Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED)


Summary

Background Most studies of the causes of diarrhoea in low-income and middle-income countries have looked at severe disease in people presenting for care, and there are few estimates of pathogen-specific diarrhoea burdens in the community.

Methods We undertook a birth cohort study with not only intensive community surveillance for diarrhoea but also routine collection of non-diarrhoeal stools from eight sites in South America, Africa, and Asia. We enrolled children within 17 days of birth, and diarrhoeal episodes (defined as maternal report of three or more loose stools in 24 h, or one loose stool with visible blood) were identified through twice-weekly home visits by fieldworkers over a follow-up period of 24 months. Non-diarrhoeal stool specimens were also collected for surveillance for months 1–12, 15, 18, 21, and 24. Stools were analysed for a broad range of enteropathogens using culture, enzyme immunoassay, and PCR. We used the adjusted attributable fraction (AF) to estimate pathogen-specific burdens of diarrhoea.

Findings Between Nov 26, 2009, and Feb 25, 2014, we tested 7318 diarrhoeal and 24 310 non-diarrhoeal stools collected from 2145 children aged 0–24 months. Pathogen detection was common in non-diarrhoeal stools but was higher with diarrhoea. Norovirus GII (AF 5·2%, 95% CI 3·0–7·1), rotavirus (4·8%, 4·5–5·0), Campylobacter spp (3·5%, 0·4–6·3), astrovirus (2·7%, 2·2–3·1), and Cryptosporidium spp (2·0%, 1·3–2·6) exhibited the highest attributable burdens of diarrhoea in the first year of life. The major pathogens associated with diarrhoea in the second year of life were Campylobacter spp (7·9%, 3·1–12·1), norovirus GII (5·4%, 2·1–7·8), rotavirus (4·9%, 4·4–5·2), astrovirus (4·2%, 3·5–4·7), and Shigella spp (4·0%, 3·6–4·3). Rotavirus had the highest AF for sites without rotavirus vaccination and the fifth highest AF for sites with the vaccination. There was substantial variation in pathogens according to geography, diarrhoea severity, and season. Bloody diarrhoea was primarily associated with Campylobacter spp and Shigella spp, fever and vomiting with rotavirus, and vomiting with norovirus GII.

Interpretation There was substantial heterogeneity in pathogen-specific burdens of diarrhoea, with important determinants including age, geography, season, rotavirus vaccine usage, and symptoms. These findings suggest that although single-pathogen strategies have an important role in the reduction of the burden of severe diarrhoeal disease, the effect of such interventions on total diarrhoeal incidence at the community level might be limited.

Funding Bill & Melinda Gates Foundation.

Copyright © Platts-Mills et al. Open Access article published under the terms of CC BY-NC-ND.

Introduction Infectious diarrhoea is the second most common cause of death in children under 5 years old in developing countries.1 Studies of the causes of diarrhoea in these settings have usually focused on children who present to health centres and, therefore, best describe pathogens associated with severe diarrhoea.2 However this approach captures only a small subset of diarrhoeal episodes which might show a different hierarchy of pathogens from that associated with mild or moderate episodes of diarrhoea.

Non-severe episodes in the community are of substantial public health importance because of their high prevalence and association with poor growth, impaired cognitive development, environmental enteropathy, and even mortality.1,4 Estimates of the pathogen-specific burdens of diarrhoea at the community level are, therefore, needed to prioritise interventions. Further, surveillance in the community allows for unbiased estimates of the associations between pathogens and distinct clinical syndromes.

The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) is a multisite birth cohort study at eight sites in South America, sub-Saharan Africa, and Asia.2 We aimed to estimate pathogen-specific burdens of diarrhoea in children aged 0–24 months at these MAL-ED study sites.

Lancet Glob Health 2015; 3: e564–75
Published Online
July 20, 2015
http://dx.doi.org/10.1016/S2214-109X(15)00151-5
See Comment page e520
*These authors contributed equally to this work
†Listed at end of paper
Division of Infectious Diseases and International Health, University of Virginia, Charlottesville, VA, USA
(A Havit MD, Ila F N Lima PhD, A M Lima MD)>
International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh
(T Ahmed MBBS, D Mondal MD, R Haque MD)>
Christian Medical College, Vellore, India
(S Baby MD, D Hariraju MSc, K G Kang MD)>
Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand
(B Rayamajhi BSc, L Bodhidatta MD, C J Mason MD)>
Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA
(P P Yori MPH, M Kosek MD)>
Asociación Benefica PRISMA, Iquitos, Peru
(M P Ollotegui BSc, P P Yori MPH, M Kosek MD)>
Aga Khan University, Karachi, Pakistan
(S Shakeri MBBS, S Qureshi MSc, F Kabir MSc, A Zadi MBBS)>
Zulfiqar Bhutta PhD)>
University of Venda, Thohoyandou, South Africa
(A Same MD, B Mufamadi B Tech, P Resso PhD)>
Haydom Lutheran Hospital, Haydom, Tanzania
(C Amour MSc, P Gratz BSc, E Mduma MPH)>
Foundation for the National...
Research in context

Evidence before this study
We searched PubMed for articles published in any language since 1990 using the terms “diarrhea/diarrhoea” and “etiology/aetiology” and “pediatric/paediatric OR infant*” and “case-control study OR cohort study.” We identified 482 publications, including 11 aetologic studies of diarrhoea which included testing for a broad range of enteropathogens. Of those, eight studied children with more severe diarrhoea presenting to health-care settings. The three remaining studies of community diarrhoea involved a single site.

Added value of this study
Our study provides multisite data on the causes of diarrhoea with longitudinal surveillance and interrogation of a broad range of pathogens, allowing unbiased estimates of pathogen-specific burdens of diarrhoea in the community as well as estimates for specific diarrhoeal syndromes. It documents the broad range of pathogens associated with diarrhoea of any severity, the heterogeneity of the main causes of diarrhoea in low-income and middle-income countries, and the diversity of pathogens associated with seasonal peaks. It also documents the effect of rotavirus vaccine.

Implications of all available evidence
These data suggest that the causes of community diarrhoea are diverse, and single pathogen interventions might not have a substantial impact on total diarrhoeal incidence across multiple populations.

Methods

Study design and participants
A detailed description of the MAL-ED study design is available elsewhere. We enrolled children from the community within 17 days of birth at eight study locations: Dhaka, Bangladesh; Fortaleza, Brazil; Vellore, India; Bhaktapur, Nepal; Loreto, Peru; Naushero Feroze, Pakistan; Venda, South Africa; and Haydom, Tanzania.

Inclusion criteria included: a mother aged 16 years or older; intention for the family to stay in the study area for at least 6 months from enrolment; that the child was from a singleton pregnancy and had no other siblings enroled in the study; and birthweight or enrolment weight greater than 1500g. We excluded children with diagnosed congenital disease or severe neonatal disease in the newborn.

Enrolment took place between November, 2009, and February, 2012. We aimed to enrol at least 200 children at every site, and we staggered enrolment to capture approximately equal number of births in each calendar month. Follow-up was for 24 months. Length, weight, and head circumference were measured every month, as described previously.

All sites received ethics approval from their respective governmental, local institutional, and collaborating institutional ethics review boards. Written informed consent was obtained from the parent or guardian of every child.

Sample and data collection
Non-diarrhoeal stool specimens were collected for surveillance for months 1–12, 15, 18, 21, and 24. Diarrhoeal episodes were collected from age 0–23 months and were identified at home visits made by fieldworkers twice a week. They were defined as maternal report of three or more loose stools in 24 h, or one loose stool with visible blood. Discrete episodes had at least 2 intervening days without diarrhoea. Diarrhoeal stool specimens had to be collected within 48 h of an episode. When a stool sample was collected between two episodes of diarrhoea that met criteria for collection, we assigned the sample to the episode closest to the time of collection.

A diarrhoea severity score was calculated for every episode using elements derived from the Vesikari score (table 1). Dehydration was defined as irritability that was difficult to console, increased thirst, loss of skin turgor, sunken eyes, or lethargy. Dysentery was defined as diarrhoea in which visible blood was reported by the child’s mother. Diarrhoea associated with fever was defined as diarrhoea with fieldworker-confirmed temperature greater than 37.5°C, and vomiting-associated diarrhoea required vomiting at any point during the episode of diarrhoea.

Diarrhoeal episodes of fewer than 7 days’ duration were classified as acute, 7–14 days as prolonged, and more than 14 days as persistent. Stools collected within 1 day of administration of a lactulose-mannitol test were excluded from analysis. Data on rotavirus vaccine administration and antibiotic use were recorded and children were referred to medical care for severe symptoms.

Stool testing
All stools were analysed in accordance with a standardised microbiology protocol, which was implemented at all
We used conventional stool culture to identify bacterial pathogens with the exception of Campylobacter spp.

Testing for diarrhoeagenic Escherichia coli was done by pooling five lactose-fermenting colonies for multiplex PCR to detect the toxin-encoding genes stx1, stx2, eae, bfpA, _iapH, aatA_, and _aaiC_, as well as those encoding heat-labile enterotoxin (LT) and heat-stable enterotoxin (ST).

Enzyme immunoassay was used for detection of Campylobacter spp, rotavirus, adenovirus, and astrovirus (ProSpecT, Remel, Lenexa, KS, USA) and _Entamoeba histolytica_, _Giardia_ spp, and _Cryptosporidium_ spp (TechLab, Blacksburg, VA, USA). Rotavirus detections were considered negative if obtained within 28 days of rotavirus vaccine administration (n=18).

We used PCR to test all diarrhoeal stool samples for norovirus. We also aimed to test all non-diarrhoeal stool samples from a randomly selected 10% subset of participants at each site.

If an additional specimen was available, we did use microscopy for identification of protozoa and helminths; however, microscopy was not required for complete testing, and microscopy results were not included for the analysis of infections for the three protozoal pathogens tested by enzyme immunoassay. If testing was incomplete, recollection was allowed within 48 h.

### Statistical analysis

Because pathogens were frequently detected in diarrhoeal and non-diarrhoeal stools, we used the adjusted attributable fraction (AF) to estimate pathogen-specific study sites and has been described in detail previously. We used conventional stool culture to identify bacterial pathogens with the exception of Campylobacter spp.

Testing for diarrhoeagenic Escherichia coli was done by pooling five lactose-fermenting colonies for multiplex PCR to detect the toxin-encoding genes _stx1_, _stx2_, _eae_, _bfpA_, _ipAH_, _aatA_, and _aaiC_, as well as those encoding heat-labile enterotoxin (LT) and heat-stable enterotoxin (ST).

Enzyme immunoassay was used for detection of Campylobacter spp, rotavirus, adenovirus, and astrovirus (ProSpecT, Remel, Lenexa, KS, USA) and _Entamoeba histolytica_, _Giardia_ spp, and _Cryptosporidium_ spp (TechLab, Blacksburg, VA, USA). Rotavirus detections were considered negative if obtained within 28 days of rotavirus vaccine administration (n=18).

We used PCR to test all diarrhoeal stool samples for norovirus. We also aimed to test all non-diarrhoeal stool samples from a randomly selected 10% subset of participants at each site.

If an additional specimen was available, we did use microscopy for identification of protozoa and helminths; however, microscopy was not required for complete testing, and microscopy results were not included for the analysis of infections for the three protozoal pathogens tested by enzyme immunoassay. If testing was incomplete, recollection was allowed within 48 h.

### Statistical analysis

Because pathogens were frequently detected in diarrhoeal and non-diarrhoeal stools, we used the adjusted attributable fraction (AF) to estimate pathogen-specific

<table>
<thead>
<tr>
<th>Pathogens detected in diarrhoeal and non-diarrhoeal stools, 0-11 months and 12-24 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAEc=enteroaggregative Escherichia coli; EIEc=enteroinvasive E. coli; aEPEC=atypical enteropathogenic E. coli; tEPEC=typical enteropathogenic E. coli; LT-TEC=LT-producing enterotoxigenic E. coli; ST-TEC=ST-producing enterotoxigenic E. coli; STEC=Shiga-toxin-producing E. coli. Pathogens present in less than 0.1% of stool samples are not shown.</td>
</tr>
</tbody>
</table>

---

**Figure 1:** Pathogens detected in diarrhoeal and non-diarrhoeal stools, 0-11 months and 12-24 months

**Table 2:** MAL-ED cohort descriptive statistics and completeness of surveillance and testing
burdens of diarrhoea, a measurement that incorporates the prevalence of detection in diarrhoeal stools and the strength of association with diarrhoea.

To analyse the strength of association between diarrhoea and detection of individual pathogens, we used generalised estimating equations (GEEs) to fit a binary logistic regression model for each site and age group to account for non-independence of stool testing within each participant. All models were adjusted for age (in days), sex, and site. We included all detected pathogens from diarrhoeal stools for each age and site, and we assumed an independent working correlation matrix. We then calculated AFs using the point estimate of the odds ratios derived from the multivariate GEEs with 95% CIs estimated using the Delta method.

We determined the pathogen-specific attributable incidence for each calendar month by first calculating the AF using the prevalence of each pathogen in diarrhoea for each calendar month and then multiplying by the number of episodes of diarrhoea during that month. To mitigate the detection of convalescent excretion of pathogens, we excluded from analysis non-diarrhoeal stools collected more than 48 h but fewer than 7 days before or after a diarrhoeal episode. The effect of prolonged excretion of enteric pathogens on AF estimates was evaluated by further restricting non-diarrhoeal specimens to those collected at least 28 days before and after any diarrhoeal episode. Pathogen-specific AFs were calculated for the subset of diarrhoeal episodes that met study definitions of acute, prolonged, persistent, mild, moderate, severe, or dysenteric diarrhoea, or diarrhoea associated with fever or with vomiting.

Figure 2: Pathogen detection and diarrhoeal episodes per child, 0–24 months
Dots show mean values with standard error bars.

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Dhaka, Bangladesh</th>
<th>Vellore, India</th>
<th>Bhaktapur, Nepal</th>
<th>Naushero Feroze, Pakistan</th>
<th>Venda, South Africa</th>
<th>Haydom, Tanzania</th>
<th>Fortaleza, Brazil</th>
<th>Loreto, Peru</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 0–11 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norovirus GII</td>
<td>–</td>
<td>–</td>
<td>8·4% (5·7–9·7)</td>
<td>–</td>
<td>8·2% (5·1–12·9)</td>
<td>–</td>
<td>5·1% (3·0–7·1)</td>
<td>5·2%</td>
<td></td>
</tr>
<tr>
<td>Rotavirus</td>
<td>9·6% (8·8–10·1)</td>
<td>6·0% (5·5–6·3)</td>
<td>6·6% (5·9–6·9)</td>
<td>3·2% (2·6–3·5)</td>
<td>9·5% (7·6–10·5)</td>
<td>–</td>
<td>4·8% (4·5–5·0)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Campylobacter spp</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>16·9% (9·0–21·6)</td>
<td>–</td>
<td>–</td>
<td>5·6% (4·7–6·5)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Astrovirus</td>
<td>2·0% (0·3–3·2)</td>
<td>4·2% (3·2–4·9)</td>
<td>–</td>
<td>2·2% (0·9–3·1)</td>
<td>–</td>
<td>–</td>
<td>2·7% (2·2–3·2)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Coxsackievirus</td>
<td>–</td>
<td>–</td>
<td>3·6% (1·9–4·8)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2·0% (1·3–2·6)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>ST-ETEC</td>
<td>4·7% (3·3–5·8)</td>
<td>1·7% (0·6–2·3)</td>
<td>2·0% (1·0–2·5)</td>
<td>1·2% (0·1–1·8)</td>
<td>3·2% (0·9–4·2)</td>
<td>–</td>
<td>1·9% (1·5–2·2)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Adenovirus</td>
<td>–</td>
<td>2·7% (0·9–3·7)</td>
<td>2·3% (0·7–3·2)</td>
<td>1·1% (0·0–1·9)</td>
<td>–</td>
<td>–</td>
<td>1·5% (0·2–2·3)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>tEPEC</td>
<td>2·2% (0·0–4·1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1·3% (0·7–1·9)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>LT-ETEC</td>
<td>2·0% (0·2–3·3)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1·3% (0·6–1·9)</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

(Table 3 continues on next page)
To analyse the association between pathogen detection and diarrhoea severity, GEEs were used to fit an ordinal regression model which was specified identically to the logistic regression models used for the analysis of diarrhoea association. For all analyses, we constructed models both with and without norovirus because of the differential testing of non-diarrhoeal specimens for this pathogen. The results we report for pathogens other than norovirus, as well as for all analyses involving aggregated pathogen testing, were derived from models that excluded norovirus. We used R version 3.0.3 (Foundation for Statistical Computing, Vienna, Austria) for all statistical analyses, with the geepack package within this program used for GEE analysis.29

Role of the funding source
The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
Between Nov 3, 2009, and Feb 29, 2012, we enrolled 2145 children (range 233–314 per site). The size of the

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Dhaka, Bangladesh</th>
<th>Vellore, India</th>
<th>Bhaktapur, Nepal</th>
<th>Naushero Feroz, Pakistan</th>
<th>Venda, South Africa*</th>
<th>Haydom, Tanzania</th>
<th>Fortaleza, Brazil*</th>
<th>Loreto, Peru*</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shigella spp</td>
<td>0.7% (0.3–0.7)</td>
<td>0.9% (0.6–1.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4%</td>
</tr>
<tr>
<td>Norovirus GI</td>
<td></td>
<td>11.2% (6.4–11.9)</td>
<td></td>
<td></td>
<td></td>
<td>19.2% (2.2–26.1)</td>
<td></td>
<td></td>
<td>11.7%</td>
</tr>
<tr>
<td>Campylobacter spp</td>
<td>8.8% (2.0–13.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.9%</td>
</tr>
<tr>
<td>Rotavirus</td>
<td>6.0% (4.8–6.6)</td>
<td>8.7% (8.7–8.7)</td>
<td>2.2% (0.7–2.9)</td>
<td></td>
<td></td>
<td>14.3% (11.5–15.1)</td>
<td>4.3% (1.7–4.9)</td>
<td>2.9%</td>
<td></td>
</tr>
<tr>
<td>Astrovirus</td>
<td>2.6% (1.7–3.7)</td>
<td>4.6% (3.2–5.3)</td>
<td></td>
<td></td>
<td></td>
<td>9.7% (8.1–12.3)</td>
<td>4.7% (3.2–5.0)</td>
<td>7.4%</td>
<td></td>
</tr>
<tr>
<td>Shigella spp</td>
<td>1.5% (0.3–2.0)</td>
<td>9.4% (8.7–9.8)</td>
<td>6.8% (5.8–7.4)</td>
<td>5.1% (3.8–5.9)</td>
<td></td>
<td>3.7% (2.3–3.8)</td>
<td></td>
<td>2.1%</td>
<td></td>
</tr>
<tr>
<td>ST-ETEC</td>
<td>8.0% (5.6–9.7)</td>
<td>5.4% (3.6–6.3)</td>
<td>4.6% (2.2–5.9)</td>
<td></td>
<td></td>
<td>9.1% (2.7–10.9)</td>
<td></td>
<td>2.0%</td>
<td></td>
</tr>
<tr>
<td>Cryptosporidium spp</td>
<td>2.5% (0.4–6.9)</td>
<td>3.2% (0.5–7.7)</td>
<td>5.5% (1.4–4.1)</td>
<td>13.0% (0.9–14.4)</td>
<td></td>
<td>13.0% (6.9–14.7)</td>
<td></td>
<td>3.8%</td>
<td></td>
</tr>
<tr>
<td>LT-ETEC</td>
<td>2.4% (0.1–3.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.1% (0.0–22.8)</td>
<td></td>
<td>1.2%</td>
<td></td>
</tr>
<tr>
<td>Enterovirus</td>
<td>3.6% (0.9–5.0)</td>
<td>3.9% (2.1–4.8)</td>
<td></td>
<td></td>
<td></td>
<td>3.8% (1.1–4.7)</td>
<td></td>
<td>0.9%</td>
<td></td>
</tr>
<tr>
<td>Entamoeba histolytica</td>
<td>0.7% (0.0–1.6)</td>
<td>0.8% (0.2–1.2)</td>
<td></td>
<td></td>
<td></td>
<td>0.7% (0.0–1.2)</td>
<td></td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td>Salmonella</td>
<td>0.7% (0.0–1.2)</td>
<td>0.5% (0.0–5.0)</td>
<td></td>
<td></td>
<td></td>
<td>0.5% (0.0–5.0)</td>
<td></td>
<td>0.3%</td>
<td></td>
</tr>
<tr>
<td>Norovirus GI</td>
<td>0.7% (1.0–1.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0% (0.1–1.2)</td>
<td></td>
<td>1.0%</td>
<td></td>
</tr>
<tr>
<td>Aeromonas</td>
<td>0.7% (0.0–1.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2% (0.0–2.0)</td>
<td></td>
<td>0.2%</td>
<td></td>
</tr>
<tr>
<td>Plesiomonas</td>
<td>0.7% (0.0–1.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0%</td>
<td></td>
</tr>
<tr>
<td>ST-EC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2% (0.0–2.0)</td>
<td></td>
<td>0.2%</td>
<td></td>
</tr>
</tbody>
</table>

*Monovalent rotavirus vaccine was introduced to the national immunisation programme at these sites before the study began.

Table 3: Adjusted attributable fraction of diarrhoea for individual pathogens in the first and second year of life
Two fieldworker visits per week were sufficient to collect most diarrhoeal stools within 48 h (79.5% overall; site range 33.0–96.1%). Collection rates were higher for longer episodes (75.5% for acute episodes and 99.3% for prolonged or persistent episodes).

A broad range of pathogens was detected, with 22 pathogens in the first year of life and 25 in the second year of life (we have not included pathogens in analysis if they were present in only very few samples—ie, less than 0.1% of all stools). For certain pathogens, detection in non-diarrhoeal stools approached, and in some cases exceeded, that noted for diarrhoeal stools (figure 1).

Enteropathogen infection began soon after birth and was common at all sites; however, the intensity varied between sites, ranging from an average of about 0.5 pathogens detected per stool by the end of the first year of life (South Africa) to almost two pathogens per stool (Pakistan; figure 2). Both the incidence of diarrhoea and the number of pathogens detected per stool increased markedly during the first year of life. At least one pathogen was detected in 76.9% (n= 15767) of diarrhoeal stools and 64.9% (15767) of non-diarrhoeal stools, and two or more pathogens were identified in 41.0% (2999) and 29.0% (7046) of stools, respectively. The number of pathogens detected was higher in diarrhoeal stools than non-diarrhoeal stools at most time points (appendix).

The presence of pathogens was associated with diarrhoea, in that each additional pathogen increased the odds of diarrhoea (odds ratio (OR) 1.20 per pathogen detection, p<0.0001). Antibiotics were administered for 4696 (46%) diarrhoeal episodes captured by surveillance with a range between sites of 20 (11%, Brazil) to 1922 (59%, Pakistan).

Overall, 19.1%, (95% CI 16.2–21.8) and 33.1% (29.0–36.7) of diarrhoeal episodes in the first and second year of life, respectively, could be attributed to pathogens. Attributable fractions did not change appreciably when the more restrictive definition of non-diarrhoeal specimens was applied, suggesting that estimates were not biased by convalescent excretion (appendix), nor did they change after controlling for child nutritional status (height-for-age Z score).

Across all sites and episodes, the highest AFs were seen for norovirus GII, rotavirus, Campylobacter spp, astrovirus, and Cryptosporidium spp in the first year of life and Campylobacter spp, norovirus GII, rotavirus, astrovirus, and Shigella spp in the second year of life (table 3 and appendix).

There was substantial heterogeneity between sites in the individual pathogen most often associated with diarrhoea, with the highest burden of diarrhoea attributed to four unique pathogens in the first year of life (Campylobacter spp, Cryptosporidium spp, norovirus GII,

![Figure 3: Prevalence and adjusted attributable fraction of diarrhoea for 3-month intervals, age 0–24 months](https://example.com/figure3.png)

**Legend:**
- EAEC=enteroaggregative Escherichia coli; EIEC=enteroinvasive E coli; aEPEC=atypical enteropathogenic E coli; tEPEC=typical enteropathogenic E coli; LT-ETEC=LT-producing enterootoxigenic E coli; ST-ETEC=ST-producing enterootoxigenic E coli; STEC=Shiga-toxin producing E coli. Data are attributable fractions (95% CI). For each organism, the first data point represents age 0–2 months, the second represents age 3–5 months, then 6–8 months, 9–11 months, 12–14 months, 15–17 months, 18–20 months, and 21–24 months.
Figure 3. First infections were more strongly associated with Entamoeba histolytica, Campylobacter spp, and ST-producing enterotoxigenic Escherichia coli (E coli) than with other pathogens; however, this did not alter AF estimates (data not shown). Helminthic infections were not associated with diarrhea for any age group, site, or diarrhoeal syndrome.

We next examined whether clinical characteristics or seasonality could aid prediction of the cause of diarrhea. Total attribution to pathogens for episodes associated with dysentery, dehydration, or admission to hospital was

Three frequently detected pathogens, namely enter-aggregative E coli, Giardia spp, and atypical enteropathic E coli, were not statistically significantly associated with diarrhea for any age group, site, or diarrhoeal syndrome. Age-related patterns were seen for several pathogens: astrovirus, norovirus GII, and rotavirus diarrhea burdens peaked during age 6–12 months, whereas Cryptosporidium spp, Shigella spp, Campylobacter spp, and ST-producing enterotoxigenic E coli continued to increase through the second year of life (figure 3). First infections were more strongly associated with diarrhea than were subsequent infections for most pathogens; however, this did not alter AF estimates (data not shown). Helminthic infections were not associated with diarrhea for any age group, site, or diarrhoeal syndrome.

<table>
<thead>
<tr>
<th>Age 0–11 months</th>
<th>Diarrhoeal stools</th>
<th>Prolonged</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Blood in stool</th>
<th>Associated fever</th>
<th>Associated vomiting</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
<td>(%) of diarrhoea</td>
</tr>
<tr>
<td>Diarrhoeal stools</td>
<td>3249 (79·9%)</td>
<td>1031 (24·1%)</td>
<td>1696 (39·6%)</td>
<td>1762 (41·2%)</td>
<td>820 (19·2%)</td>
<td>108 (4·6%)</td>
<td>204 (4·8%)</td>
<td>1235 (28·9%)</td>
<td>4280 (95·9%)</td>
</tr>
<tr>
<td>Norovirus GII</td>
<td>5·5% (5·4–5·6)</td>
<td>4·4% (4·3–4·5)</td>
<td>5·2% (5·1–5·3)</td>
<td>4·7% (4·6–4·8)</td>
<td>5·5% (5·4–5·6)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>7·5% (7·4–7·6)</td>
<td>5·2% (5·1–5·3)</td>
</tr>
<tr>
<td>Rotavirus</td>
<td>5·6% (5·5–5·7)</td>
<td>2·2% (2·1–2·3)</td>
<td>2·0% (1·9–2·1)</td>
<td>5·2% (5·1–5·3)</td>
<td>9·8% (9·7–9·9)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>7·2% (7·1–7·3)</td>
<td>4·8% (4·7–4·9)</td>
</tr>
<tr>
<td>Campylobacter spp</td>
<td>4·4% (4·3–4·5)</td>
<td>8·1% (8·0–8·2)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>23·7% (23·6–23·8)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>3·5% (3·4–3·6)</td>
</tr>
<tr>
<td>Astrovirus</td>
<td>2·9% (2·8–3·0)</td>
<td>1·8% (1·7–1·9)</td>
<td>2·7% (2·6–2·8)</td>
<td>2·3% (2·2–2·4)</td>
<td>3·4% (3·3–3·5)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>3·9% (3·8–3·9)</td>
<td>2·7% (2·6–2·8)</td>
</tr>
<tr>
<td>Cryptosporidium spp</td>
<td>1·7% (1·6–1·8)</td>
<td>3·0% (2·9–3·1)</td>
<td>1·2% (1·1–1·3)</td>
<td>2·3% (2·2–2·4)</td>
<td>3·2% (3·1–3·3)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>2·4% (2·3–2·5)</td>
<td>2·0% (1·9–2·1)</td>
</tr>
<tr>
<td>ST-ETEC</td>
<td>4·2% (4·1–4·3)</td>
<td>1·8% (1·7–1·9)</td>
<td>1·8% (1·7–1·8)</td>
<td>2·2% (2·1–2·3)</td>
<td>1·4% (1·3–1·5)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>1·9% (1·8–2·0)</td>
<td>1·9% (1·8–2·0)</td>
</tr>
<tr>
<td>Adenovirus</td>
<td>1·4% (1·3–1·5)</td>
<td>2·1% (2·0–2·2)</td>
<td>1·0% (0·9–1·1)</td>
<td>1·6% (1·5–1·7)</td>
<td>3·2% (3·1–3·3)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>3·0% (2·9–3·1)</td>
<td>1·6% (1·5–1·7)</td>
</tr>
<tr>
<td>tEPEC</td>
<td>1·2% (1·1–1·3)</td>
<td>1·6% (1·5–1·6)</td>
<td>1·4% (1·3–1·5)</td>
<td>– (–)</td>
<td>2·2% (2·1–2·3)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>1·5% (1·4–1·6)</td>
<td>1·3% (1·2–1·4)</td>
</tr>
<tr>
<td>LT-ETEC</td>
<td>0·9% (0·8–1·0)</td>
<td>2·6% (2·5–2·7)</td>
<td>1·0% (0·9–1·1)</td>
<td>1·1% (1·0–1·2)</td>
<td>2·3% (2·2–2·4)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>1·8% (1·7–1·9)</td>
<td>1·3% (1·2–1·4)</td>
</tr>
<tr>
<td>Shigella spp</td>
<td>0·3% (0·2–0·4)</td>
<td>0·6% (0·5–0·7)</td>
<td>0·3% (0·2–0·4)</td>
<td>0·4% (0·3–0·5)</td>
<td>– (–)</td>
<td>3·4% (3·3–3·5)</td>
<td>12% (11–13)</td>
<td>– (–)</td>
<td>0·4% (0·3–0·5)</td>
</tr>
<tr>
<td>EIEC</td>
<td>– (–)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>– (–)</td>
<td>8% (7–9)</td>
<td>17% (16–18)</td>
<td>– (–)</td>
<td>– (–)</td>
</tr>
</tbody>
</table>

(Table 4 continues on next page)
Articles

Persistent diarrhoea represented 4.9% and 1.8% of episodes during the first and second year of life, respectively, and was associated with LT-producing enterotoxigenic *E coli*, *astrovirus*, *Cryptosporidium* spp, *ST-producing entero-toxigenic E coli*, *ST-ETEC*, *ST-producing enterotoxigenic E coli*, and *Shigella* spp in the first year of life and *Shigella*, and *astrovirus* in India and norovirus GII, *ST-ETEC,* and *Cryptosporidium* spp were associated with a persistent diarrhoea.

The association between the attributable incidence of specific pathogens and seasonal diarrhoeal incidence varied between sites (figure 4). For many sites, peak diarrhoea incidence coincided with the peak attributable incidence for some pathogens—for example *Cryptosporidium* spp, *ST-producing enterotoxigenic E coli*, *Shigella* spp, and *astrovirus* in India and norovirus GII, *ST-producing enterotoxigenic E coli*, and *Shigella* spp in Nepal. Rotavirus incidence was strongly seasonal, and during peak season it dominated all-cause diarrhoea incidence in India, Nepal, Pakistan, and Tanzania. There was little association between rotavirus incidence and seasonality at the three sites where rotavirus vaccine had been introduced.

### Discussion

In this multicountry community-based cohort study, pathogen-specific burdens of diarrhoea varied substantially between sites. Although rotavirus diarrhoea burden was substantially decreased at sites where rotavirus vaccine had been introduced, it occupied the overall highest burden of disease at the five sites that do not have vaccination. Nevertheless, it was associated with the highest burden of diarrhoea at only three sites in the first year of life and at none in the second year.

### Table 4: Adjusted attributable fraction of diarrhoea associated with specific diarrhoeal syndromes in the first and second year of life for individual pathogens

<table>
<thead>
<tr>
<th>Age 12–24 months</th>
<th>Acute (&lt;7 days)</th>
<th>Prolonged (&gt;7 days)</th>
<th>Mild (score 1–3)</th>
<th>Moderate (score 4–6)</th>
<th>Severe (score &gt;6)</th>
<th>Blood in stool</th>
<th>Associated fever</th>
<th>Associated vomiting</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diarrhoeal stools</strong> (%) of diarrhoea</td>
<td>2568 (84.5%)</td>
<td>470 (15.5%)</td>
<td>155 (51.5%)</td>
<td>1104 (36.3%)</td>
<td>381 (12.5%)</td>
<td>159 (5.2%)</td>
<td>142 (4.7%)</td>
<td>698 (23.0%)</td>
<td>3038 (100.0%)</td>
</tr>
<tr>
<td>Campylobacter spp</td>
<td>8.9%</td>
<td>9.7%</td>
<td>8.3%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.9%</td>
</tr>
<tr>
<td>Norovirus GII (18-76)</td>
<td>5.1%</td>
<td>6.9%</td>
<td>4.5%</td>
<td>6.2%</td>
<td>6.9%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.9%</td>
</tr>
<tr>
<td>Rotavirus (47-56)</td>
<td>5.2%</td>
<td>3.8%</td>
<td>5.1%</td>
<td>7.9%</td>
<td>-</td>
<td>4.9%</td>
<td>10.1%</td>
<td>4.9%</td>
<td></td>
</tr>
<tr>
<td>Astrovirus (4-5)</td>
<td>4.5%</td>
<td>2.3%</td>
<td>4.1%</td>
<td>4.7%</td>
<td>2.8%</td>
<td>-</td>
<td>5.4%</td>
<td>4.5%</td>
<td>4.2%</td>
</tr>
<tr>
<td>Shigella spp (1-3-7)</td>
<td>3.4%</td>
<td>7.0%</td>
<td>2.7%</td>
<td>5.1%</td>
<td>5.7%</td>
<td>17.2%</td>
<td>6.9%</td>
<td>3.1%</td>
<td>4.0%</td>
</tr>
<tr>
<td>ST-ETEC (2-8-43)</td>
<td>3.6%</td>
<td>5.5%</td>
<td>3.4%</td>
<td>3.9%</td>
<td>5.8%</td>
<td>-</td>
<td>3.6%</td>
<td>5.5%</td>
<td>3.9%</td>
</tr>
<tr>
<td>Cryptosporidium spp (2-2-43)</td>
<td>3.4%</td>
<td>6.1%</td>
<td>3.0%</td>
<td>4.5%</td>
<td>3.2%</td>
<td>-</td>
<td>-</td>
<td>3.8%</td>
<td>3.8%</td>
</tr>
<tr>
<td>LT-ETEC (1-3)</td>
<td>1.3%</td>
<td>-</td>
<td>1.5%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.0%</td>
<td>2.2%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Adenovirus (0-2-19)</td>
<td>1.0%</td>
<td>-</td>
<td>0.8%</td>
<td>1.9%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.9%</td>
<td>0.9%</td>
</tr>
<tr>
<td>EIEC (0-1-3)</td>
<td>0.8%</td>
<td>-</td>
<td>0.9%</td>
<td>1.2%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.9%</td>
<td>0.8%</td>
</tr>
<tr>
<td>E histolytica (0-3-0)</td>
<td>0.7%</td>
<td>-</td>
<td>1.1%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.7%</td>
</tr>
<tr>
<td>Salmonella (0-1-5)</td>
<td>0.4%</td>
<td>-</td>
<td>0.4%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.3%</td>
<td>0.3%</td>
</tr>
<tr>
<td>Aeromonas spp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.3%</td>
<td>1.2%</td>
<td>0.9%</td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td>Plesiomonas spp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.6%</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

EIEC = enteroinvasive *Escherichia coli*; TPEC = typical enteropathogenic *E coli*; LT-ETEC = LT-producing enterotoxigenic *E coli*; ST-ETEC = ST-producing enterotoxigenic *E coli*; STEC = Shiga-toxin producing *E coli*. Data are n or attributable fractions (95% CI). The subset of pathogens assayed that were significant in at least one syndrome or age group are shown in descending order of average attributable fraction for study-defined diarrhoea. For cells with a dash, the pathogen was either not detected or was not statistically significantly associated with diarrhoea.

(Continued from previous page)
Astrovirus—pathogens that have rarely been examined in such a large study with modern diagnostic tools,2 or have not been noted as important in case-control studies.2,3,30 The number and diversity of pathogens associated with community diarrhoea suggests that single pathogen interventions, apart from rotavirus vaccination, might not have an effect on the incidence of diarrhoeal episodes across populations.

This multisite longitudinal study design allowed us to uncover an unbiased picture of the association between specific pathogens and specific clinical features, including duration, severity, dysentery, febrile illness, and vomiting.

---

**Figure 4: Association between individual pathogens and seasonal diarrhoeal incidence**

TEPEC=typical enteropathogenic Escherichia coli; LT-ETEC=LT-producing enterotoxigenic E.coli; ST-ETEC=ST-producing enterotoxigenic E.coli. Primary y-axis shows percent of total attributable incidence of diarrhoea for individual pathogens; secondary y-axis (and dotted line) shows annual diarrhoeal incidence by calendar month. *Monovalent rotavirus vaccine was introduced to the national immunisation programme before the study began.
Dysentery in the first year of life was predominantly associated with *Campylobacter* spp; however, *Campylobacter*-associated diarrhoea was, otherwise, mild when assessed with a severity score that did not include the presence of blood. By contrast, dysentery associated with *Shigella* spp was often severe and of surprisingly long duration. Rotavirus and norovirus GII were associated with vomiting.

*Campylobacter* spp were the most frequently detected pathogens and had the highest burden of diarrhoea in Brazil, Peru, and South Africa in the first year of life. Such a high burden of *Campylobacter* spp early in the first year of life, often with dysentery, has been observed in some studies but not others. This pathogen did not show strong seasonal trends. We have previously shown that culture substantially underdetects *Campylobacter* whereas EIA broadly detects *Campylobacter* spp, including species other than C. jejuni and C. coli. We expect most of the episodes associated with *Campylobacter* spp to be caused by *C. jejuni* or *C. coli*, but culture identification was only done on a subset of stools in our study and further work is needed.

We documented a substantial burden of diarrhoea associated with norovirus GII infection at the sites in Nepal, South Africa, Tanzania, and Peru, as well as in the overall analysis. As in developed countries, norovirus GII appeared to be a significant contributor to overall diarrhoeal incidence at several sites. There has been substantial variation in previous estimates of the global burden of norovirus, in part because detection of norovirus GII is often high in asymptomatic control participants matched for age, community, and season.

Astrovirus is known to be a common cause of sporadic diarrhoea that is less severe than that associated with rotavirus, and astrovirus often exists as a co-infection. Our study shows the global importance of astrovirus diarrhoea, with a substantial burden of disease in most sites. Adenovirus had a low overall attributable fraction, but, when present, was associated with diarrhoea classified as “severe” by an adapted Vesikari score. We used a pan-adenovirus ELISA without typing for the major gastrointestinal subtypes 40/41; however, we would not expect the AF for adenovirus to increase significantly given its low prevalence. Helmint infections were rare in this study, except for *Ascaris* in the second year of life, and were not associated with diarrhoea.

This study also documents frequent detection of a wide range of pathogens, including *Campylobacter* spp, enteroaggregative *E. coli*, norovirus, *Giardia*, LT-producing enterotoxigenic *E. coli*, and typical and atypical enteropathogenic *E. coli* in routinely collected non-diarrhoeal stools. Whether the presence of these pathogens is associated with more insidious phenotypes such as poor growth, impaired cognitive development, environmental enteropathy, or impaired mucosal immunity is unclear and further study is warranted in this area.

Our study has some limitations. In light of the variation between sites in diarrhoeal incidence, the study was not powered to identify all associations between pathogens and diarrhoea at individual sites. Furthermore, because short episodes of diarrhoea are more difficult to capture with community-based surveillance than are longer periods of diarrhoea, especially in rural settings, burden estimates might be biased against pathogens associated with a short duration of symptoms. Additionally, we used a modified severity score that only partly recapitulates a score derived from rotavirus studies and may not be generalisable. Therefore, we also looked at the subset of diarrhoea associated with dysentery, dehydration, or hospital admission in addition to looking at specific diarrhoeal syndromes. Finally, the diagnostic approach used a diverse set of detection methods with differing performance characteristics. It is possible, for example, that culture for bacterial pathogens is insensitive and was affected by the frequent use of antibiotics for diarrhoea in these settings, such that the use of culture for detection may have resulted in underestimates of bacterial presence. Molecular testing, in particular quantification of pathogen load and quantitative analysis, could revise estimates of the burden of diarrhoea for these organisms.

The longitudinal nature of this study allowed us to look at causes of diarrhoea in ways that are not possible with other study designs, including use of unbiased estimates of causes of diarrhoea at the community level and evaluation of assumptions about appropriate control specimens. Detection of pathogens in non-diarrhoeal stool samples might represent convalescent excretion of certain pathogens rather than true asymptomatic infection, in which case we may underestimate the burden of diarrhoea associated with these organisms. Malnourished children may be particularly likely to have prolonged excretion of enteropathogens. However, controlling for nutritional status did not appreciably alter AF estimates.

This study documents a diverse range of pathogens associated with community diarrhoea in children in low-income and middle-income countries, which contrasts with the smaller set of pathogens associated with severe diarrhoea. The hierarchy of pathogen-specific diarrhoea varied between sites and high rates of enteropathogens were detected in non-diarrhoeal samples.

Consistent with previous studies, a high burden of childhood diarrhoea was attributed to rotavirus, ST-ETEC, *Shigella* spp, and *Cryptosporidium* spp. However, our results suggest that *Campylobacter* spp, norovirus GII, and astrovirus also contribute substantially to the burden of diarrhoea in children.
Additional MAL-ED Network Investigators


Declaration of interests

We declare no competing interests.

Acknowledgments

The Etiology, Risk Factors and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) is a collaborative project of the Malnutrition as an Enteric Infectious Disease with Long-term Effects (MAL-ED) Newtork Investigators.

Bill & Melinda Gates Foundation, the Foundation for the NIH, and the National Institutes of Health, Fogarty International Center. The authors thank the staff and participants of the MAL-ED Network Project for their important contributions.

References


