 1%)		(100%)
<input type="checkbox"/>	Protocol deviation	1,685 (< 1%)	1,685 (< 1%)		(100%)
<input type="checkbox"/>	Week 6	693 (< 1%)	693 (< 1%)		(100%)
<input type="checkbox"/>	Week 10	647 (< 1%)	647 (< 1%)		(100%)
<input type="checkbox"/>	Week 12	633 (< 1%)	633 (< 1%)		(100%)
<input type="checkbox"/>	Week 14	639 (< 1%)	639 (< 1%)		(100%)
<input checked="" type="checkbox"/>	Week 17	622 (< 1%)	622 (< 1%)		(100%)
<input type="checkbox"/>	Week 18	609 (< 1%)	609 (< 1%)		(100%)
<input type="checkbox"/>	Week 24	604 (< 1%)	604 (< 1%)		(100%)
<input type="checkbox"/>	Week 39	614 (< 1%)	614 (< 1%)		(100%)
<input type="checkbox"/>	Week 40	612 (< 1%)	612 (< 1%)		(100%)
<input type="checkbox"/>	Week 52	606 (< 1%)	606 (< 1%)		(100%)
<input type="checkbox"/>	Week 53-56	594 (< 1%)	594 (< 1%)		(100%)
<input type="checkbox"/>	Week 65	578 (< 1%)	578 (< 1%)		(100%)
<input type="checkbox"/>	Week 78	565 (< 1%)	565 (< 1%)		(100%)
<input type="checkbox"/>	Week 91	562 (< 1%)	562 (< 1%)		(100%)

7. How should we complete the sentence below to properly combine our related observations (Search 2: observations occurring post-17 weeks of age) with our observations (Search 1: observations where *Rotavirus* was detected)?⁶

Keep / Remove Observation that are # to # days Before / After the Related Observation specified below

8. You should be left with 385 observations where:
- Participants were treated according to the study protocol, and either received Rotarix vaccine on schedule or were not vaccinated (if in the control arm).
 - Participants were followed until they were at least 2 years of age.
 - Rotavirus* was detected AFTER the participant was 17 weeks of age.

⁶ **KEEP** observations that are **0 to 756** days **AFTER** the related observations specified below. We want to KEEP observations (Search 1, where *Rotavirus* was detected) that occurred AFTER the participants were 17 weeks old (the related observations, Search 2).

Why was the range of days for the related observations set to “0 to 756”?

- We are interested in observations that occurred at ANY point in the study after the participant was 17 weeks old, so the minimum was set to 0 (when the participant was 17 weeks old).
- When we look at “Age (days)” we can see that the maximum age in the study is 756 days of age. In order to return the full set of observations where the participant was at least 17 weeks of age, we could have completed the range with any number greater than the difference between the maximum age (756) and the age at 17 weeks (119). It's easier to just use 756 instead of calculating that difference.



- Now that we have limited our observations to those where *Rotavirus* was detected when our population of interest was at least 17 weeks old, we are ready to answer (B): How many of these *Rotavirus* detections occurred in participants randomized to receive Rotarix? Let's click on the blue "View 385 observations" button in the search wizard to get to the results page.

We are interested in "**Rotarix vaccine randomization**", which will tell us which of these observations occurred in participants who were in the control arm (randomized to receive no rotavirus vaccine) versus the treatment arm (and received the Rotarix vaccine). Click on "Add columns" button on the right-hand side above the Observations table to add this variable.

- Click on the histogram icon next to the "Rotarix vaccine randomization" column header to view the distribution of observations in participants who received versus did not receive the vaccine. How many observations occurred in participants who received both doses of Rotarix according to the study schedule?⁷ *Hint: hover over each of the bars in the histogram, or click on the "Data" tab.*

385 PROVIDE Observations [Revise this search](#)

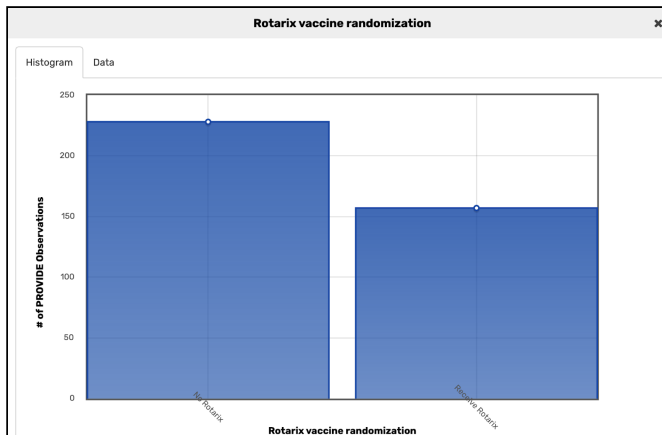
PROVIDE Observation Results [Analyze Results](#)

Rows per page: 20

[Download](#) [Add to Basket](#) [Add Columns](#)

Observation Id	Observation type	Observation date	Rotarix vaccine randomization
obp1027_-9_2013-11-02	Daily surveillance/diarrheal episode	2013-11-02 00:00:00.0	Receive Rotarix
obp1056_-9_2012-05-05	Daily surveillance/diarrheal episode	2012-05-05 00:00:00.0	No Rotarix
obp1061_-9_2013-04-21	Daily surveillance/diarrheal episode	2013-04-21 00:00:00.0	Receive Rotarix

⁷ 157 observations occurred in participants who received both doses of Rotarix.



11. Now let's answer question (C): In the set of samples that corresponds to the *Rotavirus* detections in (A), how many samples were also positive for *Norovirus*?

To answer this question, let's transform our set of observations into the set of samples where *Rotavirus* was detected. Click on "Add a step" in your search strategy, and follow the instructions in the popup windows to transform these 385 observations into their corresponding set of samples.

My Search Strategies

[Opened \(1\)](#) [All \(56\)](#) [Public \(26\)](#) [Help](#)

Unnamed Search Strategy *

PROVIDE

385 Observations

Step 1

+ Add a step

385 PROVIDE Observations [Revise this search](#)

PROVIDE Observation Results [Analyze Results](#)

1

2

3

...

20

Rows per page: 20

[Download](#)
[Add to Basket](#)
[Add Columns](#)

Observation Id	Observation type	Observation date	Rotarix vaccine randomization
obp1027_-9_2013-11-02	Daily surveillance/diarrheal episode	2013-11-02 00:00:00.0	Receive Rotarix

← Add a step to your search strategy ?

Combine with other PROVIDE Observations

Transform into related records

Transform 385 PROVIDE Observations into...

PROVIDE Participants

PROVIDE Samples

← Add a step to your search strategy ?

Your Observations from Step 1 will be converted into Samples

Run Step

? Give this search a name (optional)

? Give this search a weight (optional)

12. We are interested in “**Norovirus, by multiplex qRT-PCR**”, which will tell us whether these samples had *Norovirus* detected or not. Click on “Add columns” button on the right-hand side above the Samples table to add this variable.

385 PROVIDE Samples [Revise this search](#)

PROVIDE Sample Results

Rows per page: 20

Sample Id	Poliovirus, by culture	Rotavirus, by qPCR	Hematocrit (%)	Blood test >> Interferon gamma (pg/mL)
sbp1027_9_2013-11-02	N/A	Positive	N/A	N/A
sbp1056_9_2012-05-05	N/A	Positive	N/A	N/A
sbp1061_9_2013-04-21	N/A	Positive	N/A	N/A
sbp1095_9_2012-11-22	N/A	Positive	N/A	N/A

Select Columns

Update Columns

add these | clear these | select only these
select all | clear all | reset to current | reset to default

norovirus

☐ Public health and epidemiology

☐ Sample

☐ Norovirus Ct value, by qRT-PCR result

☐ Norovirus GI Ct value, by TAC result

☐ Norovirus GII Ct value, by TAC result

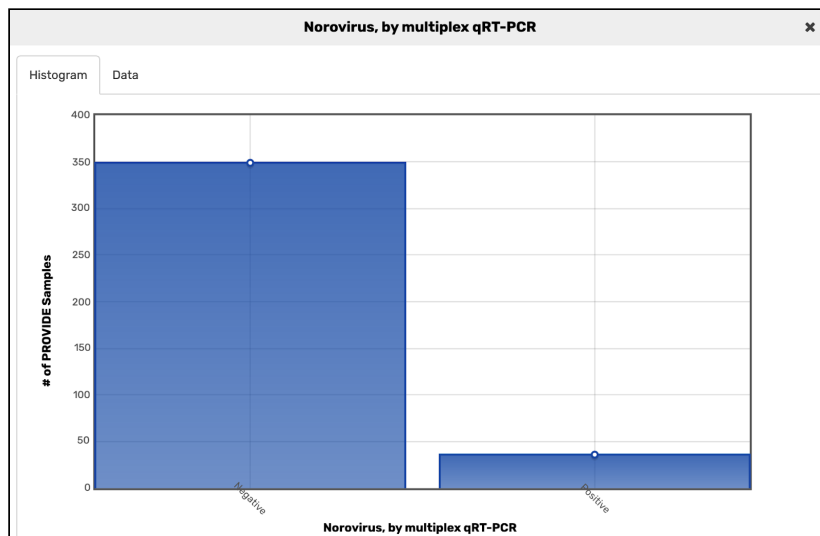
☒ Norovirus, by multiplex qRT-PCR

add these | clear these | select only these
select all | clear all | reset to current | reset to default

Update Columns

Download Add to Basket Add Columns

13. Click on the histogram icon next to the “Norovirus, by multiplex qRT-PCR” column header to view the distribution of samples that tested positive versus negative for *Norovirus*. How many of our samples of interest were also positive for *Norovirus*?⁸



⁸ 36 of our samples of interest were also positive for *Norovirus*.