Characterization of enteric disease in children using a low-cost specimen preservation method

Amanda K. Debes1*, Shaoming Xiao1, Jie Liu2, Allison Shaffer1, Paul Scalzo1, Etienne Guenou3,
Landry Beyala3, Goura Andre Pascal3, Anthony Njimbia Chebe3, Hirma Tchio-Nighie3, Nafack
Sonkeng Sonia3, Malathi Ram1, David A. Sack1, Jerome Ateudjieu3,4,5

1 Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland;
2 Department of Biomedical Sciences,
3 University of Virginia, Charlottesville, VA, USA.
4 M.A. SANTE (Meilleur Accès aux Soins de Santé), Yaoundé, Cameroon
5 Department of Public Health, Faculty of Medicine and Pharmaceutical Sciences, University of
Dschang, Dschang, Cameroon;
6 Clinical Research Unit, Division of Health Operations Research, Ministry of Public Health,
7 Yaoundé, Cameroon

* Correspondence: Amanda K. Debes PhD, MS
8 Johns Hopkins Bloomberg School of Public Health, Department of International Health
9 615 N. Wolfe Street, E5036, Baltimore, MD 21205
10 adebes1@jhu.edu

Running Title: Characterizing enterics in children

Keywords: Shigella, ETEC, enterics, diagnostics, pediatrics

Wordcount: Text – 2951; Abstract - 247
Abstract

BACKGROUND:
Diarrhea is a leading cause of death in children under five. Molecular methods exist for the rapid
detection of enteric pathogens; however, the logistical costs of storing stool specimens limit
applicability. We sought to demonstrate that dried filter paper specimen preservation can identify
diarrheal diseases causing significant morbidity among children in resource constrained
countries.

METHODS:
A sub-study was nested into cholera surveillance in Cameroon. Enrollment criteria included:
enrollment between 8/1/16 - 10/1/18; age < 18 years; a stool specimen; ≥ three loose stools
within 24 hours with the presence of dehydration and/or blood. 7227 persons were enrolled, for
which 2746 met enrollment criteria and 337 were included in this analysis using the enteric
TaqMan Array Card. Bacterial pathogens were compared to severity of diarrhea, age and sex,
among other variables.

RESULTS:
107 were ETEC positive of which: 40.2% (N=43) LT-STh, 19.6% (N=21) LT-STp; and 49.5%
(53) LT-only. Major CFs were present in 43.9% of ETEC-positives. 96 were positive for
Shigella, of which 14 (14.6%) reported dysentery. Model-derived quantitative cutoffs identified
116 (34.4%) with one highly diarrhea-associated pathogen and 16 (4.7%) with ≥
two. Shigella and rotavirus were most strongly associated with diarrhea in children with mixed
infections.

CONCLUSION:
Dried filter paper preserved specimens eliminate the need for frozen stool specimens and will facilitate enteric surveillance and contribute to the understanding of disease burden, which is needed to guide vaccine development and introduction. This study confirms Rotavirus, Shigella and ETEC as major contributors to pediatric diarrheal disease in two regions of Cameroon.
Background

Worldwide, nearly half a million children under five years of age die annually from diarrheal diseases, while millions more suffer from multiple episodes throughout early childhood. The morbidity associated with pediatric diarrheal disease has long-term consequences including malnutrition and associated physical and cognitive developmental delays. Malnutrition and resulting stunting increases the risk for a number of chronic disease states later in life. Nearly one-third of pediatric diarrheal deaths are due to Enterotoxigenic *Escherichia coli* (ETEC) and *Shigella*. In the reanalysis of the Global Enterics Multicenter Study (GEMS), ETEC and *Shigella* were among the top four pathogens, with *Shigella* having the highest rate of pathogen-attributable disease, which caused moderate-to-severe (MSD) diarrhea among children enrolled in the study. The use of the Taqman Array Card (TAC) demonstrated that *Shigella* and ETEC were two times greater than reported previously using classical microbiological methods of detection. Key stakeholders often lack the disease burden information they need to guide decisions regarding the prioritization of enteric vaccine development. Classical methods are often not feasible due to lack of laboratory capacity in the remote areas where vaccine trials need to be conducted. PCR methods such as the TAC have been shown to improve detection capabilities, allowing for nearly 90% of diarrheal episodes to be attributed to a specific pathogen, as compared to approximately 50% attribution using classical methods. The TAC also helps resolve case attribution in mixed infections, which are common among children under five in low-middle income countries (LMIC). In addition, more stringent quantitative cutoffs can be applied, such as “highly diarrhoea associated” applied in the GEMS studies.
Freezer storage is required for whole stool preservation with significant cost as it is not commonly available in LMICs. Therefore, we validated a modified version of the filter paper specimen preservation method previously developed for cholera surveillance(13) that allows for sustained integrity of the nucleic acid in LMICs.

Diarrhea surveillance was implemented in two regions of Cameroon previously identified as cholera endemic. The surveillance found high rates of cholera negative diarrhea among persons presenting for the study. All participants were screened by cholera rapid diagnostic test (RDT). The high rates of cholera negative diarrhea presented the opportunity to evaluate our specimen preservation method as a tool to improve understanding of pathogen-specific diarrheal disease burden. We nested a pediatric subgroup into on-going cholera surveillance hypothesizing that ETEC and Shigellae cause a large proportion of MSD which may be controlled with vaccines currently being developed. This study validated filter paper for specimen preservation at ambient temperature; and demonstrated that filter paper specimens can be utilized for PCR detection of diarrheal pathogens to improve understanding of disease burden in various age groups of children in Cameroon. Additionally, our burden estimate resulting from this analysis were consistent with other large-scale pediatric diarrheal studies based on culture and qPCR analysis of fresh or frozen stool samples.

Methods

Study Design

The Far North of Cameroon (FNC) is part of the Lake Chad Basin region of Africa. This rural area is unique in that in spite of its isolation it is a crossroads for communication among the bordering countries and, thus, is a center for commercial activities. Food security, water
availability, and access to health care are the poorest in Cameroon, likely further exacerbated by the extreme climate situation. These characteristics make this region vulnerable to diarrheal diseases.

In contrast, Douala, Littoral region, is Cameroon’s largest city. Douala has significant rainfall and a short dry season. The flat relief of the city, the high population density in the slums and the limited water drainage systems favor floods during the rainy season, contaminating wells and other drinking water sources. As a result, Douala is endemic for cholera.

In 2013, diarrheal surveillance began in these two regions including routinely collecting specimens from acute watery diarrhea (AWD) patients presenting to select health facilities. Consenting participants completed a questionnaire including data on access to WASH, participant and household characteristics, clinical information including the presence of blood in stool. Stool specimens were tested via RDT (Crystal-VC, Arkray, India) and 2 drops (~50ul) of stool were spotted and preserved on Whatman® 903 Protein Saver Card (Sigma-Aldrich, St. Louis, MO), stored at room temperature (RT).

Inclusion criteria for this sub-study included: enrollment between 8/1/2016 and 10/1/2018; patients <18 years of age who presented with three or more loose stools within 24 hours and who presented with dehydration (determined based on established WHO criteria) and/or blood and/or mucous in stool.

Sample Size Calculation:

We hypothesized that 20% of the diarrhea patients would be infected with Shigella, with a precision of 10%, we needed to recruit at least 62 children in each age group (0-2y, 3-5y, and 6-17y).

Nucleic Acid Extraction:
DNA was extracted from filter paper specimens using Chelex® 100 Chelating Resin (Bio-Rad, Hercules, CA) and the TAC was performed the same day. For each filter paper specimen, half of the filter paper circle was excised and extracted per previously published methods. (17) The bacteriophage MS2 and Phocine Herpesvirus 1 (PhHV) served as extrinsic controls. Acid-washed glass beads (Sigma-Aldrich, St. Louis, MO) were added and the specimen was homogenized for 2 minutes (Mini Beadbeater-8, BioSpec Products, Bartlesville, OK), prior to the boil-extraction step.

Additional controls included a template negative control and an extraction blank. Results from any contaminated targets were excluded from the analysis for that extraction batch. The standard curves for each target were created using synthetic targets, plasmids for DNA targets and in vitro transcripts for RNA targets. The standards were run in triplicates and served as the positive controls. To prepare specimens for the TAC, 80 µL of AgPath-ID One-Step RT-PCR Reagents (Thermo Fisher Scientific, Waltham, MA) was mixed with 20 µL of nucleic acid extract. After loading sample wells, the TAC was centrifuged twice at 1.2K rpm for 1 minute, sealed and loaded into the QuantStudio™ 12K Flex system (Applied Biosystems, Foster City, CA). The reaction volume was 1 µL for each well, and the cycling conditions were set as follows: 1 cycle at 45°C for 20 min, 1 cycle at 95°C for 10 min, then 40 cycles of 95°C for 15 sec and 60°C for 1 min.

Analysis of the TAC results:
The results were analyzed using the QuantStudio 12K Flex Software. Predefined cycle thresholds were set into the software template for each target. For a sample, the negative results of RNA or DNA targets were excluded from the analysis if the quantification cycle (C_q) for the corresponding extrinsic control, MS2 or PhHV, was ≥35. A target was considered negative if the C_q ≥35.
To compare the demographic information for the three age strata, Student’s t test was performed using either the pooled or Satterthwaite method depending on whether the variances were equal. Chi-square test or Fisher’s exact test was performed for categorical variables. The prevalence of pathogen target was calculated based on the number of positive results for the target ($C_q<35$). The prevalence of highly-diarrhea associated pathogen targets were calculated based on quantities where the point estimate of the odds ratio exceeded 2. The proportion of highly-diarrhea associated pathogen targets were calculated by the ratio of the two prevalences described above. The prevalence and proportion were calculated for the total study population as well as by age strata. Wilson Confidence Interval was used for prevalence or proportion. A negative binomial regression model was used to determine if age, gender, region, dysentery, acute diarrhea, or historical antibiotic use were risk factors for a high number of pathogens in the stool specimen. We defined dysentery based on the questionnaire as the presence of ‘mucus and blood’ or ‘blood’ or ‘watery and blood’. Acute diarrhea was defined as diarrhea in the previous 24 hours. A logistic regression model was used to examine whether the covariates described above plus the number of pathogens in the stool were indicative of having a highly diarrhea-associated pathogen. All statistical analyses were conducted in R version 4.0.3 (R Core Team, 2020).

### Results

#### Study population:

From August 2016 to October 2018, 7227 children and adolescents under 18 years of age were enrolled in the cholera surveillance study. 2746 participants met the inclusion criteria for the
pediatric substudy: 0-2 years old: 2234 participants; 3-5 years old: 277 participants; 6-17 years old: 235 participants. We randomly selected 114 (0-2 years), 112 (3-5 years) and 112 (6-17 years) children from each age group for the nested subgroup analysis (Table 1). We tested 338 specimens using the TAC,(5, 18) of which 337 produced valid results. One specimen (3-5 age group) was excluded due to poor sample quality. Of the remaining 337, three samples failed to produce valid results for RNA targets and their RNA targets were removed from the prevalence calculation and the pathogen count analysis. Three additional samples failed to produce valid results for Adenovirus, thus their Adenovirus targets were removed from the prevalence calculation, and they were removed from the pathogen count analysis.

A secondary aim of this study was to evaluate and adapt a low-cost filter paper preservation method for use with TAC methodologies. Extraction methods were evaluated using spiked stool specimens prior to processing the participant specimens to ensure the filter paper preserved DNA was extracted sufficiently for processing via the TAC (data not shown).

Positivity in the qualitative analysis is the presence of any pathogen with a cycle threshold (CT) of less than 35. Figure 1 uses this threshold to determine which pathogens are present as well as those most prevalent among participant stools. Derived variables are grouped according to defined primer targets (Supp Table 1). Stratifying by age groups, EAEC remains most prevalent, with Shigella and ETEC among the five top pathogens (Table 2).

The mean number of pathogens was 3.12 pathogens per person, with a range of 0 to 10 pathogens per person (Supp Figure 2). 93% of specimens had pathogens detectable in their stool, with nearly 75% having 2 or more pathogens in their stool. The negative binomial regression model demonstrated that the expected number of pathogens in the stool for those meeting the study definition of dysentery was 1.37 times of those that did not have dysentery (P<0.05). The
expected number of pathogens in the stool of those with non-acute diarrhea was 1.38 times of that of those with acute diarrhea (P<0.05) (Supp Table 2).

Applying the GEMS quantitative cut-offs for each pathogen, the prevalence of highly diarrhea associated pathogens did not differ significantly if compared to the total population or by age group (Table 3). 40% of participants had one or more highly associated pathogen present in their stool. Shigella was the most prevalent highly diarrhea-associated pathogen in the older adolescent age group (6-17) at 23% and the second most prevalent in the younger two age groups.

After adjusting for other covariates, the logistic regression demonstrated that the presence of one or more pathogen in the stool increased the odds of having a highly diarrhea-associated pathogen present by 45% (p<0.05). The odds of having a highly diarrhea-associated pathogen present are increased in males by 62% (p<0.05) and increased if from Douala, Littoral Region, by 130% (p<0.05) (Supp Table 3). The overall proportion of pathogens that are highly diarrhea associated demonstrates that of the 96 participants with Shigella in their stool, 56% present with Shigella that is highly diarrhea-associated (Table 3).

Of the 337 specimens, 96 individuals (28.5%) were positive at any quantity for Shigella/EIEC, characterized by the gene IpaH. Shigella/EIEC was the sixth most common pathogen seen in individuals with dysentery, with 37.8% (14/37) individuals with dysentery testing positive at any quantity. 29.5% (N=33) of participants > 5 years of age were Shigella positive. Shigella/EIEC was the most common pathogen associated with dysentery, accounting for 52.9% (9/17) of dysentery with an attributable cause, and 24.3% (9 in 37) of the total dysentery cases. Genes indicative of S. flexneri, including primer targets for any flexneri, were seen in 59.2% (32 in 54) of all shigella-attributable cases and 46.9% (45 in 96) of all Shigellae cases at any quantity.
(67%) were from Douala and 15 (33%) were from the Far North. *S. sonnei* was identified in
20.4% (11 in 54) of all shigella-attributable cases and 12.5% (12/96) of all *Shigellae* cases at any
quantity, the majority were from Douala (92%, 11 in 12). 85% (82/96) of participants positive
for *Shigellae* at any quantity did not present with dysentery, further, 83% (45/54) of the
attributable incidence of Shigella was non-dysenteric.

107 (32%) were positive at any quantity for ETEC, defined as positive for LT, STh, or STp. Of
these, 53 (49.5%) were positive for LT-only, 10 (9.3%) were positive for ST only and 44
(41.1%) were positive for LT and ST. Of those positive for ST, 11 (20.4%) were positive for
STp, 33 (61.1%) were positive for STh, and 10 (18.5) were positive for both STp and STh.
31.3% (N=35) of participants > 5 years of age were ETEC positive. Of which, all experienced
watery diarrhea. STh was present at attributable quantities in 10 samples. LT nor STp had
quantitative cut-offs available in the literature.

The distribution of ETEC toxin and 6 major CF types were analyzed. Overall, CFs were present
in 43.9% (47 in 107) of ETEC positive individuals. CS6 was the most common, followed by
CFA I, CS2, CS3, and CS1. We detected a CS in 6 patients from whom we did not detect LT or
ST (Supp Table 4).

Campylobacter was the second most common pathogen detected, present in 33.2% (n=112) of
specimens. 66 (58.9%) were positive for *C. jejuni* or *C. coli*. Despite high presence in the
qualitative analysis, campylobacter did not meet the cut-off to suggest disease attribution.

Campylobacter was present in 77 (68.8%) specimens from Far North, and 35 (31.2%) from
Douala, a statistically significant difference in prevalence. Figure 2 presents the prevalence of
the top 5 pathogens, stratified by region. Seasonality was evaluated but did not demonstrate a
significant difference by pathogen or region (data not shown).
Regardless of which study thresholds were used, comparing GEMS (0-5 age strata), and MAL-ED (0-2 age strata), the results remain similarly aligned (Figure 3). Per MAL-ED, shigella is the most prevalent highly diarrhea-associated pathogen.

Discussion

This study characterized pediatric diarrheal samples from multiple locations with high rates of diarrhea, where previously there was limited information on the circulating etiologic agents. This study demonstrated that through optimization of nucleic acid extraction from low-cost filter paper preserved specimens we can examine stool specimens for a range of enteric and serotype specific information of several key targets including Enterotoxigenic E. coli (ETEC) and Shigella.

We conducted both qualitative and quantitative analyses to assess the primary outcome confirming our primary hypothesis of high rates of ETEC and Shigella in the pediatric population. The top five pathogens demonstrate that Shigella, ETEC and Campylobacter are highly present among the participants. The most common pathogens by age group were similar (Supp Figures 1; Figure 2). The regional differences implore the need for further investigation. Importantly, the data demonstrate that the five most prevalent pathogens remain the same regardless of age. Research on the casual pathways contributing to environmental enteropathy, stunting and other long-term negative health outcomes among infants and young children in LMICs indicates that even asymptomatic colonization with these enteropathogens should be a public health concern since it may impact growth, cognitive development and long-term economic productivity(3, 19, 20) Additionally, we found high rates of Shigella and ETEC in age
groups older than five, confirming the importance of these two pathogens in among older children and adolescents. (21)

A secondary aim was to use qPCR to determine the quantity of the bacterial DNA present and to correlate degrees of clinical illness and case attribution. Since the parent study did not involve the enrollment of controls, we were unable to establish independent cut-offs and applied previously established cut-offs. (5) While neither GEMS nor MAL-ED had a site in Cameroon, both large studies included sites in sub-Saharan Africa focusing on pediatric populations. We applied the “highly diarrhoea-associated quantity” for more stringent cutoffs in our analysis. We were able to identify etiological pathogens for 132 (39.9%) of the samples tested. Further, using the filter paper preservation, we were successful at this rate in a population where more than 90% had taken antibiotics. 7 of 14 pathogens were attributable quantities, rotavirus was the most commonly present attributable disease, followed by Shigella/EIEC and ETEC. STh. 116 (87.8%) cases had one diarrhea-associated pathogen detected, 15 (11.4%) had two, and 1 (0.8%) had three. Campylobacter and EAEC were two of the most common pathogens found at any quantity, but no cases were attributable to either.

Limitations to our analysis include the lack of controls due to the nested design. Further, we utilized filter paper preserved specimens which contained 50 ul of dried stool specimen. To address these two limitations, we used the GEMS derived cutoffs and for simplicity, we applied the more stringent highly-diarrhea associated cut-off values for CQ rather. Utilizing the GEMS cut-offs to attribute disease with the broad inclusion of bacteria, viruses and parasites identified several interesting findings for which we hope to build future studies. The filter paper cards were preserved at RT; while we have not previously observed degradation using this method, the TAC used in this analysis looks at both RNA and DNA viruses. Additional research is needed to fully
understand the stability of filter paper preserved specimens stored at various temperatures. Our results do indicate similar finding to studies that used such methods. In conclusion, we were able to utilize low-cost filter paper preserved specimens to identify diarrheal pathogens prevalent in a pediatric population. We confirmed that ETEC and *Shigella* cause a large proportion of MSD illnesses which may be controlled with vaccines currently being developed. We applied stringent cutoffs to minimize the likelihood of false-attribution and improve applicability for vaccine trials. Ongoing trials into current Shigella and ETEC vaccines will need validation in additional endemic countries, and this study suggests that Shigella and ETEC endemic in the pediatric population in Cameroon. The findings from this study support the use filter paper preservation of stool specimens for PCR confirmation, and therefore, more accurate understanding of diarrheal disease burden in the pediatric population in Cameroon.
Funding

This work was supported by The Thrasher Research Fund Early Career Award [grant number 1870291656]; the Bill and Melinda Gates Foundation [grant number OPP1148763], which provided financial support through the John Hopkins University Delivery of Oral Cholera Vaccine Effectively (DOVE) project and the National Institute of Allergy and Infectious Disease [grant number R01AI123422].
Acknowledgements:

We are grateful to all the study participants at the many study sites. We thank the staff of the Kousseri Health District, Mada Health District, Makary Health District, Nylon and Goulfey Health District for their participation in this study. We also thank the Regional Delegate for Public Health, from the Far North both Dr. Rebecca Djao and Dr Fanne Mahamat, and from the Littoral Region Dr. Dissongo Jean II. We thank the Division of Health Operations Research, Cameroon Ministry of Public Health, for their support of this project. We thank Dr. Tahaniyat Lalani and Michele Tisdale from the Infectious Disease Clinical Research Program, Department of Preventive Medicine & Biostatistics, Uniformed Services University of the Health Sciences, Bethesda, USA and The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, USA, for their guidance on the TAC platform. Lastly, we thank Dr. A. Lou Bourgeois from PATH’s Center for Vaccine Innovation and Access for providing his expertise and editorial input from the conception of the sub-study through manuscript completion.


Table 1: Demographic characteristics

<table>
<thead>
<tr>
<th>Demographic characteristics: n (%)</th>
<th>All age (N = 337)</th>
<th>1 – 2 years (N = 114)</th>
<th>3 – 5 years (N = 111)</th>
<th>6 – 17 years (N = 112)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>175(52)</td>
<td>57(50)</td>
<td>50(45)</td>
<td>68(61)</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Douala</td>
<td>199(59)</td>
<td>59(52)</td>
<td>69(62)</td>
<td>71(63)</td>
</tr>
<tr>
<td>Far North</td>
<td>138(41)</td>
<td>55(48)</td>
<td>42(38)</td>
<td>41(37)</td>
</tr>
<tr>
<td>Dysentery</td>
<td>37(11)</td>
<td>17(15)</td>
<td>11(10)</td>
<td>9(8)</td>
</tr>
<tr>
<td>Acute diarrhea</td>
<td>292(87)</td>
<td>92(81)</td>
<td>103(93)</td>
<td>97(87)</td>
</tr>
<tr>
<td>Reported any antibiotics use</td>
<td>315(93)</td>
<td>110(96)</td>
<td>106(95)</td>
<td>99(88)</td>
</tr>
<tr>
<td>Number of pathogens: Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average number of pathogens</td>
<td>3.1(2.0)</td>
<td>3.5 (1.9)</td>
<td>2.8 (1.9)</td>
<td>3 (2.2)</td>
</tr>
<tr>
<td>Number of individuals with highly diarrhea-associated pathogen: n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of individuals with highly diarrhea-associated pathogen¹</td>
<td>132(40)</td>
<td>40(36)</td>
<td>45(41)</td>
<td>47(42)</td>
</tr>
</tbody>
</table>

¹ 6 sample failed to produce valid results for part of pathogens and were removed from this counting, leaving 311 samples in total.
Table 2: Prevalence of Pathogens (top five), prevalence of diarrhea-associated pathogens of all samples, stratified by age.

<table>
<thead>
<tr>
<th>Rank</th>
<th>All age (N = 337)</th>
<th>1 – 2 years (N = 114)</th>
<th>3 – 5 years (N = 111)</th>
<th>6 – 17 years (N = 112)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pathogens</td>
<td>Prevalence of pathogens (95% CI)</td>
<td>Prevalence of pathogens (95% CI)</td>
<td>Prevalence of pathogens (95% CI)</td>
</tr>
<tr>
<td>1</td>
<td>EAEC</td>
<td>0.45 (0.40,0.51)</td>
<td>0.57 (0.49,0.66)</td>
<td>0.40 (0.31,0.49)</td>
</tr>
<tr>
<td>2</td>
<td>Campylobacter</td>
<td>0.33 (0.28,0.38)</td>
<td>0.43 (0.34,0.52)</td>
<td>0.32 (0.24,0.41)</td>
</tr>
<tr>
<td>3</td>
<td>ETEC</td>
<td>0.32 (0.27,0.37)</td>
<td>0.41 (0.33,0.50)</td>
<td>0.28 (0.21,0.38)</td>
</tr>
<tr>
<td>4</td>
<td>EPEC</td>
<td>0.31 (0.26,0.36)</td>
<td>0.32 (0.25,0.42)</td>
<td>0.27 (0.20,0.36)</td>
</tr>
<tr>
<td>5</td>
<td>Shigella/EIEC</td>
<td>0.28 (0.24,0.34)</td>
<td>0.29 (0.21,0.38)</td>
<td>0.23 (0.17,0.32)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pathsogens</th>
<th>Prevalence of highly diarrhea-associated pathogens (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rotavirus 0.17 (0.13,0.21) Rotavirus 0.15 (0.09,0.23)</td>
</tr>
<tr>
<td>2</td>
<td>Shigella/EIEC 0.16 (0.12,0.20) Shigella/EIEC 0.13 (0.08,0.21)</td>
</tr>
<tr>
<td>3</td>
<td>V.cholerae 0.03 (0.02,0.05) Adenovirus.40.41 0.04 (0.01,0.09)</td>
</tr>
<tr>
<td>4</td>
<td>ETEC.STh 0.03 (0.02,0.05) V.cholerae 0.04 (0.01,0.09)</td>
</tr>
<tr>
<td>5</td>
<td>Adenovirus.40.41 0.02 (0.01,0.04) Salmonella 0.02 (0.00,0.06)</td>
</tr>
<tr>
<td>6</td>
<td>Salmonella 0.02 (0.01,0.04) ETEC.STh 0.02 (0.00,0.06)</td>
</tr>
<tr>
<td>7</td>
<td>Cryptosporidium 0.01 (0.01,0.03) Cryptosporidium 0.01 (0.01,0.05)</td>
</tr>
</tbody>
</table>
Table 3: Proportion of Pathogens that are highly diarrhea-associated (95% CI).

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Presence</th>
<th>diarrhea</th>
<th>All age</th>
<th>1 – 2 years</th>
<th>3 - 5 years</th>
<th>6 – 17 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotavirus</td>
<td>68</td>
<td>56</td>
<td>82% (72, 90)</td>
<td>85% (64, 95)</td>
<td>87% (71, 95)</td>
<td>71% (47, 87)</td>
</tr>
<tr>
<td>Shigella/EIEC</td>
<td>96</td>
<td>54</td>
<td>56% (46, 66)</td>
<td>45% (30, 62)</td>
<td>43% (27, 61)</td>
<td>79% (62, 89)</td>
</tr>
<tr>
<td>V.cholerae</td>
<td>15</td>
<td>11</td>
<td>73% (48, 89)</td>
<td>67% (30, 90)</td>
<td>0</td>
<td>88% (53, 99)</td>
</tr>
<tr>
<td>ETEC.STh</td>
<td>43</td>
<td>10</td>
<td>23% (13, 38)</td>
<td>13% (4, 38)</td>
<td>50% (25, 75)</td>
<td>13% (3, 36)</td>
</tr>
<tr>
<td>Adenovirus.40.41</td>
<td>72</td>
<td>7</td>
<td>10% (5, 19)</td>
<td>13% (5, 30)</td>
<td>5% (8, 33)</td>
<td>10% (3, 30)</td>
</tr>
<tr>
<td>Salmonella</td>
<td>10</td>
<td>7</td>
<td>70% (40, 89)</td>
<td>67% (21, 98)</td>
<td>50% (15, 85)</td>
<td>100% (44, 100)</td>
</tr>
<tr>
<td>Cryptosporidium</td>
<td>28</td>
<td>5</td>
<td>18% (8, 36)</td>
<td>7% (1, 18)</td>
<td>20% (1, 62)</td>
<td>88% (53, 99)</td>
</tr>
</tbody>
</table>
Figure 1: The prevalence of the top 10 pathogens

Figure 2: The prevalence of the top 5 pathogens, stratified by region

Figure 3: Alignment of our study results: A) in comparison to the prevalence of highly diarrhea-associated pathogens in GEMS and B) in comparison to the prevalence of highly diarrhea-associated pathogens in MAL-ED
Figure 1: The prevalence of the top 10 pathogens
Figure 2: The prevalence of the top 5 pathogens, stratified by region
Figure 3: Alignment of our study results: A) in comparison to the prevalence of highly diarrhea associated pathogens in GEMS and B) in comparison to the prevalence of highly diarrhea-associated pathogens in MAL-ED