Associations between detection of enteropathogens and microbial source tracking markers in the environment and child enteric infections and growth: an individual participant data meta-analysis

Andrew Mertens, Jack Colford, Oliver Cumming, Joe Brown, Jill Stewart, David Holcomb, Drew Capone, Jackie Knee, Tom Clasen, Heather Reese, Amy Pickering, Clair Null, Steve Luby, Jessica Grembi, Ben Arnold, Audrie Lin, Jade Benjamin-Chung, Laura Kwong, Lauren Steinbaum, Ali Boehm, Kara Nelson, Erica Fuhrmeister, Mahbubur Rahman, Sammy Njenga, Rassul Nala, Ruwan Thilakaratne, Ayse Ercumen (middle order not finalized)

**Abstract word count: 304**

**Current word count: 3887**

## Abstract

**Background:** Fecal contamination is typically measured using fecal indicator bacteria (FIB). FIB in environmental samples have been associated with increased risk of diarrhea and reduced linear growth in children. However, FIB are imperfect predictors of enteropathogens, and few studies have assessed associations between detection of enteropathogens and microbial source tracking (MST) markers in the environment and child health outcomes.

**Methods:** We conducted an individual participant data (IPD) meta-analysis to pool data from WASH trials to assess associations between the presence of pathogens and/or MST markers in the environment and enteric infections with specific pathogens, caregiver-reported diarrhea and height-for-age Z-scores (HAZ) in children. We used covariate-adjusted regression models with robust standard errors to estimate associations and pooled results across studies. For the infection and diarrhea outcomes, we used environmental data from up to four months prior to the measurement of health endpoints; we used all available environmental data prior to the HAZ measurement.

**Findings:** We identified and received data from nine studies. Detection of a specific pathogen in environmental samples was consistently associated with increased risk of subsequent child infection with the same pathogen. However, there was no consistent association between detection of enteropathogens or MST markers in the environment and subsequent caregiver-reported diarrhea, except for a significant increase in diarrhea risk associated with any pathogen detection in any sample type during wet seasons. Detection of any pathogen in any sample type was consistently associated with slightly lower HAZ (adjusted pooled mean difference: -0.08 (95% CI: -0.15, -0.01)); there was no association between MST markers and HAZ.

**Interpretation:** Detection of enteropathogens in the environment was associated with increased risk of pathogen-specific infections and lower HAZ but not with caregiver-reported diarrhea, highlighting the limitation of reported diarrheal symptoms as an outcome measure. MST markers showed no consistent associations with any child health outcome.

**Funding:** The Bill & Melinda Gates Foundation.

## Introduction

Undernutrition is a leading contributor to child mortality and morbidity in low and middle income countries, and growth failure from undernutrition is associated with reduced cognitive development and adult income.1 Growth failure leaves children vulnerable to disease, as children with less tissue mass have weakened immune systems, and conversely, infections can lead to growth failure as children use nutrients for their immune systems instead of growth.2,3 Enteric infections may particularly cause growth failure, as both subclinical changes to the gut and symptomatic diarrhea lead to nutrient loss, and diarrhea also is a leading cause of death of children younger than 5 years, causing an estimated 534,000 deaths in 2017.4 Poor water, sanitation, and hygiene (WASH) conditions lead to the contamination of the local environment with fecal pathogens, with an estimated 62% of deaths from diarrhea and 16% of growth failure among children under 5 years attributed to fecal exposure from poor WASH.5

However, several large, recent trials of WASH interventions found small or null effects on child diarrhea and growth, which may be because the interventions failed to reduce environmental fecal contamination, or because environmental fecal contamination from inadequate WASH was not the primary cause of child diarrhea or growth failure in those populations.6–8 Fecal contamination is usually assessed by testing for the presence of fecal indicator bacteria (FIB) such as *E. coli*, which have been associated with increased risk of diarrhea and reduced linear growth in children.9 FIB are general markers of fecal contamination which can originate from non-fecal sources,10, cannot differentiate between human vs. animal fecal sources,11, and cannot tell if fecal contamination detected in a household’s environment contains the enteropathogens that cause diarrheal disease in children in the household.

Few studies have assessed associations between DNA-based detection of specific enteropathogens and microbial source tracking (MST) markers in the household environment and child health outcomes. We assessed associations between detection of enteropathogens and MST markers in the environment and child diarrheal disease, pathogen-specific enteric infections, and growth failure with a systematic review and an individual participant data (IPD) meta-analysis. By examining specific enteropathogens and source tracking markers in different types of household samples (source and stored drinking water, soil, captured flies, and mothers’ and children’s hands) as well as specific pathogens in children’s stool, we aimed to investigate the specific pathways through which environmental contamination influence child health. This may help understand the failure of previous WASH interventions and guide the development and implementation of future WASH interventions.

## Methods

We conducted a systematic literature search to identify WASH intervention studies that have measured pathogens and/or MST markers in environmental samples as well as at least one of the following health outcomes in children: caregiver-reported diarrhea, growth or pathogen detection in stool. We included studies meeting the following inclusion criteria: 1) prospective studies with a water, sanitation, or hygiene intervention and concurrent control (i.e., randomized controlled trial, matched cohort, controlled before-and-after study), 2) measured pathogens and/or MST markers in environmental samples, and 3) measured child anthropometry, diarrheal disease, or pathogen-specific infections for use in a companion manuscript [Mertens et al. 2022 in prep]. We excluded studies that only measured fecal indicator bacteria such as coliforms or *E. coli*. We only included studies published after 2000 to capture more recently developed advanced pathogen detection methods. Details on the search strategy are in Mertens et al. 2022.

We examined associations between enteropathogens and MST markers in the environment and child health outcomes, including enteropathogen-specific infections, caregiver-reported diarrheal disease and growth . We classified enteropathogens and MST markers in the environment into multiple exposure variables. Our primary exposure variables were the prevalence of any enteropathogen or any MST markers in any type of environmental sample. We also tabulated prevalence separately for each sample type (source or stored drinking water, mothers’ and children’s hands, household and latrine soil, flies). Secondary exposure variables included the prevalence of specific pathogen types (any viruses, any bacteria, any protozoa, any helminths), the prevalence of MST markers from specific host types (human or animal), and the prevalence and abundance of individual enteropathogens and MST markers. The primary outcomes were height-for-age Z-scores (HAZ) and 7-day prevalence of caregiver-reported diarrheal disease. For specific enteropathogens detected in the environment, primary outcomes also included child infection with the same pathogen ascertained by stool testing. Secondary outcomes included Z-scores for weight-for-age (WAZ) and weight-for-length (WLZ) and the prevalence of stunting, underweight and wasting, defined as a Z-score below -2 for HAZ, WAZ and WHZ, respectively. For the growth outcomes, we used data from all environmental samples collected over the child’s lifetime prior to the anthropometry measurement, and used the child growth measurement taken closest to environmental sampling if there were repeated measurements. For diarrheal disease and enteropathogen-specific infections, we only used environmental samples collected up to four months before the sampling of the child.

For binary outcomes (prevalence of pathogen-specific infection, diarrhea, stunting, underweight, and wasting), we estimated prevalence ratios associated with the different exposure variables using modified Poisson regression.12 For continuous outcomes (child anthropometry Z-scores), we used linear regression to estimate mean differences. Because of repeated sampling or clustered designs in some studies, we used the Huber Sandwich Estimator to calculate robust standard errors.13 All analyses were adjusted for potential confounders. We included child age and asset-based household wealth as adjustment covariates for all adjusted estimates. Other covariates were prescreened using likelihood ratio tests, and only variables associated with the outcome with a p-value < 0.2 were included in the model for each outcome. We included the following variables in the prescreening set if they were measured within an included study: study arm, child age, child sex, maternal age, household food security status, number of people in the household, age and education of primary caregiver in the household, asset-based household wealth, number of rooms, construction materials (walls, floor, roof), access to electricity, land ownership and if anyone in the household works in agriculture. Within each study, we only estimated associations when there were at least 5 cases of the binary outcome in the rarest stratum of the exposure.

Given the heterogeneity in study settings (e.g., local WASH and nutrition conditions, climate, urbanization, population density, region-specific infectious disease patterns), we reported individual study-specific estimates for all analyses. For outcomes where data were available from four or more studies, we tested for heterogeneity in estimates using Cochran’s Q-test.14 If there was no significant heterogeneity (p-value>0.2), we pooled estimates using fixed-effects models, otherwise we pooled estimates using random-effects models.

We conducted subgroup analyses by child age (immobile vs. crawling vs. walking pre-school-age vs. school-age) and sex, animal ownership in the household, season (dry vs. wet), study setting (rural vs. urban). The wet season for each study was defined as the 6 months of highest average rainfall, obtained from <https://www.weather-atlas.com/>.15. For age, sex, animal presence and season, we included interaction terms between the exposure and the indicator variable for the subgroup in the regression models; a p-value <0.2 on the interaction term was considered evidence of effect modification. There was no variation in urbanicity within individual studies; we separately pooled estimates from urban vs. rural studies to assess effect modification.

As sensitivity analyses, we compared covariate-adjusted estimates with unadjusted estimates. We also compared adjusted estimates from parametric regression models with adjusted estimates from flexible machine-learning based targeted maximum likelihood estimation models.16 Additionally, to assess the impact of the assumption that environmental contamination is related to diarrheal disease within four months, we re-estimated associations between environmental contamination and child diarrhea using environmental data collected within a month prior to the diarrhea measurement, as well as environmental data collected at any time with respect to the diarrhea measurement.

All analyses were conducted in R 4.1.4, and analysis scripts are publicly available (<https://github.com/amertens/wash-ipd>). The systematic review search strategies and the analysis plan were pre-registered on Open Science Framework (<https://osf.io/8sgzn/>). Our PRISMA checklist can be found in Supplementary Table S1.

## Results

### Search results and data acquisition

The systematic review was conducted on 1/19/2021 and returned 3,376, of which nine both met the inclusion criteria and had study authors who agreed to share data. The nine publications reported findings from five unique intervention studies: WASH Benefits Bangladesh and Kenya trials,17 the Maputo Sanitation (MapSan) study in Mozambique,18 the Gram Vikas study in India,19 and the Odisha Total Sanitation Campaign trial in India,20. Because the different individual studies collected environmental measurements from different subsets of trial participants at different times, we report results stratified by publication rather than trial. For the Odisha Total Sanitation Campaign trial, only village-level source water quality data were shared. The studies collected a range of sample types (source and stored drinking water, child and mother hands, soil from the courtyard, household and latrine areas, food, and flies caught in the compound’s latrine and kitchen areas), and measured a range of bacterial, viral, protozoan and helminthic pathogens, including pathogenic *E. coli, V. cholerae, Shigella, Campylobacter, Salmonella, Yersinia, C. difficile*, rotavirus, norovirus, sapovirus, adenovirus, astrovirus, pan-enterovirus, *Cryptosporidium, Giardia, Entamoeba histolytica, Ascaris lumbricoides and Trichuris trichiura*. Details on study designs, intervention design and uptake, environmental sample collection, and laboratory methods are in Mertens et al. 2022. The number of child diarrhea observations with time-matched environmental samples ranged from 210-2034 observations and diarrhea prevalence ranged from 6.1-25.9 across studies (Table 1). The number of HAZ observations with time-matched environmental samples ranged from 202 to 1800 observations and mean HAZ ranged from -1.90 to -1.33 (Table 1).

#### Associations with pathogen-specific infections

Detection of a specific enteropathogens in the compound environment was associated with increased prevalence of subsequent infection with the same pathogen in children living in the compound; trends were consistent across different enteropathogens and sample types (Figure 1). *Clostridium difficile*, *Ascaris* and *Trichuris* detected in courtyard soil and *Shigella* and *Trichuris* detected in flies were significantly associated with increasing prevalence of infection with the same pathogens. Pathogenic *E. coli* and *Giardia* detected in soil and/or flies also had borderline associations with increasing infection prevalence. Few studies had time-matched data on water/hand samples and child infections; these studies found no association between pathogens measured in water or on hands and child infections with the same pathogen (Figure 1).

#### Associations with diarrhea

The presence of any enteropathogen in any type of environmental sample was not associated with diarrheal disease, except for significantly increased diarrhea prevalence associated with any enteropathogen detection on child hands in WASH Benefits Bangladesh (Figure 2); we note that the only pathogen investigated in this study was rotavirus.21 When broken down by groups of pathogens, bacteria on child hands22 and protozoa in soil23 were also borderline associated with increasing risk of diarrhea but most associations between pathogen groups and child diarrhea were null (Figure S1). Similiarly, most associations between specific pathogens in the environment and diarrhea were null, but rotavirus on child hands21 and *Giardia* in latrine soil23 were both significantly associated with an approximately two-fold increase in diarrhea risk (Figure S2). Detection of *Ascaris*, astrovirus and *Clostridium difficile* in soil23,24 and pathogenic *E. coli* on child hands22 was also borderline associated with increased risk (Figure S2). Examining enteropathogen abundance revealed similar trends; increasing abundance of *Ascaris* and rotavirus in soil21,24 and rotavirus on child hands21 was associated with increasing risk of diarrhea (Figure S3).

There was no significant associations between the presence of any MST marker or specific groups of MST markers (human or animal) and child diarrheal disease in any sample type (Figure 2, Figure S1). Among specific markers, detection of the avian marker GFD in soil was significantly associated with an over two-fold increase in diarrhea risk,25 and the same marker in stored water and on child hands was also borderline associated with increased diarrhea (Figure S4). Most other human and animal markers were not associated with diarrhea.

#### Associations with child growth

The presence of any enteropathogen in any environmental sample was significantly associated with lower HAZ when pooled across studies (adjusted mean difference: -0.08 (95% CI: -0.15, 0.00), Figure 3). This was driven primarily by the number of slightly harmful but insignificant effects rather than by any strong effect in specific studies, with the exception of water samples with any enteropathogen presence being significantly associated with lower mean HAZ.21 There was also a borderline association between detection of any pathogen in household soil and lower HAZ when pooled across studies (adjusted mean difference: -0.07 (95% CI: -0.15, 0.02), Figure 3).

When broken down by groups of enteropathogens, presence of viruses in stored water and protozoa on child hands was significantly associated with a reduction in HAZ on the order of 0.5 in individual studies21,22 while presence of bacteria in source or stored water and STH in soil were borderline associated with lower HAZ (Figure S1).21,26 Individual pathogens whose detection was significantly associated with reduced HAZ were *Ascaris* in soil and flies, *E. histolytica* in soil, *Giardia* on child hands and rotavirus in water (Figure S2). However, many associations between individual pathogens and HAZ were null, and multiple pathogens in different sample types were associated with higher HAZ (Figure S2). Similarly, there were inconsistent associations between the abundance of specific enteropathogens and child HAZ, with most estimates having null effects, and with significant effects occurring in both harmful and protective directions, though the direction of effects were consistent between prevalence and abundance exposures (Figure S3). For other measures of growth, associations between the presence/abundance of enteropathogens and WAZ, WHZ, stunting and wasting were mostly inconsistent but the presence and increasing abundance of rotavirus in stored water was consistently associated with both reduced WAZ and WHZ (Figures S1-S3) and many individual pathogens showed some degree of association with increased risk of being underweight (Figures S1-S3).

The presence of any MST in any environmental sample was not associated with HAZ when pooled across studies, and individual studies showed significant associations both in the harmful and protective direction (Figure 3). Among groups of MST markers, the associations were null for most combinations of markers and sample types but aggregate and animal fecal markers in stored water were associated with lower HAZ while human markers showed associations in both directions (Figure S1). There were inconsistent associations between the presence of specific MST markers and HAZ, with most estimates having null effects (Figure S4). Of the statistically significant associations, half of the sample-specific estimates were associated with increased HAZ and half were associated with decreased HAZ. Associations between the abundance of specific MST markers and HAZ were similarly inconsistent (Figure S5). For other measures of growth, there were inconsistent associations between the presence or abundance of any MST marker and WAZ, WHZ, stunting, underweight and wasting across studies, with most estimates having null effects, and with significant effects occurring in both harmful and protective directions (Figures S1, S4, S5). However, some markers were consistently associated with reduced growth across multiple metrics within individual studies, such as the animal marker BacCow in multiple sample types,23 and the avian marker GFD and ruminant marker BacR in stored water [Boehm 2016] (Figures S4, S5). The abundance of MST markers generally had similar association with health outcomes as the corresponding prevalences, though the abundance but not presence of BacCow was associated with lower HAZ and more stunting in household soil, stored water, and flies, and the abundance but not presence of HumM2 was associated with more stunting and wasting in household soil.

#### Subgroup analyses

There were no consistent differences in associations between enteropathogens or MST Markers and diarrhea or between enteropathogens or MST Markers and HAZ when analyses were stratified by child age (Figures S6, S7). However, most studies did not have children measured in all age categories. There was also no significant effect of child sex on associations between environmental pathogens or MST markers and diarrheal disease (Figure S8). Pooled across studies, pathogen presence was associated with twice the reduction in HAZ in boys (adjusted mean difference: -0.08 (95% CI: -0.16, -0.01)) than in girls (adjusted mean difference: -0.18 (95% CI: -0.30, -0.06), Figure S9). The decrease in HAZ associated with the presence of any pathogen in any environmental sample was higher among boys than in most individual studies, though the difference was only significant in Kenya (Figure S9).24 There was no effect modification by sex on associations between MST markers and HAZ.

Diarrheal disease was too sparse to estimate differences in associations between households with and without animals, but, when pooled across studies, there was a significant decrease in child HAZ in compounds with any sample with any enteropathogen detected when the child lived in a compound with no animals, but not in compounds that had animals (Figure S10). There was no interaction by animal presence and MST.

Also when pooled across studies, there was a significant increase in child diarrheal disease risk in compounds with any sample with any enteropathogen detected when the child diarrheal disease occurred during the wet season (Figure S11). There was no association with between MST markers and diarrhea in either season. There were also no significant differences in pooled estimates between the one urban study (Holcomb et al. 2020) and the four rural studies for any combination of exposures and outcomes.

#### Sensitivity analyses

Most covariates were not strongly associated with enteropathogen or MST marker presence in the environment, suggesting they are not strong confounders of the relationship between these exposures and our child health outcomes (Figure S12). Measures of household wealth generally had the strongest association with environmental contamination, though the association varied by study and microbial target. Additionally, data sparsity allowed controlling for a small number of covariates in most analyses. On average, covariate adjustment had small effects on the results; adjusted estimates were slightly larger in magnitude than unadjusted estimates and the effect of adjustment was slightly more pronounced when a larger number of covariates were used for adjustment (Figures S13-S14). Comparison between associations estimated with generalized linear models (GLM) vs. machine-learning based targeted likelihood estimation models (TMLE) showed no major differences, indicating that the linear assumptions and lack of interactions in the GLMs used for the primary analysis did not lead to greater residual confounding compared to more flexible methods (Figures S15-S16). Lastly, results were similar when we used data from environmental samples up to four months prior, one month prior or at any time with respect to diarrhea measurements (Figure S17).

## Discussion

Detection of enteropathogens in the compound environment was associated with increased risk of subsequent infection with the same pathogen among children living in the same compound, as well as with lower HAZ pooled across studies, especially among boys. Enteropathogen detection in the environment overall was not associated with risk of subsequent diarrhea, except during the rainy season, but we observed associations between some individual pathogens and higher diarrhea risk. MST markers were generally were not associated with diarrhea, except for the avian GFD marker. Associations between MST markers and child growth outcomes were inconsistent across studies but detection of specific animal markers in environmental samples was associated with substantially reduced growth metrics within individual studies.

Positive associations between detection of pathogens in the environment and subsequent detection in child stool samples demonstrates environmental transmission and provides a link in the causal chain between environmental contamination and lower child HAZ. However, the reduction in HAZ associated with enteropathogens in the environment was modest; small non-significant effects in individual studies became significant when pooled, highlighting the strength of IPD meta-analyses. The diarrhea outcome did not capture these risks, highlighting the limitations of self-reported all-cause diarrhea as an outcome to assess the impact on environmental pathogen contamination. This is consistent with research on pathogens in recreational water, where specific enteropathogens had limited association with all-cause diarrhea.27

Pathogens in the environment that had associations with increased risk of child diarrhea included rotavirus, *Giardia* and, to a smaller extent, pathogenic E. coli, *Ascaris*, astrovirus and *Clostridium difficile*. These findings are consistent with multi-country case-control studies that have identified rotavirus, pathogenic *E. coli*and astrovirus among the pathogens with the highest attributable burden of diarrhea in low-income countries.28,29 Other dominant pathogens in these studies included Cryptosporidium, *Shigella*, Campylobacter and norovirus; we note that we did not have sufficient time-matched data to estimate associations between detection of these pathogens in the environment and child diarrhea.

The avian marker GFD was the only MST marker associated with increased risk of diarrhea and detection of animal but not human fecal markers was consistently associated with reduced child growth. These findings support growing evidence that animals, specifically poultry, are a major source of diarrhea transmission in low-income countries.30

Soil and child hands stood out as dominant pathways of environmental diarrhea transmission, and pathways associated with reduced child growth also included source and stored water. A recent meta-analysis showed increased risk of diarrhea associated with increasing levels of fecal indicator bacteria in drinking water and on child hands, and reduced HAZ associated with increasing levels of fecal indicator bacteria in drinking water.9 Child hands have been identified a major source of children’s fecal exposure, in terms of frequency of mouth contacts,31 estimated *E. coli* ingestion32 and associations with diarrhea.33 Similarly, soil has been shown to account for a significant portion of estimated *E. coli*ingestion for children34 and ingestion of soil has been associated with environmental enteric dysfunction and stunting in children.35,36 Our findings corroborate the role of child hands and soil in diarrhea transmission.

Our analysis had several limitations. Due to the smaller sample size of the environmental samples within the eligible studies, the low prevalence of diarrheal disease in children in many individual studies, and the rare detection of many of the enteropathogens in environmental samples, data sparsity limited the feasible analyses. Many exposure-outcome associations were not estimated due to sparse data and there was only a small number of pathogens measured in both the environment and subsequently in children’s stool. Additionally, we could only adjust for a small subset of potentially confounding covariates in some analyses due to the small number of available observations. However, most covariates were weakly associated with measures of environmental contamination, and our unadjusted and adjusted estimates were similar, even when controlling for a larger number of covariates. Flexible covariate adjustment through TMLE did not change the associations between environmental contamination and diarrheal disease or HAZ. Therefore, we believe our modeling approach adequately adjusted for measured confounding but unmeasured confounding may bias our results. We did not correct for multiple comparisons, and so some significant associations are likely type-1 errors, especially when inconsistent with results across sample types and individual studies. The differences in the lengths of time between environmental and child health measurements across different studies may have also led to inconsistencies in associations between studies. However, shrinking or expanding the window we allowed between environmental and diarrhea measurements in our analyses did not change our findings (Figure S17).

Future studies investigating the associations between environmental fecal contamination and child health should ascertain health outcomes soon after environmental sampling, and should both focus on enteropathogens in the environment instead of MST markers, which had limited association with child health and only indicate the fecal source rather than the pathogenicity of the environmental contamination, and on enteropathogen specific infections instead of all cause diarrhea.

## References

1. Black, M. M. *et al.* Early childhood development coming of age: Science through the life course. *Lancet (London, England)* **389**, 77–90 (2017).

2. Chang, C. Y. *et al.* Children successfully treated for moderate acute malnutrition remain at risk for malnutrition and death in the subsequent year after recovery. *The Journal of Nutrition* **143**, 215–220 (2013).

3. McDonald, C. M. *et al.* The effect of multiple anthropometric deficits on child mortality: Meta-analysis of individual data in 10 prospective studies from developing countries. *The American Journal of Clinical Nutrition* **97**, 896–901 (2013).

4. James, S. L. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392**, 1789–1858 (2018).

5. Prüss-Ustün, A. *et al.* Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: An updated analysis with a focus on low- and middle-income countries. *International Journal of Hygiene and Environmental Health* **222**, 765–777 (2019).

6. Humphrey, J. H. *et al.* Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: A cluster-randomised trial. *The Lancet. Global Health* **7**, e132–e147 (2019).

7. Luby, S. P. *et al.* Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised controlled trial. *The Lancet Global Health* **6**, e302–e315 (2018).

8. Null, C. *et al.* Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: A cluster-randomised controlled trial. *The Lancet. Global Health* **6**, e316–e329 (2018).

9. Goddard, F. G. B. *et al.* Faecal contamination of the environment and child health: A systematic review and individual participant data meta-analysis. *The Lancet Planetary Health* **4**, e405–e415 (2020).

10. Hardina, C. M. & Fujioka, R. S. Soil: The environmental source of Escherichia coli and enterococci in Hawaii’s streams. *Environmental Toxicology and Water Quality* **6**, 185–195 (1991).

11. Gruber, J. S., Ercumen, A. & Colford, J. M. Coliform bacteria as indicators of diarrheal risk in household drinking water: Systematic review and meta-analysis. *PloS One* **9**, e107429 (2014).

12. Zou, G. A modified poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology* **159**, 702–706 (2004).

13. Freedman, D. A. On The So-Called ‘Huber Sandwich Estimator’ and ‘Robust Standard Errors’. *The American Statistician* **60**, 299–302 (2006).

14. Cochran, W. G. The Combination of Estimates from Different Experiments. *Biometrics* **10**, 101–129 (1954).

15. d.o.o, Y. M. G. Weather Atlas | Weather forecast and Climate information for cities all over the Globe. *Weather Atlas*.

16. Schuler, M. S. & Rose, S. Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. *American Journal of Epidemiology* **185**, 65–73 (2017).

17. Arnold, B. F. *et al.* Cluster-randomised controlled trials of individual and combined water, sanitation, hygiene and nutritional interventions in rural Bangladesh and Kenya: The WASH Benefits study design and rationale. *BMJ Open* **3**, e003476 (2013).

18. Brown, J. *et al.* A controlled, before-and-after trial of an urban sanitation intervention to reduce enteric infections in children: Research protocol for the Maputo Sanitation (MapSan) study, Mozambique. *BMJ Open* **5**, e008215 (2015).

19. Reese, H. *et al.* Assessing longer-term effectiveness of a combined household-level piped water and sanitation intervention on child diarrhoea, acute respiratory infection, soil-transmitted helminth infection and nutritional status: A matched cohort study in rural Odisha, India. *International Journal of Epidemiology* **48**, 1757–1767 (2019).

20. Clasen, T. *et al.* Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: A cluster-randomised trial. *The Lancet. Global Health* **2**, e645–653 (2014).

21. Boehm, A. B. *et al.* Occurrence of Host-Associated Fecal Markers on Child Hands, Household Soil, and Drinking Water in Rural Bangladeshi Households. *Environmental Science & Technology Letters* **3**, 393–398 (2016).

22. Fuhrmeister, E. R. *et al.* Effect of Sanitation Improvements on Pathogens and Microbial Source Tracking Markers in the Rural Bangladeshi Household Environment. *Environmental Science & Technology* **54**, 4316–4326 (2020).

23. Capone, D. *et al.* Impact of an Urban Sanitation Intervention on Enteric Pathogen Detection in Soils. *Environmental Science & Technology* **55**, 9989–10000 (2021).

24. Steinbaum, L. *et al.* Effect of a sanitation intervention on soil-transmitted helminth prevalence and concentration in household soil: A cluster-randomized controlled trial and risk factor analysis. *PLoS neglected tropical diseases* **13**, e0007180 (2019).

25. Holcomb, D. A. *et al.* Human fecal contamination of water, soil, and surfaces in households sharing poor-quality sanitation facilities in Maputo, Mozambique. *International Journal of Hygiene and Environmental Health* **226**, 113496 (2020).

26. Reese, H. Effectiveness of a combined sanitation and household-level piped water intervention on infrastructure coverage, availability and use, environmental fecal contamination, and child health in rural Odisha, India: A matched cohort study. (Emory University, 2017).

27. Griffith, J. F. *et al.* Epidemiologic evaluation of multiple alternate microbial water quality monitoring indicators at three California beaches. *Water Research* **94**, 371–381 (2016).

28. Kotloff, K. L. *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *The Lancet* **382**, 209–222 (2013).

29. Platts-Mills, J. A. *et al.* Pathogen-specific burdens of community diarrhoea in developing countries: A multisite birth cohort study (MAL-ED). *The Lancet Global Health* **3**, e564–e575 (2015).

30. Ercumen, A. *et al.* Poultry Ownership Associated with Increased Risk of Child Diarrhea: Cross-Sectional Evidence from Uganda. *The American Journal of Tropical Medicine and Hygiene* **102**, 526–533 (2020).

31. Kwong, L. H. *et al.* Hand- and Object-Mouthing of Rural Bangladeshi Children 3-18 Months Old. *International Journal of Environmental Research and Public Health* **13**, E563 (2016).

32. Mattioli, M. C. M., Davis, J. & Boehm, A. B. Hand-to-mouth contacts result in greater ingestion of feces than dietary water consumption in Tanzania: A quantitative fecal exposure assessment model. *Environmental Science & Technology* **49**, 1912–1920 (2015).

33. Pickering, A. J. *et al.* Fecal Indicator Bacteria along Multiple Environmental Transmission Pathways (Water, Hands, Food, Soil, Flies) and Subsequent Child Diarrhea in Rural Bangladesh. *Environmental Science & Technology* **52**, 7928 (2018).

34. Kwong, L. H. *et al.* Effect of sanitation improvements on soil-transmitted helminth eggs in courtyard soil from rural Bangladesh: Evidence from a cluster-randomized controlled trial. *bioRxiv* 2020.09.29.318097 (2020) doi:[10.1101/2020.09.29.318097](https://doi.org/10.1101/2020.09.29.318097).

35. George, C. M. *et al.* Geophagy Is Associated with Environmental Enteropathy and Stunting in Children in Rural Bangladesh. *The American Journal of Tropical Medicine and Hygiene* **92**, 1117–1124 (2015).

36. Morita, T. *et al.* Mouthing of Soil Contaminated Objects is Associated with Environmental Enteropathy in Young Children. *Tropical medicine & international health: TM & IH* **22**, 670–678 (2017).