

Exploring Enteric Disease Datasets

Comparing pathogen detection methods

With continuing advances in molecular diagnostics many epidemiologic studies are beginning to adopt new, culture-independent diagnostics for detecting pathogens. In ClinEpiDB we have loaded both conventional microbiology test results as well as the TaqMan gene expression array cards (TAC) for GEMS1 and MAL-ED. The conventional diagnostics are generally presented as categorical variables in **(a)** their unprocessed form where each value indicates the specific *Shigella* subgroup of serotype that was identified and/or as **(b)** a binary analytic variable which indicates whether or not any *Shigella* was found in the stool sample. Note that **(a)** and **(b)** are two different versions of the same result. Alternatively, **(c)** indicates a completely separate test result - the Ct value from TAC. The maximum Ct value for the majority of the TAC results are truncated at 35.0, which was done during data cleaning as a means of establishing the analytical limit of detection.

(a)

<input type="checkbox"/>	Shigella, by bacteriology result	Remaining Observatio... ?	Observatio... ?	Distribution ?	% ?
		39,837 (100%)	39,837 (100%)		
<input type="checkbox"/>	No	39,492 (99%)	39,492 (99%)	<div></div>	(100%)
<input type="checkbox"/>	<i>Shigella boydii</i> GpC	55 (0%)	55 (0%)		(100%)
<input type="checkbox"/>	<i>Shigella dysenteriae</i> GpA	35 (0%)	35 (0%)		(100%)
<input type="checkbox"/>	<i>Shigella flexneri</i> GpB	142 (0%)	142 (0%)		(100%)
<input type="checkbox"/>	<i>Shigella sonnei</i> GpD	105 (0%)	105 (0%)		(100%)
<input type="checkbox"/>	<i>Shigella</i> sp non-typable	9 (0%)	9 (0%)		(100%)

(b)

Bacteria	Remaining Observations	All Observations	Distribution	%
Shigella				
Shigella, by bacteriology	39,837 (100%)	39,837 (100%)	1344486 Observations have no data	
<input type="checkbox"/> No	39,492 (99%)	39,492 (99%)	<div></div>	(100%)
<input type="checkbox"/> Yes	346 (1%)	346 (1%)		(100%)

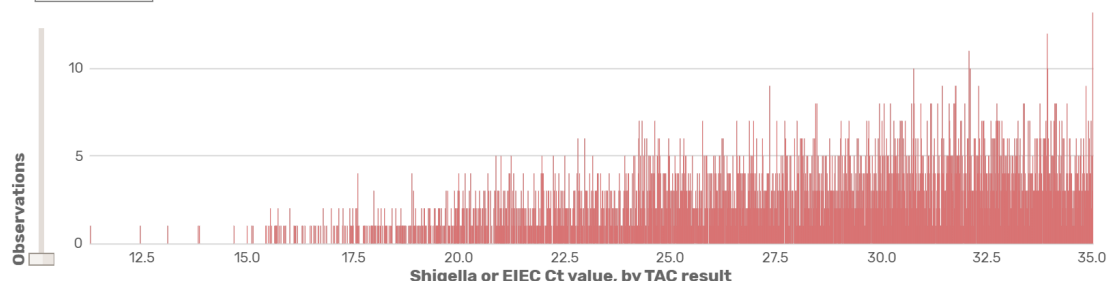
(c)

Shigella or EIEC Ct value, by TAC result

Mean: 34.23 Min: 11.27 Max: 35

Select Shigella or EIEC Ct value, by TAC result from to

41,405 (3%) of 1,384,323 Observations have data for this filter



Use the stool microbiology test results from the MAL-ED study to answer the following questions:

- # of participants who had a Shigella or EIEC Ct value, by TAC result of exactly 35 _____¹
- # of participants who had a Shigella or EIEC Ct value, by TAC result less than 25 [Note to get less than any # you will either have to point and select very carefully on the histogram or type ##.99 into the result selection maximum.] _____²
- # of participant who had a bacteriology result indicating Shigella flexneri Group B
_____³
- # of participants who had Cryptosporidium identified by modified acid stain _____⁴
- # of participants who had a positive Cryptosporidium, by ELISA result _____⁵
- # of participants who had a Cryptosporidium Ct value less than 35 _____⁶

Notice that the conventional microbiology results often differ from the molecular method results. Generally, molecular diagnostics may have a greater sensitivity but a lower specificity than culture-based methods. Both GEMS and MAL-ED reanalyzed primary study objectives using quantitative molecular diagnostic assays following publication of conventional method results. Without even downloading the dataset we can use ClinEpiDB to do exploratory analyses on how results differ between the conventional and diagnostic methods.

¹ 1,715

² 739

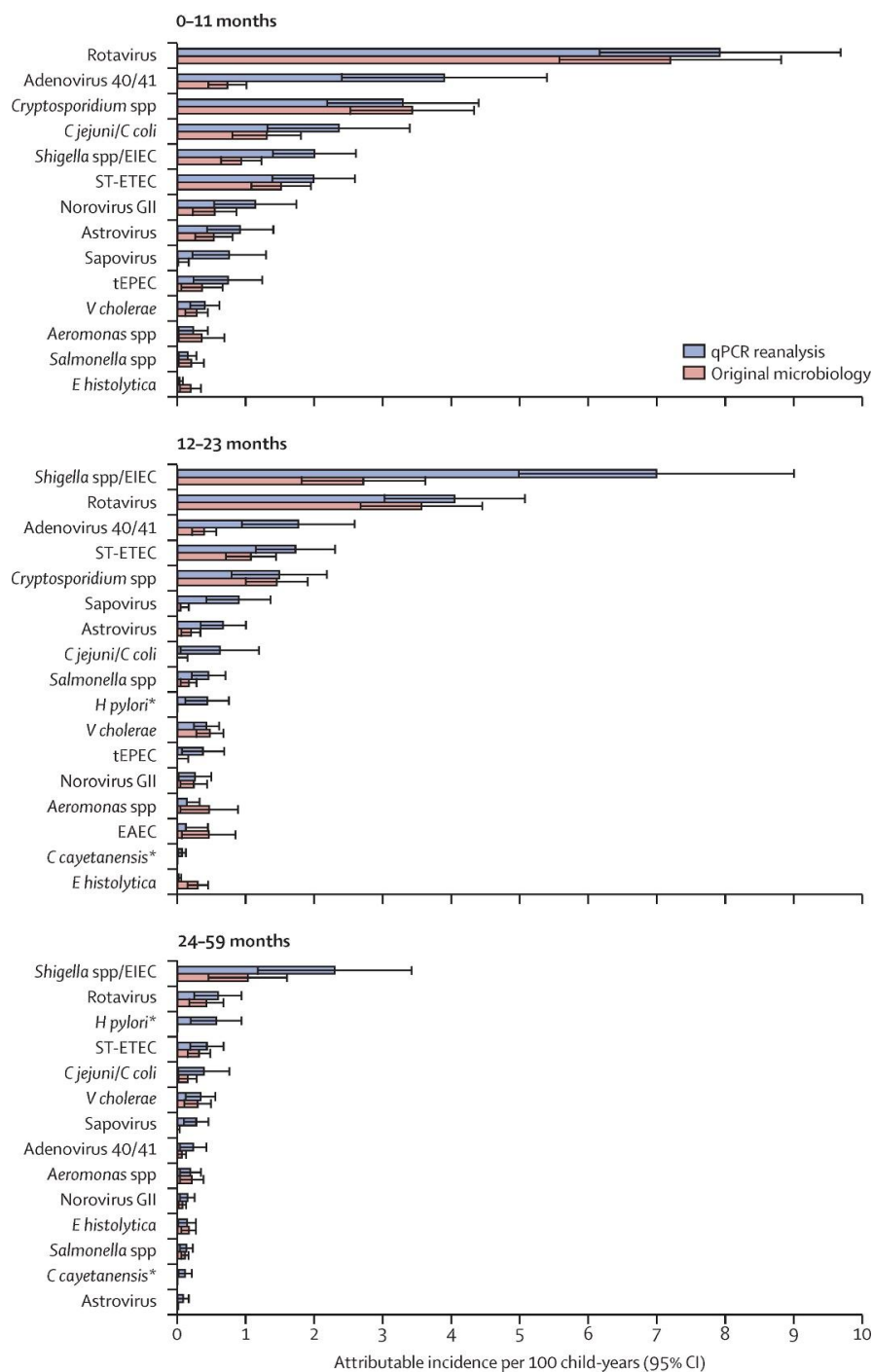
³ 204

⁴ 648

⁵ 1007

⁶ 541

Examining the difference between conventional and quantitative molecular diagnostics of ST-ETEC in GEMS



A remarkable finding from the GEMS study is that re-analysis of the stool samples using quantitative PCR (qPCR) produced higher estimates for *Shigella*, increasing the attributable incidence by about 2-fold for *Shigella* and 1.5-fold for ST-EPEC (STh-producing enterotoxigenic *E. coli*).⁷ We can estimate an approximation of this result using ClinEpiDB.

⁷ J. Liu, J.A. Platts-Mills, J. Juma, et al. 2016. Lancet. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study.

For ST-EPEC in GEMS1, how many participants overall tested positive by conventional PCR? (EPEC ST-pos, by PCR) _____ Proportion: _____⁸

Compare this to, how many participants overall tested positive by qPCR (TAC)? (EPEC ST-pos Ct value, by TAC result <35) _____ Proportion: _____⁹

Remember that if a diagnostic test is more sensitive it may increase detection in diarrheal disease cases *but also* controls. The increased detection will not result in an impact to attributable incidence *unless* it is differential in cases versus controls.

Cases: EPEC ST-pos, by PCR _____ Proportion _____¹⁰

Controls: EPEC ST-pos, by PCR _____ Proportion _____¹¹

Difference in Case vs. Control Proportion: _____¹²

Cases: EPEC ST-pos Ct value, by TAC result <35) _____ Proportion _____¹³

Controls: EPEC ST-pos Ct value, by TAC result <35) _____ Proportion _____¹⁴

Difference in Case vs. Control Proportion: _____¹⁵

Which detection method had a greater difference in the proportion positive in cases vs. controls?¹⁶

Finally, look at what happens to the conventional test results when you vary the threshold for positivity by Ct value. The advantage of a quantitative test is that it may give you more information about pathogen load and/or disease severity.

While you cannot do a direct comparison to the MAL-ED study because of the difference in study design and distinct case definition for GEMS, take these next few minutes to explore MAL-ED in the same way. Do the results of the conventional detection method for ST-EPEC differ from the TAC results?

⁸ 1,009, 4%

⁹ 1,957, 8.7%

¹⁰ 645, 6.8%

¹¹ 364, 2.8%

¹² 4.0%

¹³ 1,127, 11.9%

¹⁴ 830, 6.3%

¹⁵ 5.6%

¹⁶ TAC