

# Protection From Natural Immunity Against Enteric Infections and Etiology-Specific Diarrhea in a Longitudinal Birth Cohort

Elizabeth T. Rogawski McQuade,<sup>1,2</sup> Jie Liu,<sup>2</sup> Gagandeep Kang,<sup>3,6</sup> Margaret N. Kosek,<sup>2,4</sup> Aldo A. M. Lima,<sup>5</sup> Pascal O. Bessong,<sup>6</sup> Amidou Samie,<sup>6</sup> Rashidul Haque,<sup>7</sup> Estomih R. Mduma,<sup>8</sup> Sanjaya Shrestha,<sup>9</sup> Jose Paulo Leite,<sup>10</sup> Ladaporn Bodhidatta,<sup>11</sup> Najeeha Iqbal,<sup>12</sup> Nicola Page,<sup>13</sup> Ireen Kiwelu,<sup>14</sup> Zulfiqar Bhutta,<sup>12</sup> Tahmeed Ahmed,<sup>7,9</sup> Eric R. Houpt,<sup>2</sup> and James A. Platts-Mills<sup>2</sup>

<sup>1</sup>Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, USA, <sup>2</sup>Division of Infectious Diseases & International Health, University of Virginia, Charlottesville, Virginia, USA, <sup>3</sup>Christian Medical College, Vellore, India, <sup>4</sup>Asociación Benéfica PRISMA, Iquitos, Peru, <sup>5</sup>Federal University of Ceara, Fortaleza, Brazil, <sup>6</sup>University of Venda, Thohoyandou, South Africa, <sup>7</sup>International Centre for Diarrheal Disease Research, Dhaka, Bangladesh, <sup>8</sup>Haydom Global Health Research Centre, Haydom, Tanzania, <sup>9</sup>Walter Reed/AFRIMS Research Unit, Kathmandu, Nepal, <sup>10</sup>Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, Brazil, <sup>11</sup>Armed Forces Research Institute of Medical Sciences (AFRIMS), Bangkok, Thailand, <sup>12</sup>Aga Khan University, Karachi, Pakistan, <sup>13</sup>National Institute for Communicable Diseases, Johannesburg, South Africa, <sup>14</sup>Kilimanjaro Clinical Research Institute, Moshi, Tanzania

(See the Editorial Commentary by Lopman and Baker, on pages 1764–7.)

**Background.** The degree of protection conferred by natural immunity is unknown for many enteropathogens, but it is important to support the development of enteric vaccines.

**Methods.** We used the Andersen–Gill extension of the Cox model to estimate the effects of previous infections on the incidence of subsequent subclinical infections and diarrhea in children under 2 using quantitative molecular diagnostics in the MAL-ED cohort. We used cross-pathogen negative control associations to correct bias due to confounding by unmeasured heterogeneity of exposure and susceptibility.

**Results.** Prior rotavirus infection was associated with a 50% lower hazard (calibrated hazard ratio [cHR], 0.50; 95% confidence interval [CI], 0.41–0.62) of subsequent rotavirus diarrhea. Strong protection was evident against *Cryptosporidium* diarrhea (cHR, 0.32; 95% CI, 0.20–0.51). There was also protection due to prior infections for norovirus GII (cHR against diarrhea, 0.67; 95% CI, 0.49–0.91), astrovirus (cHR, 0.62; 95% CI, 0.48–0.81), and *Shigella* (cHR, 0.79; 95% CI, 0.65–0.95). Minimal protection was observed for other bacteria, adenovirus 40/41, and sapovirus.

**Conclusions.** Natural immunity was generally stronger for the enteric viruses than bacteria, potentially due to less antigenic diversity. Vaccines against major causes of diarrhea may be feasible but likely need to be more immunogenic than natural infection.

**Keywords.** bias analysis; diarrhea; enteric infections; natural immunity; negative control.

Vaccines are currently under development for several leading causes of diarrhea among children in low-resource settings, including *Shigella* [1], enterotoxigenic *Escherichia coli* [2], norovirus [3], and *Campylobacter* [4]. The development of immunity after natural infection is an important guide towards a successful vaccine. For rotavirus, observational analyses of natural immunity in low-resource settings found levels of protection that were comparable to that of vaccine efficacy [5–7], suggesting that natural immunity was a good predictor of vaccine performance. Assessment of natural immunity has informed cholera vaccine development, and levels of natural immunity

are frequently used as a benchmark against which to judge vaccine candidates [8, 9]. However, for many of the highest burden diarrhea etiologies, the degree of protection conferred by natural immunity is unknown, and immunologic surrogates of protection are imperfect. As large-scale investments in new vaccines are being made, better understanding of the magnitude, if any, of natural immunity to enteric pathogens is needed to inform expectations for vaccine effectiveness.

Estimates of natural immunity are confounded by the fact that children who are infected with an enteric pathogen may be more likely to be infected again due to high exposure risk and/or greater host susceptibility compared with other children. For example, children with a water source that is contaminated with *Cryptosporidium* will be more likely than other children to be infected again, even if they acquire some immunity after their first infection. Because this bias is expected to be in the opposite direction of a protective effect of natural immunity, confounding may completely mask evidence of protection or identify prior infection as a risk factor for subsequent infection.

Although observed factors such as sociodemographics, environmental characteristics, and markers of malnutrition may be able to explain some of this heterogeneity, these variables

Received 21 October 2019; editorial decision 12 December 2019; accepted 22 January 2020; published online January 27, 2020.

Presented in part: Vaccines for Enteric Diseases Annual Meeting, Lausanne, Switzerland, October 17, 2019.

Correspondence: Elizabeth T. Rogawski McQuade, PhD, MSPH, Department of Public Health Sciences, University of Virginia, P.O. Box 801379, Charlottesville, Virginia 22908-1379 (etr5m@virginia.edu).

The Journal of Infectious Diseases® 2020;222:1858–68

© The Author(s) 2020. Published by Oxford University Press for the Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. DOI: 10.1093/infdis/jiaa031

have been only modestly associated with acquisition of specific enteric infections among children in low-resource settings [10–14], suggesting that it is difficult to predict exposure and susceptibility based on readily observed characteristics. Unmeasured factors, such as innate immunity (major histocompatibility complex variation), are likely also important to characterize individual susceptibility, but they are difficult to account for in this setting. Alternatively, exposure to other enteropathogens can be used as negative controls because prior exposure to other pathogens would not be expected to elicit immunity, but it may be a good proxy for exposure and susceptibility due to common transmission pathways and host-related risk factors.

In this study, we estimated the protective effects of natural infection by common enteric pathogens against subsequent subclinical infection and etiology-specific diarrhea identified by quantitative molecular diagnostics in a longitudinal birth cohort. We used cross-pathogen estimates of protection as negative controls to estimate the magnitude of confounding bias due to unmeasured heterogeneity of exposure and susceptibility and calibrate effect estimates to account for this systematic error.

## METHODS

The study design and methods of the MAL-ED study have been previously described [15]. In brief, children were enrolled within 17 days of birth from November 2009 to February 2012 at 8 sites: Dhaka, Bangladesh; Fortaleza, Brazil; Vellore, India; Bhaktapur, Nepal; Naushahro Feroze, Pakistan; Loreto, Peru; Venda, South Africa; and Haydom, Tanzania. Each site obtained ethical approval from their institutions, and written informed consent was obtained from participants. Children were excluded if their mother was <16 years of age, their family intended to move from the study area, they were from a multiple pregnancy, their birthweight was  $\leq 1500$  grams, or they were diagnosed with congenital or severe neonatal disease. Surveillance was conducted for diarrhea, defined as maternal report of  $\geq 3$  loose stools in 24 hours or 1 stool with visible blood, at twice-weekly home visits until 2 years of age. Stool samples were collected monthly and during diarrhea episodes.

Stool samples from the subset of children with complete follow-up to 2 years of age were tested for 29 enteropathogens (Supplementary Table S1) by quantitative polymerase chain reaction (qPCR) using custom-designed TaqMan Array Cards (Thermo Fisher Scientific, Carlsbad, CA). Details of the assays and quality control have been previously described [16, 17]. Pathogen-attributable diarrhea episodes were defined using adjusted attributable fractions (AFes) for each episode to account for subclinical infections [16, 18]. In brief, we used the pathogen quantity, age, and sex-specific odds ratios (ORs) for diarrhea to estimate AFes for each diarrhea episode as follows:  $1 - 1/\text{OR}$ . We defined pathogen-attributable episodes when the pathogen quantity-derived AFe was .5 or higher (ie, majority

attribution), as previously described [19]. In a sensitivity analysis, we excluded diarrhea episodes in which more than 1 etiology was identified. Severe pathogen-attributable diarrhea episodes were defined by a severity score greater than 6, derived from components of the Vesikari score [20].

Because low-quantity detection of pathogen nucleic acid by qPCR may not indicate an established infection, we defined infections that could confer natural immunity as any detection at a quantity corresponding to qPCR cycle threshold (Cq) value  $\leq 30$ . For pathogens in which lower quantities were associated with diarrhea (where AFe  $\geq .5$ ), the quantity associated with diarrhea was used (rotavirus, Cq  $\leq 32.638$ ; *Shigella*, Cq  $\leq 30.507$ ; adenovirus 40/41, Cq  $\leq 30.424$ ; and norovirus GII, Cq  $\leq 30.357$ ). In sensitivity analyses, we (1) used the more sensitive analytical cutoff of Cq  $< 35$  to define infections and (2) considered only attributable diarrhea episodes as able to confer natural immunity, a more specific definition. We further limited to new infections that occurred at least 21 days after a previous detection. In sensitivity analyses, we defined new infections after 14 and 31 days.

## Data Analysis

To estimate protection due to natural immunity, we estimated the effects of prior infections on the hazard of subsequent subclinical infections, diarrhea, and severe diarrhea episodes due to the same pathogen. We assessed natural immunity to the 10 most common causes of diarrhea in MAL-ED: rotavirus, adenovirus 40/41, norovirus GII, sapovirus, astrovirus, *Shigella*, *Campylobacter jejuni*/*Campylobacter coli*, heat-stable enterotoxigenic *E coli* (ST-EPEC), typical enteropathogenic *E coli* (tEPEC), and *Cryptosporidium* [18]. We also assessed protection against subclinical infection for enteric pathogens with prevalence  $\geq .5\%$ : norovirus GI, enteroaggregative *E coli*, atypical EPEC, heat-labile enterotoxigenic *E coli* (LT-EPEC), *Aeromonas*, *Helicobacter pylori*, *Plesiomonas*, STEC, *Salmonella*, *Giardia*, *Enterocytozoon bieneusi*, and *Cyclospora*.

To model these effects, we used the Andersen and Gill extension of the Cox model for recurrent events [21] with a counting process formulation. Each risk period was defined by birth or age at 21 days after a prior infection to age at subsequent outcome, and each child contributed multiple risk periods from birth to 2 years of age (Supplementary Figure S1). Therefore, baseline hazards by age were estimated non-parametrically, and hazards of subsequent outcomes were conditioned on age. We estimated protection due to natural immunity as follows:  $(1 - \text{hazard ratio [HR]}) \times 100$ , in which the HR compared children who had experienced 1 or 2+ prior infections to children who experienced no prior infections. Robust variance accounted for correlation between risk periods within each child [22]. Models were adjusted for site and prespecified risk factors for enteric infections identified in previous analyses of MAL-ED [10–14]: sex, socioeconomic status (WAMI index [23]), enrollment weight-for-age z-score, maternal height, maternal education, crowding in the

home, and percentage of days exclusively breastfed (from birth to the current risk period up to 6 months of age). In a sensitivity analysis, we stratified effects by age.

### Bias Calibration

We used cross-pathogen estimates of protection, that is, associations between each enteric infection outcome and prior exposure to a different pathogen (eg, the protection due to prior *Campylobacter* infection on subsequent *Shigella* diarrhea) as negative control associations. We considered pathogens of the same type (bacteria, viruses, and parasites) as negative controls for each of the enteric infections. Because there were only 4 parasites, both bacteria and parasites were included as negative controls for *Cryptosporidium*. In sensitivity analyses, we included all pathogens as negative controls and compared calibrated estimates to estimates additionally adjusted for prior exposure to other pathogens.

We calibrated estimates with the negative controls using methods previously described [24, 25] with the EmpiricalCalibration package in R. After estimating the negative control associations using the models above, we used maximum likelihood to generate a systematic error model for each pathogen outcome that fit a Gaussian distribution to the negative control estimates and accounted for the sampling error of each estimate. Assuming the systematic error does not change as a function of the true effect size, we then estimated a calibrated distribution for the effect of interest that incorporated both random error and the systematic error model estimated above. We computed calibrated effect estimates and confidence intervals (CIs) as the .5, .025, and .975 percentiles of the corresponding cumulative distribution function [25].

## RESULTS

Among 1715 children with follow-up to 2 years and molecular testing of stool samples, there were a total of 52 382 detections of the top 10 causes of diarrhea (Table 1). Half of these detections ( $n = 24\ 520$ , 46.8%) were at quantities of  $Cq \leq 30$  (or greater than the diarrhea-associated quantity) and 21 days distant from a prior infection. Among these infections, the children experienced 3526 attributable diarrhea episodes and 539 severe attributable diarrhea episodes. For each pathogen, more than half of children had at least 1 infection, except for rotavirus (48.9% of children). *Shigella* ( $n = 719$  episodes; 29.6% of children with 1+ episodes) and rotavirus ( $n = 552$ ; 26.0%) were the most common causes of diarrhea (Table 1).

Covariate-adjusted estimates of protection against subsequent subclinical infection and attributable diarrhea due to prior infection (subclinical or diarrhea) (Figure 1A) were nearly equivalent to the corresponding unadjusted estimates (Supplementary Table S2). However, cross-pathogen negative control associations generated systematic bias distributions that were above the null for almost all pathogen outcomes,

suggesting that the adjusted estimates were biased upwards and underestimated natural immunity (Supplementary Figure S2). The magnitude of estimated bias (Table 2) was generally larger for the diarrhea outcomes than the infection outcomes, and it was larger for estimates assessing 2+ prior infections compared with 1 prior infection. *Shigella* demonstrated the largest magnitude of bias. Hazard ratios between *Shigella* diarrhea and 2+ prior infections by negative control pathogens were systematically 42% higher than the expected null association (mean adjusted HR for systematic error, 1.42; 95% CI, 1.34–1.51). The viruses, rotavirus, astrovirus, sapovirus, and adenovirus 40/41 had similar bias distributions, with negative control associations an average of 18% higher than the null for the diarrhea outcomes. In contrast, there was no systematic error identified for norovirus GII or *C jejuni/C coli* (Table 2).

After calibrating the adjusted estimates, estimates of protection generally increased (Figure 1B, Supplementary Table S2). One prior rotavirus infection was associated with an adjusted 26% lower hazard (calibrated HR [cHR], .74; 95% CI, .61–.91) of subsequent rotavirus infection and 47% lower hazard (cHR, .53; 95% CI, .42–.66) of subsequent rotavirus diarrhea (Figure 1A, Supplementary Table S2). Two or more prior rotavirus infections were associated with similar protection against infection but larger protection against diarrhea (57%; cHR, .43; 95% CI, .31–.61). There was slightly less protection for norovirus GII (cHR against diarrhea for 1+ prior infections, 0.67; 95% CI, 0.49–0.91) and astrovirus (cHR, 0.62; 95% CI, 0.48) (Supplementary Table S3).

The strongest protection was observed for *Cryptosporidium* diarrhea, and it was larger after multiple prior infections compared with 1 prior infection (cHR for 1 prior infection .035, 95% CI 0.21–0.59 vs cHR for 2+ prior infections 0.25, 95% CI 0.11–0.53). The calibrated estimates for *Shigella* diarrhea indicated modest protection after 1 prior infection (15%; cHR, .085; 95% CI, 0.69–1.04) and substantial protection after 2+ prior infections (33%; cHR, 0.67; 95% CI, 0.53–0.84). There was no evidence of protection after infections with the other bacterial pathogens (*C jejuni/C coli*, ETEC, tEPEC) and limited protection for adenovirus 40/41 and sapovirus (Figure 1B, Supplementary Table S2). Site-specific estimates of protection were generally consistent but were highly imprecise (Supplementary Table S4).

Estimates of protection against severe diarrhea were larger than those for episodes of any severity (Table 3). There was strong protection against severe diarrhea due to rotavirus (69%; cHR, 0.31; 95% CI, 0.20–0.48), astrovirus (66%; cHR, 0.34; 95% CI, 0.16–0.74), and *Cryptosporidium* (84%; cHR, 0.16; 95% CI, 0.02–1.72) after 1+ prior infections. Slightly less protection against severe *Shigella* diarrhea (43%, cHR, 0.57; 95% CI, 0.33–0.97) was also observed.

There was no evidence of protection against bacteria or parasites that were infrequently associated with diarrhea (Table 4). In contrast, prior infection was strongly associated

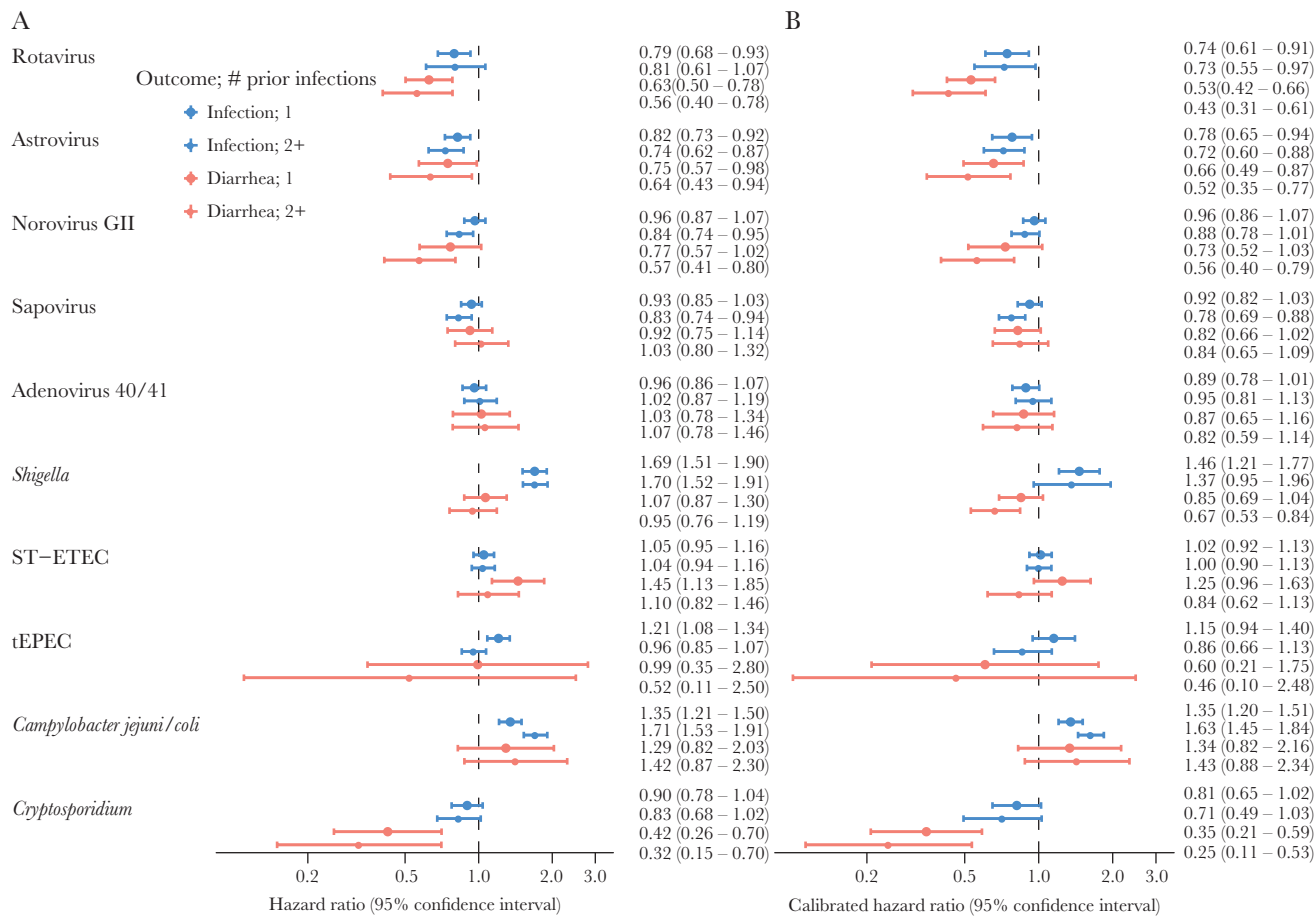
**Table 1. Number of Etiology-Specific Infections, Attributable Diarrhea Episodes, and Severe Attributable Diarrhea Episodes Among 1715 Children in the MAL-ED Cohort**

Number (%)	Rotavirus	Astrovirus	Norovirus GII	Sapovirus	Adenovirus 40/41	Shigella	ST-EPEC	tEPEC	Campylobacter jejuni/ Campylobacter coli	Cryptosporidium
Detections (Cq <35)	2239	5037	5526	6094	6112	4744	6314	5197	7734	3385
Infections (Cq ≤30 <sup>a</sup> )										
Total	1307 (58.4)	1886 (37.4)	2951 (53.4)	3043 (49.9)	2443 (40.0)	2901 (61.2)	3343 (62.9)	3041 (58.5)	3860 (49.9)	1625 (48.0)
New <sup>b</sup>	1217 (93.1)	1759 (93.3)	2717 (92.1)	2801 (92.0)	2264 (92.7)	2648 (91.3)	3216 (96.2)	2849 (93.7)	3596 (93.2)	1453 (89.4)
Mean age (days; SD)	339 (188.0)	388 (191.5)	340 (174.9)	419 (172.7)	398 (190.5)	508 (159.2)	435 (178.5)	381 (183.6)	391 (178.2)	467 (164.0)
Children with 1	568 (33.1)	566 (33.0)	511 (29.8)	517 (30.1)	570 (33.2)	422 (24.6)	419 (24.4)	416 (24.3)	341 (19.9)	603 (35.2)
Children with 2+	271 (15.8)	482 (28.1)	799 (46.6)	833 (48.6)	591 (34.5)	719 (41.9)	825 (48.1)	835 (48.7)	850 (49.6)	347 (20.2)
Attributable Diarrhea Episodes										
Total	566	308	296	542	411	755	452	31	162	119
New <sup>b</sup>	552 (97.5)	302 (98.1)	293 (99.0)	526 (97.0)	390 (94.9)	719 (95.2)	446 (98.7)	31 (100)	156 (96.3)	111 (93.3)
Mean age (days; SD)	348 (170.1)	396 (177.7)	377 (160.5)	428 (150.2)	357 (169.3)	507 (152.0)	442 (162.9)	223 (144.3)	338 (160.1)	478 (170.1)
Children with 1	353 (20.6)	224 (13.1)	218 (12.7)	276 (16.1)	157 (9.2)	347 (20.2)	248 (14.5)	27 (1.6)	99 (5.8)	93 (5.4)
Children with 2+	93 (5.4)	37 (2.2)	36 (2.1)	112 (6.5)	90 (5.2)	160 (9.3)	82 (4.8)	2 (0.1)	25 (1.5)	9 (0.5)
Severe Attributable Diarrhea Episodes										
Total	156	46	40	67	61	90	62	8	12	16
New <sup>b</sup>	152 (97.4)	46 (100)	40 (100)	65 (97.0)	56 (91.8)	87 (96.7)	61 (98.4)	8 (100)	10 (83.3)	14 (87.5)
Mean age (days; SD)	302 (144.4)	306 (149.5)	389 (150.9)	413 (150.6)	313 (143.2)	459 (165.3)	386 (167.8)	204 (136.7)	364 (112.5)	384 (175.7)
Children with 1	137 (8.0)	46 (2.7)	38 (2.2)	52 (3.0)	49 (2.9)	77 (4.5)	59 (3.4)	8 (0.5)	10 (0.6)	14 (0.8)
Children with 2+	7 (0.4)	0	1 (0.1)	6 (0.4)	3 (0.2)	5 (0.3)	1 (0.1)	0	0	0

Abbreviations: Cq, quantitative polymerase chain reaction cycle threshold; SD, standard deviation.

<sup>a</sup>For pathogens in which lower quantities were associated with diarrhea (where adjusted attributable fraction ≥5), the quantity associated with diarrhea was used to define infections (rotavirus, Cq ≤32.638; Shigella, Cq ≤30.507; adenovirus 40/41, Cq ≤30.424; and norovirus GII, Cq ≤30.357).

<sup>b</sup>Infections and episodes were considered new and included in the analysis if they occurred at least 21 days distant from a previous detection.



**Figure 1.** Estimates of protection against subclinical infection and attributable diarrhea due to 1 or 2 or more prior infections (subclinical or diarrhea) from the same pathogen in the MAL-ED cohort. (A) Hazard ratios adjusted for site, socioeconomic status, sex, enrollment weight-for-age Z-score, maternal education, maternal height, crowding, and exclusive breastfeeding in first 6 months; (B) hazard ratios adjusted for the same variables and calibrated based on negative control estimates. Abbreviations: ST-EPEC, heat-stable enterotoxigenic *E coli*; tEPEC, typical enteropathogenic *E coli*.

with subsequent infection for many, most strikingly for *Giardia*, *Salmonella*, *H pylori*, and *Cyclospora*. Estimates of natural protection against infection for norovirus GI were consistent with those for norovirus GII.

### Sensitivity Analyses

Greater protection in the second year of life compared with the first year of life was observed for adenovirus 40/41 and *Cryptosporidium* infection and tEPEC diarrhea. In contrast, prior sapovirus infection was protective against subsequent diarrhea in the first year but not the second (Supplementary Table S5).

Estimates of protection against rotavirus diarrhea were similar in the sites that had (Brazil, Peru, and South Africa) and had not introduced rotavirus vaccine. Protection against rotavirus infection was slightly stronger in sites that had introduced vaccine (cHR, 0.65; 95% CI, 0.42–1.00 for 1+ prior infection) compared with sites that had not (cHR, 0.76; 95% CI, 0.63–0.92).

Defining infections at a quantity cutoff of Cq <35 instead of Cq ≤30 resulted in inconsistent shifts in the estimates (Supplementary Table S6). For example, estimates of natural protection against diarrhea for rotavirus and astrovirus were closer to the null; estimates for norovirus GII and sapovirus were further from the null. Prior infections with *Cryptosporidium* were strongly predictive of subsequent infections, which may indicate that low-quantity detections are more likely to identify persistent infections rather than new infections (Supplementary Table S6).

Modification of the minimum duration between new infections (14 and 31 days instead of 21) resulted in minor changes to the estimates (Supplementary Tables S7 and S8). Defining prior exposure by prior attributable diarrhea instead of prior infection generally resulted in smaller estimates of protection, suggesting potential misclassification of immunity acquired in the absence of diarrheal symptoms (Supplementary Table S9). The exclusion of diarrhea episodes with multiple attributable pathogens resulted in

**Table 2. Estimates of the Systematic Bias Distribution From Negative Control Cross-Pathogen Associations Between Pathogen Outcomes and Prior Exposure to Other Pathogens in the Same Group**

Pathogen	Outcome	No. of Previous Infections	Mean Bias <sup>a</sup>	Bias Standard Deviation	Hazard Ratio <sup>b</sup> (95% CI)
Rotavirus	Infection	1	0.06	0.07	1.07 (0.93–1.22)
		2+	0.10	0.03	1.10 (1.04–1.18)
	Diarrhea	1	0.17	0.02	1.18 (1.13–1.24)
		2+	0.26	0.05	1.30 (1.19–1.42)
Astrovirus	Infection	1	0.05	0.07	1.05 (0.91–1.21)
		2+	0.02	0.05	1.02 (0.93–1.12)
	Diarrhea	1	0.13	0.04	1.14 (1.06–1.22)
		2+	0.21	0.04	1.23 (1.14–1.33)
Norovirus GII	Infection	1	0.01	0.02	1.01 (0.98–1.04)
		2+	–0.05	0.02	0.95 (0.91–0.99)
	Diarrhea	1	0.05	0.10	1.05 (0.87–1.27)
		2+	0.02	0.04	1.02 (0.95–1.09)
Sapovirus	Infection	1	0.02	0.03	1.02 (0.96–1.08)
		2+	0.07	0.02	1.07 (1.03–1.11)
	Diarrhea	1	0.11	0.02	1.12 (1.07–1.17)
		2+	0.20	0.03	1.22 (1.14–1.31)
Adenovirus 40/41	Infection	1	0.08	0.03	1.08 (1.02–1.15)
		2+	0.07	0.03	1.07 (1.00–1.14)
	Diarrhea	1	0.17	0.05	1.18 (1.07–1.30)
		2+	0.26	0.05	1.30 (1.18–1.43)
<i>Shigella</i>	Infection	1	0.15	0.08	1.16 (0.99–1.35)
		2+	0.22	0.17	1.25 (0.89–1.75)
	Diarrhea	1	0.23	0.02	1.26 (1.20–1.32)
		2+	0.35	0.03	1.42 (1.34–1.51)
ST-ETEC	Infection	1	0.03	0.02	1.03 (0.99–1.07)
		2+	0.04	0.02	1.04 (1.00–1.08)
	Diarrhea	1	0.15	0.05	1.16 (1.05–1.28)
		2+	0.27	0.05	1.31 (1.20–1.43)
tEPEC	Infection	1	0.05	0.09	1.05 (0.89–1.24)
		2+	0.10	0.13	1.11 (0.87–1.42)
	Diarrhea	1	0.49	0.12	1.64 (1.30–2.06)
		2+	0.12	0.31	1.13 (0.61–2.07)
<i>Campylobacter jejuni/Campylobacter coli</i>	Infection	1	0.00	0.02	1.00 (0.96–1.04)
		2+	0.05	0.02	1.05 (1.00–1.10)
	Diarrhea	1	–0.03	0.09	0.97 (0.82–1.14)
		2+	–0.01	0.04	0.99 (0.92–1.07)
<i>Cryptosporidium</i>	Infection	1	0.10	0.09	1.10 (0.93–1.31)
		2+	0.15	0.15	1.17 (0.86–1.58)
	Diarrhea	1	0.20	0.06	1.22 (1.08–1.37)
		2+	0.28	0.05	1.32 (1.21–1.45)

Abbreviations: CI, confidence interval; ST-ETEC, heat-stable enterotoxigenic *E coli*; tEPEC, typical enteropathogenic *E coli*.

<sup>a</sup>Mean bias is estimated on the natural log scale.

<sup>b</sup>Hazard ratio =  $e^{(\text{mean bias})}$ .

slightly stronger estimates of protection for the enteric viruses (Supplementary Table S10). Including all pathogens as negative controls resulted in smaller estimates of bias, especially for the viruses, with calibrated estimates generally closer to the covariate-adjusted estimates (Supplementary Table S11). Finally, adjusting for prior exposure to other pathogens did not completely correct the bias identified by the negative control calibration approach (Supplementary Table S12).

## DISCUSSION

We estimated strong protection due to prior infection against rotavirus, *Cryptosporidium*, and astrovirus diarrhea, which suggests that vaccine development for the latter 2 may be relatively feasible. Norovirus GII and *Shigella* also exhibited protection against diarrhea, but more strongly after multiple infections, which may reflect low potency or heterotypic protection. Less protection was observed for sapovirus and adenovirus 40/41. Estimates of protection against infection were universally

**Table 3. Estimates of Protection Against Severe Attributable Diarrhea Episodes Due to 1 or More Prior Infections From the Same Pathogen in the MAL-ED Cohort**

Pathogen	No. of Previous Infections	N	N (%) With Subsequent Severe Diarrhea	Unadjusted <sup>a</sup> HR (95% CI)	Adjusted <sup>b</sup> HR (95% CI)	Calibrated <sup>c</sup> HR (95% CI)
Rotavirus	0	1715	121 (7.1)	1.	1.	1.
	1+	839	29 (3.5)	0.37 (0.24–0.58)	0.37 (0.24–0.58)	0.31 (0.20–0.48)
Astrovirus	0	1715	35 (2.0)	1.	1.	1.
	1+	1048	11 (1.0)	0.40 (0.19–0.85)	0.39 (0.18–0.84)	0.34 (0.16–0.74)
Norovirus GII	0	1715	15 (0.9)	1.	1.	1.
	1+	1310	24 (1.8)	0.70 (0.31–1.55)	0.76 (0.34–1.70)	0.73 (0.32–1.65)
Sapovirus	0	1715	27 (1.6)	1.	1.	1.
	1+	1350	34 (2.5)	1.12 (0.64–1.97)	1.11 (0.62–1.96)	0.97 (0.55–1.72)
Adenovirus 40/41	0	1715	29 (1.7)	1.	1.	1.
	1+	1161	26 (2.2)	0.91 (0.54–1.51)	0.89 (0.51–1.54)	0.74 (0.42–1.29)
<i>Shigella</i>	0	1715	62 (3.6)	1.	1.	1.
	1+	1141	24 (2.1)	0.78 (0.45–1.34)	0.73 (0.43–1.25)	0.57 (0.33–0.97)
ST-EPEC	0	1715	28 (1.6)	1.	1.	1.
	1+	1244	33 (2.7)	1.01 (0.53–1.89)	0.97 (0.52–1.83)	0.82 (0.43–1.56)
tEPEC	0	1715	8 (0.5)	—	—	—
	1+	1251	0	—	—	—
<i>Campylobacter jejuni/Campylobacter coli</i>	0	1715	3 (0.2)	1.	1.	1.
	1+	1191	7 (0.6)	1.82 (0.49–6.70)	1.71 (0.46–6.33)	1.81 (0.49–6.70)
<i>Cryptosporidium</i>	0	1715	13 (0.8)	1.	1.	1.
	1+	950	1 (0.1)	0.21 (0.02–2.15)	0.20 (0.02–2.11)	0.16 (0.02–1.72)

Abbreviations: CI, confidence interval; HR, hazard ratio; ST-EPEC, heat-stable enterotoxigenic *E coli*; tEPEC, typical enteropathogenic *E coli*.

<sup>a</sup>Estimates adjusted for site.

<sup>b</sup>Estimates adjusted for site, socioeconomic status, sex, enrollment weight-for-age z-score, maternal education, maternal height, crowding, and exclusive breastfeeding in first 6 months.

<sup>c</sup>Estimates adjusted for the same variables above and calibrated based on negative control estimates for the associations between 1 or more previous infections and attributable diarrhea of any severity.

smaller in magnitude, whereas estimates against severe diarrhea were larger, compared with diarrhea of any severity.

Observed levels of natural protection were generally consistent with previous literature. Estimates for rotavirus were similar to those in a previous study from India (40%–80% protection against diarrhea) [5], and they were smaller in magnitude to estimates from Mexico [6] and Guinea-Bissau [7] (70%–90%). These estimates were also consistent with estimates of rotavirus vaccine efficacy in low-resource settings [26], which supports the use of the models for other pathogens. Levels of natural protection against norovirus GII were consistent with previous analyses of MAL-ED using a subset of samples (25%–30% protection against diarrhea) [27] and smaller than estimates from Peru (50%–80%) [28]. In contrast, no natural immunity to norovirus was observed in a study in Ecuador [29]. Estimates for astrovirus were larger than those from previous analyses of MAL-ED using the enzyme immunoassay in a subset of samples [30] and those from a study in rural Egypt [31]. Protection for *Cryptosporidium* was larger against diarrhea (70% vs 25%) and smaller against subclinical infection (20% vs 50+%) compared with a study in India [32]. The evidence for natural immunity to *Cryptosporidium* is supported by protection observed among adults with pre-existing serum antibodies [33] and delayed cryptosporidiosis among children with higher levels of anti-*Cryptosporidium* fecal immunoglobulin A (IgA) [34].

Because many of the pathogens included in this analysis are immunologically heterogeneous, these results should be interpreted as estimates of “functional protection” that reflect both homotypic and heterotypic immunity at the population-level. Different degrees of pathogen heterogeneity may explain variations in levels of protection. *Shigella* is antigenically diverse with 4 species and more than 50 serotypes [35], such that poor cross-protection may explain the relatively limited observed natural protection. A previous analysis of infection-derived immunity to *Shigella* in Chile found 14% protection overall but more than 70% serotype-specific protection [36].

The lack of natural protection observed for sapovirus, ST-EPEC, and *C jejuni/C coli* may also be explained by antigenic diversity. Sapovirus has 4 genogroups and 16 genotypes [37], with a lack of cross-protection [38]. Protection against ETEC is conferred by immune responses to more than 20 different colonization factors [39, 40], and it was previously only observed for ETEC infections of the same toxin-colonization factor profile [41]. Likewise, protection against *Campylobacter* may be related to the polysaccharide capsule, which has a broad diversity of types [42].

The bias analysis using negative controls identified substantial unmeasured confounding in the covariate-adjusted estimates. Larger magnitudes of bias observed for the diarrhea than infection outcomes suggest that confounding by heterogeneity





of susceptibility may be more important than heterogeneity of exposure because the latter would not be expected to differ based on the severity of the outcome. One source of heterogeneity in host susceptibility may be histoblood group antigens status [43]. It is interesting that bias was not observed for norovirus GII and *C jejuni/C coli*, which may reflect the tendency for norovirus to spread indiscriminately in populations (because it is a cause of outbreaks in other settings) [44] and the uniformly high frequency of *Campylobacter* detection [10, 19]. Persistent carriage of *Giardia* [11], *Salmonella* [45], and *H pylori* [46] may explain the strong positive associations with repeated detections for these pathogens, because subsequent detections may not be capturing new infections.

This analysis leverages previous work by using molecular diagnostics across pathogens and attributable fractions to adjudicate diarrhea etiology in the context of frequent subclinical infection. This study design allowed for the novel bias correction based on multiple negative control associations, which generated an empirical null distribution that captured a common distribution of biases [24], but did not require the structure and magnitude of confounding to be identical [47]. However, these analyses were limited by the inability to assess homotypic immunity. The poor sensitivities of typing assays applied directly to stool specimens for rotavirus, *Shigella*, and ST-EPEC resulted in more than 50% of detections being typed as other, which prohibited assessment of homotypic immunity. Furthermore, speciation and typing directly from stool rather than isolates limited our ability to ensure that (1) G and P types for rotavirus and (2) toxins and colonization factors for EPEC were detected in the same organism. In addition, with only monthly sampling, there was likely under-ascertainment of infections. In the 5 sites that had not introduced rotavirus vaccine, 30.4% (n = 238) of first rotavirus infections based on serologic testing at 7 and 15 months of age (IgA  $\geq$ 20 U/mL) were not detected by qPCR. A total of 36.6% of these children were qPCR positive at an older age, at a median age delay of 5.0 months (interquartile range, 2.8–9.1). Because we were unable to make this comparison for the other pathogens, serology was not included as evidence of prior infection in the analysis.

## CONCLUSIONS

Estimates of natural protection for most enteric pathogens were modest and suggest that vaccines that simulate natural infections for the enteric bacteria are unlikely to provide important levels of protection. Vaccines currently in development, such as for *Shigella*, will likely need to provide heterotypic protection, be conjugated, and/or require boosting. Furthermore, protection against infection was limited, such that vaccines are unlikely to provide sterilizing immunity. Therefore, although vaccines could limit the acute burden of diarrheal illness and mortality, they may not effectively address the long-term impacts of subclinical infections, such as growth impairment [19].

## Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

## Notes

**Acknowledgments.** We thank Martijn Schuemie and Laura Hester for their assistance with the methods for empirical calibration using negative controls. We also thank the staff and participants of the MAL-ED Network Project for their important contributions.

**Author contributions.** E. T. R. M. conceived the study and led data analysis and writing of the report. J. L. led data acquisition and interpretation. G. K., M. N. K., A. A. M. L., P. O. B., A. S., R. H., E. R. M., S. S., J. P. L., L. B., N. I., N. P., I. K., Z. B., and T. A. contributed to site data acquisition, the design of the original protocol, and data interpretation. E. R. H. contributed to data acquisition, analysis, and interpretation. J. A. P.-M. conceived the study and contributed to data analysis and interpretation.

**Financial support.** This work was funded by the National Institute of Allergy and Infectious Diseases at the National Institutes of Health (K01AI130326; to E. T. R. M.). The Etiology, Risk Factors and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) was carried out as a collaborative project funded by the Bill & Melinda Gates Foundation (OPP1131125), the Foundation for the National Institutes of Health, and the Fogarty International Center.

**Potential conflicts of interest.** All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

## References

1. Mani S, Wierzbica T, Walker RI. Status of vaccine research and development for *Shigella*. *Vaccine* **2016**; 34:2887–94.
2. Bourgeois AL, Wierzbica TF, Walker RI. Status of vaccine research and development for enterotoxigenic *Escherichia coli*. *Vaccine* **2016**; 34:2880–6.
3. Lucero Y, Vidal R, O’Ryan G M. Norovirus vaccines under development. *Vaccine* **2018**; 36:5435–41.
4. Giersing BK, Modjarrad K, Kaslow DC, et al. Report from the World Health Organization’s Product Development for Vaccines Advisory Committee (PDVAC) meeting, Geneva, 7–9th Sep 2015. *Vaccine* **2016**; 34:2865–9.
5. Gladstone BP, Ramani S, Mukhopadhyaya I, et al. Protective effect of natural rotavirus infection in an Indian birth cohort. *N Engl J Med* **2011**; 365:337–46.
6. Velázquez FR, Matson DO, Calva JJ, et al. Rotavirus infection in infants as protection against subsequent infections. *N Engl J Med* **1996**; 335:1022–8.

7. Fischer TK, Valentiner-Branth P, Steinsland H, et al. Protective immunity after natural rotavirus infection: a community cohort study of newborn children in Guinea-Bissau, west Africa. *J Infect Dis* **2002**; 186:593–7.
8. Pasetti MF, Levine MM. Insights from natural infection-derived immunity to cholera instruct vaccine efforts. *Clin Vaccine Immunol* **2012**; 19:1707–11.
9. Ali M, Emch M, Park JK, Yunus M, Clemens J. Natural cholera infection-derived immunity in an endemic setting. *J Infect Dis* **2011**; 204:912–8.
10. Amour C, Gratz J, Mduma E, et al.; Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) Network Investigators. Epidemiology and impact of *Campylobacter* infection in children in 8 low-resource settings: results from the MAL-ED Study. *Clin Infect Dis* **2016**; 63:1171–9.
11. Rogawski ET, Bartelt LA, Platts-Mills JA, et al.; MAL-ED Network Investigators. Determinants and impact of giardia infection in the first 2 years of life in the MAL-ED birth cohort. *J Pediatric Infect Dis Soc* **2017**; 6:153–60.
12. Rogawski ET, Guerrant RL, Havt A, et al.; MAL-ED Network Investigators. Epidemiology of enteroaggregative *Escherichia coli* infections and associated outcomes in the MAL-ED birth cohort. *PLoS Negl Trop Dis* **2017**; 11:e0005798.
13. Mohan VR, Karthikeyan R, Babji S, et al.; Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development (MAL-ED) Network Investigators. Rotavirus infection and disease in a multisite birth cohort: results from the MAL-ED study. *J Infect Dis* **2017**; 216:305–16.
14. Korpe PS, Valencia C, Haque R, et al. Epidemiology and risk factors for cryptosporidiosis in children from 8 low-income sites: results from the MAL-ED study. *Clin Infect Dis* **2018**; 67:1660–9.
15. MAL-ED Network Investigators. The MAL-ED study: a multinational and multidisciplinary approach to understand the relationship between enteric pathogens, malnutrition, gut physiology, physical growth, cognitive development, and immune responses in infants and children up to 2 years of age in resource-poor environments. *Clin Infect Dis* **2014**; 59(Suppl 4):S193–206.
16. Liu J, Platts-Mills JA, Juma J, et al. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. *Lancet* **2016**; 388:1291–301.
17. Liu J, Gratz J, Amour C, et al. Optimization of quantitative PCR methods for enteropathogen detection. *PLoS One* **2016**; 11:e0158199.
18. Platts-Mills JA, Liu J, Rogawski ET, et al.; MAL-ED Network Investigators. Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. *Lancet Glob Health* **2018**; 6:e1309–18.
19. Rogawski ET, Liu J, Platts-Mills JA, et al.; MAL-ED Network Investigators. Use of quantitative molecular diagnostic methods to investigate the effect of enteropathogen infections on linear growth in children in low-resource settings: longitudinal analysis of results from the MAL-ED cohort study. *Lancet Glob Health* **2018**; 6:e1319–28.
20. Lee GO, Richard SA, Kang G, et al.; MAL-ED Network Investigators. A comparison of diarrheal severity scores in the MAL-ED Multisite Community-Based Cohort Study. *J Pediatr Gastroenterol Nutr* **2016**; 63:466–73.
21. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Annals of Statistics* **1982**; 10:1100–20.
22. Lin DY, Wei LJ. The robust inference for the cox proportional hazards model. *J Am Stat Assoc* **1989**; 84:1074–8.
23. Psaki SR, Seidman JC, Miller M, et al.; MAL-ED Network Investigators. Measuring socioeconomic status in multicountry studies: results from the eight-country MAL-ED study. *Popul Health Metr* **2014**; 12:8.
24. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med* **2016**; 35:3883–8.
25. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proc Natl Acad Sci U S A* **2018**; 115:2571–7.
26. Babji S, Kang G. Rotavirus vaccination in developing countries. *Curr Opin Virol* **2012**; 2:443–8.
27. Rouhani S, Peñataro Yori P, Paredes Olortegui M, et al.; Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) Network Investigators. Norovirus infection and acquired immunity in 8 countries: results from the MAL-ED study. *Clin Infect Dis* **2016**; 62:1210–7.
28. Saito M, Goel-Apaza S, Espetia S, et al.; Norovirus Working Group in Peru. Multiple norovirus infections in a birth cohort in a Peruvian Periurban community. *Clin Infect Dis* **2014**; 58:483–91.
29. Lopman BA, Trivedi T, Vicuña Y, et al. Norovirus infection and disease in an Ecuadorian birth cohort: association of certain norovirus genotypes with host FUT2 secretor status. *J Infect Dis* **2015**; 211:1813–21.
30. Olortegui MP, Rouhani S, Yori PP, et al. Astrovirus infection and diarrhea in 8 countries. *Pediatrics* **2018**; 141:e20171326.
31. Naficy AB, Rao MR, Holmes JL, et al. Astrovirus diarrhea in Egyptian children. *J Infect Dis* **2000**; 182:685–90.

32. Kattula D, Jeyavelu N, Prabhakaran AD, et al. Natural history of cryptosporidiosis in a birth cohort in Southern India. *Clin Infect Dis* **2017**; 64:347–54.
33. Chappell CL, Okhuysen PC, Sterling CR, Wang C, Jakubowski W, Dupont HL. Infectivity of *Cryptosporidium parvum* in healthy adults with pre-existing anti-*C. parvum* serum immunoglobulin G. *Am J Trop Med Hyg* **1999**; 60:157–64.
34. Steiner KL, Kabir M, Priest JW, et al. Fecal immunoglobulin A against a sporozoite antigen at 12 months is associated with delayed time to subsequent cryptosporidiosis in urban Bangladesh: a prospective cohort study. *Clin Infect Dis* **2020**; 70:323–6.
35. Barry EM, Pasetti MF, Sztein MB, Fasano A, Kotloff KL, Levine MM. Progress and pitfalls in *Shigella* vaccine research. *Nat Rev Gastroenterol Hepatol* **2013**; 10:245–55.
36. Ferreccio C, Prado V, Ojeda A, et al. Epidemiologic patterns of acute diarrhea and endemic *Shigella* infections in children in a poor periurban setting in Santiago, Chile. *Am J Epidemiol* **1991**; 134:614–27.
37. Oka T, Mori K, Iritani N, et al. Human sapovirus classification based on complete capsid nucleotide sequences. *Arch Virol* **2012**; 157:349–52.
38. Sánchez GJ, Mayta H, Pajuelo MJ, et al.; Sapovirus Working Group. Epidemiology of sapovirus infections in a birth cohort in Peru. *Clin Infect Dis* **2018**; 66:1858–63.
39. Levine MM, Barry EM, Chen WH. A roadmap for enterotoxigenic *Escherichia coli* vaccine development based on volunteer challenge studies. *Hum Vaccin Immunother* **2019**; 15:1357–78.
40. Vidal RM, Muhsen K, Tennant SM, et al. Colonization factors among enterotoxigenic *Escherichia coli* isolates from children with moderate-to-severe diarrhea and from matched controls in the Global Enteric Multicenter Study (GEMS). *PLoS Negl Trop Dis* **2019**; 13:e0007037.
41. Steinsland H, Valentiner-Branth P, Gjessing HK, Aaby P, Mølbak K, Sommerfelt H. Protection from natural infections with enterotoxigenic *Escherichia coli*: longitudinal study. *Lancet* **2003**; 362:286–91.
42. Guerry P, Poly F, Riddle M, Maue AC, Chen YH, Monteiro MA. *Campylobacter* polysaccharide capsules: virulence and vaccines. *Front Cell Infect Microbiol* **2012**; 2:7.
43. Colston JM, Francois R, Pisanic N, et al. Effects of child and maternal histo-blood group antigen status on symptomatic and asymptomatic enteric infections in early childhood. *J Infect Dis* **2019**; 220:151–62.
44. Graaf M de, Beek J van, Koopmans MPG. Human norovirus transmission and evolution in a changing world. *Nat Rev Microbiol* **2016**; 14:421–33.
45. Gunn JS, Marshall JM, Baker S, Dongol S, Charles RC, Ryan ET. Salmonella chronic carriage: epidemiology, diagnosis, and gallbladder persistence. *Trends Microbiol* **2014**; 22:648–55.
46. Taylor NS, Fox JG, Akopyants NS, et al. Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA fingerprinting. *J Clin Microbiol* **1995**; 33:918–23.
47. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* **2010**; 21:383–8.