Effects of water, sanitation, and hygiene interventions on detection of enteropathogens and host-specific faecal markers in the environment: an individual-participant data meta-analysis

Andrew Mertens, Benjamin F. Arnold, Jade Benjamin-Chung, Alexandria Boehm, Joe Brown, Drew Capone, Thomas Clasen, Erica Fuhrmeister, Jessica Grembi, David Holcomb, Jackie Knee, Laura H Kwong, Audrie Lin, Stephen P. Luby, Rassul Nala, Kara Nelson, Sammy Njenga, Clair Null, Amy J. Pickering, Mahbubur Rahman, Heather Reese, Lauren Steinbaum, Jill Stewart, Ruwan Thilakaratne, Oliver Cumming, John M. Colford Jr., Ayse Ercumen

Abstract

Background: Water, sanitation, and hygiene (WASH) improvements are promoted to reduce diarrhoea in low-income countries. However, recent trials have found mixed effects of household- and community-level WASH interventions on child health. Measuring pathogens and host-specific faecal markers in the environment can help investigate causal pathways between WASH and health by quantifying if and how much interventions reduce environmental exposure to enteric pathogens and faecal contamination from human and different animal sources.

Methods: We conducted a systematic review and individual participant data meta-analysis to assess the effects of WASH interventions on enteropathogens and microbial source tracking (MST) markers in environmental samples. We used covariate-adjusted regression models with robust standard errors to estimate intervention effects and pooled results across studies.

Findings: Few trials have measured the effect of sanitation interventions on pathogens/MST markers in the environment, and we identified no trials of drinking water or hygiene interventions. We extracted individual participant data on nine comparisons from five eligible trials. Environmental sampling was primarily focused on onsite sanitation interventions and included drinking water, hand rinses, soil and flies. Studies consistently indicated that interventions were associated with reduced pathogen detection in the environment but effect estimates in most individual studies could not be distinguished from chance. Pooled across studies, we found a small but significant reduction in the prevalence of any pathogen in any sample type (prevalence ratio [PR]: 0·94 (95% CI: 0·90, 0·99)). There was no effect on MST markers from humans (pooled PR: 1·00 (95% CI: 0·88, 1·13)) or animals (pooled PR: 1·00 (95% CI: 0·97, 1·03)).

Interpretation: The small effect of these sanitation interventions on pathogen detection and lack of effect on human/animal faecal markers in the studies are consistent with the limited health impact previously reported in these trials. Our findings suggest that the basic sanitation interventions implemented in these studies failed to contain human waste and reduce exposure to enteropathogens in the environment.

Funding: The Bill & Melinda Gates Foundation.

Research in context

Evidence before this study. Children in areas with poor drinking water, sanitation, and hygiene (WASH) conditions experience increased diarrhoeal disease and reduced growth. Recent rigorous trials of WASH interventions have shown mixed effectiveness in reducing diarrhoeal disease in children and no improvements in child growth. Quantifying the effect of WASH improvements on enteric pathogens in environmental samples and on contamination originating from human vs. animal sources may help elucidate if interventions successfully interrupt the causal pathway between poor WASH, environmental exposure to faecal pathogens, and child health. Most previous studies and meta-analyses on the effect of WASH interventions on faecal contamination in the environment have focused on faecal indicator bacteria (FIB). Studies have shown FIB reductions in water and on hands due to water treatment and handwashing, respectively, but no effects from sanitation. However, naturalized FIB can be present in the environment without faecal contamination, and FIB also poorly correlate with actual pathogen presence, limiting the interpretation of these findings. FIB also cannot distinguish between faecal contamination from humans and animals; this information can help illuminate whether lack of health effects from sanitation interventions are due to zoonotic disease transmission from unsafely managed animal feces. Recent applications of advanced analytic techniques to environmental sampling in low-income countries allow detection and enumeration of a range of enteropathogens, as well as faecal markers associated with specific hosts for microbial source tracking (MST). We conducted a systematic review and individual participant data (IPD) meta-analysis of WASH intervention studies to assess if interventions reduced the prevalence and abundance of enteropathogens or MST markers in the domestic environment.

Added value of this study. We identified 6 eligible intervention studies that measured enteropathogens and MST markers in environmental samples and obtained data from 5 studies. Studies consistently indicated that the interventions were associated with reduced detection of pathogens, and some pathogens (e.g., adenovirus, Shigella) showed large significant reductions but most effect estimates in individual studies were not statistically significant due to small sample sizes and infrequent detection of some of the targets. The IPD meta-analysis approach with increased statistical power allowed us to detect a small but significant reduction in the prevalence of any type of pathogen in any type of sample by pooling across all studies. There was no overall intervention effect on the prevalence of human or animal MST markers. This study takes advantage of advanced methods to enumerate enteropathogens and host-specific faecal markers in a range of environmental samples, including understudied reservoirs such as soil. We provide the first synthesis of evidence of the effect of WASH interventions on these important targets to advance our understanding of the environmental mechanisms of interventions beyond the available evidence on effects on FIB.

Implications of all the available science. The environmental sampling in the studies in our review was mostly focused on onsite sanitation interventions that provided or promoted improved latrines with pits or septic tanks at the household or community level. The small reduction in pathogen prevalence in the environment when pooled across all studies may explain the small effect these interventions had on child health in the parent studies. Taken together, this evidence suggests that the sanitation interventions implemented in the studies in our review achieved a small reduction in faecal contamination in the environment. More comprehensive WASH interventions, such as safely managed water and sanitation, including safe disposal and/or treatment of excreta from both humans and animals, are potentially needed to reduce environmental contamination enough to improve child health. We note that only a small number of intervention studies measured our targets of interest, and we identified no studies that assessed the effect of water supply/treatment, hygiene or safely managed and/or sewered sanitation interventions on pathogens and MST markers in the domestic environment. Also, pathogen targets and analytic methods varied by study, limiting comparability. Future research would benefit from environmental sampling following implementation of a more diverse and comprehensive set of WASH interventions. Such studies should enumerate a common range of pathogen targets and use standardized laboratory methods for a given target and environmental matrix.

Introduction

Every year, diarrhoea kills an estimated 525,000 children under five.[1] Enteropathogens (pathogens causing infections of the intestinal track) are transmitted from infected individuals' feces to new hosts through a diverse set of interconnected environmental pathways. Water, sanitation and hygiene (WASH) improvements have been promoted to reduce childhood enteric infections by reducing environmental

exposure to faecal-orally transmitted pathogens. Until recently, trials of WASH interventions have primarily focused on documenting health outcomes such as caregiver-reported diarrhoea without measuring intermediate outcomes along the causal chain, such as pathogens in environmental samples to characterize exposure. Such measurements can illuminate underlying mechanisms of interventions and offer explanations for intervention success or failure. Inspecting the causal chain is especially important given the small or null effects on child diarrhoea and growth in recent well-conducted WASH trials.[2–6]

To date, WASH intervention studies measuring environmental contamination have mostly relied on faecal indicator bacteria (FIB) such as *E. coli* as a proxy for a wide range of enteropathogens, including bacteria, viruses, protozoa, and helminths. Sampling has primarily focused on drinking water (and, to a smaller extent, hands and food) while other pathways such as soil and surfaces have received less attention.[7] Household water treatment and handwashing have been associated with reduced faecal contamination of drinking water[8] and hands,[9,10] respectively, while sanitation interventions have had limited impact on FIB in/on drinking water, hands, objects, surfaces, soil and flies.[7] However, FIB are imperfect predictors of faecal contamination, pathogen presence and ultimate health risk. While *E. coli* in drinking water is correlated with increased risk of diarrhoea,[11] FIB can also originate from non-faecal sources[12] and generally correlate poorly with pathogens in the environment.[13] In addition, FIB are found in both human and animal feces, and their detection in the environment therefore cannot differentiate the source of contamination.[11,12,14]

Recent applications of advanced molecular methods to environmental sampling in low-income settings can offer advantages over FIB measurements in characterizing environmental contamination. These methods can detect a range of enteropathogens,[15,16] or distinguish between human vs. animal faecal sources through microbial source tracking (MST).[17] We assessed the effect of WASH interventions in low-income countries on enteropathogens and human- and animal-specific MST markers in the domestic environment with a systematic review and individual participant data (IPD) meta-analysis, which allows combining observation-level data from studies with standardized statistical methods.

Methods

Search methods

We searched the PubMed, Embase, CAB Direct Global Health, Agricultural & Environmental Science Database, Web Of Science, and Scopus databases (Tables S1 and S2). 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), consistent with prior WHO burden of disease reviews,[18,19] 2) measured pathogens and/or MST markers in environmental samples, and 3) measured child anthropometry, diarrhoea, or pathogen-specific infections. We restricted the search to studies measuring child health outcomes to estimate associations between environmental contamination and child health in a separate analysis.[20] We included studies published after 2000 to reflect recent advances in laboratory methods but we did not limit our search to any specific method (e.g., molecular, culture-based, microscopy). We excluded studies that only measured FIB. We limited our search to studies in English. One reviewer (AM) screened abstracts, and two independent reviewers (AM, RT) examined the full texts of short-listed articles with differences resolved with a third reviewer (AE). We followed PRISMA reporting guidelines (Figure S1, Table S3). We used an adapted version of the Newcastle-Ottawa scale to evaluate bias (Table S4).[21]

Data collection and analysis

For each eligible study, we requested individual data from authors. If the corresponding author was unwilling to share individual data, the study was excluded from our analysis. Prior to sharing data, all personal identifiers such as GPS locations were removed, and indirect identifiers such as sampling dates were coarsened to a monthly resolution.

We did not pre-specify specific enteropathogens or markers as outcomes because each study measured a different set of targets. We used the pathogens and MST markers measured in the included studies to generate two composite measures as our primary outcomes: (i) detection of any pathogenic target or (ii) any MST target, in any sample type collected during the same sampling round from the same compound, where a compound was defined by the original studies as a set of households with common courtyards, water sources, or latrines. Because many targets were infrequently detected in individual studies,

composite outcomes allowed us to pool information from studies that focused on different targets and sample types, leveraging the IPD approach for increased statistical precision. We also analyzed the prevalence of any pathogen and any MST marker separately for each sample type (e.g., water, hand rinses, soil, flies). Secondary outcomes included the prevalence of specific pathogen class (any viruses, any bacteria, any protozoa, any helminths), the prevalence of MST markers from specific host types (human, animal), and the prevalence and abundance of individual enteropathogens and MST markers. We excluded general MST markers that are not host-specific from our analysis.

We compared outcomes between the intervention and control arms of each study. We estimated prevalence ratios (PRs) using modified Poisson regressions.[22] For abundance outcomes, we used linear regressions to estimate differences in $\log_{10}$-transformed gene copies and negative binomial regressions to estimate ratios of soil-transmitted helminth (STH) egg counts. Because of repeated sampling or clustered designs in some studies, we used the Huber Sandwich Estimator to calculate robust standard errors.[23] For abundance measures, we imputed values below the limit of detection (LOD) with half the LOD and values below the limit of quantification (LOQ) with the midpoint between the LOD and LOQ. We limited our analysis of abundance measures to targets where >50% of samples were within the range of quantification (ROQ).

All analyses were adjusted for potential confounders and treatment arm. While estimated effects from randomized trials should be unconfounded, covariate adjustment may increase statistical efficiency and improve exchangeability with matched cohorts and non-randomized trials.[24] We prescreened covariates using likelihood ratio tests and included those associated with the outcome with a p-value <0·2 in the model for each outcome. We prescreened the following variables if they were measured within an included study: number of people in household, age and education of primary caregiver, 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. These variables reflect socio-demographic conditions that are commonly considered potential confounders in WASH studies. When analyzing binary outcomes, we only included one potential confounder per 10 positive samples, or per 10 negative samples if <50% of samples were negative. We did not estimate prevalence ratios for targets with less than five positive/negative values for a given sample type.

Given the heterogeneity across studies (e.g., local WASH conditions, climate, urbanization, population density, regional infectious disease patterns, intervention type), we individually estimated study-specific effects. For outcomes where data were available from four or more studies, we pooled using random-effects models fit using restricted-maximum likelihood with the metafor package.[25]. We did not pool abundance estimates because of issues in standardizing qPCR methods across sites and the small number of available abundance estimates.[26,27]

We conducted subgroup analyses by season (dry vs. wet), animal ownership (at least one vs. no animal owned) and pathogens with vs. without zoonotic transmission. The wet season for each study was defined as the six months of highest country-level average rainfall.[28] The pathogens we considered as potentially zoonotic were *Campylobacter jejuni/coli, Salmonella, Yersinia enterocolitica, Clostridium difficile, Cryptosporidium, Giardia* and *Ascaris*.[29] We classified *Ascaris* as potentially zoonotic because *Ascaris lumbricoides* and *Ascaris suum* cross-infect humans and pigs, and the microscopy methods used in the studies in our review do not distinguish between them .[29–31] When studies detected virulence genes associated with specific *E. coli* pathotypes (EAEC, EPEC/EHEC, STEC, EIEC, ETEC), we classified STEC and EPEC (due to atypical EPEC) as zoonotic.[29] We used linear regression models estimating prevalence differences to assess additive interaction by examining the p-values on the interaction terms between the treatment and the indicator variable for the subgroup; additive interaction has been argued to better capture public health importance than multiplicative interaction.[32] A conservative p-value cut-off of <0·2 on the interaction term was considered evidence of effect modification because of the low power of interaction analyses.[33]

We assessed heterogeneity by study-level characteristics, including setting, study design, intervention uptake and time between intervention onset and environmental sampling. There was limited heterogeneity in urbanicity within any individual study. Therefore, we pooled estimates separately for rural vs. urban studies. We also separately pooled estimates from randomized vs. quasi-experimental

studies, studies with high vs. low intervention uptake and studies with shorter (≤ 1 year) vs. longer (>1 year) follow-up between intervention onset and sampling. We compared pooled estimates between strata with Wald tests.

Analyses were conducted in R 4·0·4. Analysis scripts are publicly available (https://github.com/amertens/wash-ipd). Our systematic review search strategy and analysis plan were pre-registered on Open Science Framework (https://osf.io/8sgzn/).

Results

Search results and data acquisition

The systematic review was conducted on 19 January 2021 and returned 3,376 results after removing duplicates. Of these, 3,253 were excluded by abstract screening, and of 123 short-listed studies, nine were eligible after full-text screening. The nine articles reported findings from six unique intervention studies: the WASH Benefits Bangladesh (WBB) and Kenya (WBK) trials,[34] the Maputo Sanitation (MapSan) study in Mozambique,[35] the Gram Vikas study in India,[36] the Total Sanitation Campaign (TSC) trial in India,[6] and the CHoBI7 trial in Bangladesh[37] (Table 1). Data were obtained from all studies except CHoBI7 where individual participant data were not shared; this trial was excluded from our analysis. For Mapsan, additional data from an unpublished analysis were also shared. For the TSC trial, only village-level source water quality data were available. For WBB and Mapsan, multiple substudies within the trials collected samples from different subsets of participants at different times; therefore, we report the results of individual publications separately rather than combined by trial.

Characteristics of included studies

Three studies were cluster-randomized controlled trials (WBB, WBK, TSC). MapSan was a controlled before-and-after study with control and intervention sites matched on compound size and time of enrollment. Gram Vikas was a matched cohort study where control and intervention villages were matched on 12 pre-intervention WASH and socio-economic characteristics. Using the Newcastle-Ottawa scale, studies had low risk of bias due to blinded outcome assessments, with the Gram Vikas and MapSan studies having a lower rating due to higher loss to follow-up and lack of randomization (Table S4). WBB, WBK, TSC and Gram Vikas were conducted in rural settings while MapSan was urban. All included studies evaluated sanitation interventions (Table 1). TSC and MapSan focused on sanitation alone. The WBB and WBK trials included individual and combined water, hygiene, sanitation, and nutrition interventions but pathogens and MST markers in environmental samples were only measured in the sanitation and control arms. The Gram Vikas study evaluated a combined piped drinking water and sanitation intervention. No included studies evaluated drinking water supply/treatment or hygiene interventions alone.

All sanitation interventions evaluated were onsite (i.e. non-sewered) technologies delivered at the household or community level. None of the interventions met the Sustainable Development Goal standard of “safely managed sanitation” and would be classified as “basic” or “limited” sanitation. The WASH Benefits studies provided new or upgraded improved latrines for each household in enrolled compounds, child potties and sani-scoops for feces removal. WBB latrines were dual-pit latrines with a water seal and in WBK plastic latrine slabs were used to improve existing latrines. MapSan provided pour-flush latrines draining to septic tanks, shared by multiple households. TSC promoted construction of a pour-flush latrine with a single pit and Y-joint for a second pit, subsidized post-hoc by government funding. In the Gram Vikas study, a non-governmental organization provided materials for the construction of pour-flush latrines in each household in selected villages and built community water tanks and piped distribution systems providing household connections. When every household in the village completed latrine construction, the water system was turned on for the whole village.

Latrine access and use was higher in intervention households than control households in all studies. Definitions of latrine quality varied, including improved/clean/hygienic/functional latrines or latrines with a functional water seal, as observed by field staff. In four studies, 78-97% of intervention recipients had access to these types of facilities, compared to 18-45% of controls.[3,4,36,38] The TSC trial had the lowest effect on latrine access, with 38% of intervention compounds having functional latrines compared to 10% of controls.[6] Latrine use in intervention households was variable and especially low among children, and safe management of child and animal feces was uncommon. In WBB, 94% of adults were observed to

defecate in a hygienic latrine in structured observations but only 54% of children were observed using the latrine or potty and 15% of animal feces were observed to be removed with the sani-scoop.[39] In WBK, reported safe disposal of child feces dropped from 77% one year after intervention to 37% after two years.[4] In TSC, 50% of households reported children using a latrine,[6] and in Gram Vikas, 35% of intervention villages reported disposing of child feces in improved latrines.[36]

Environmental sample types and targets

Environmental samples were collected from 4 months[40] to 6-10 years[36] after intervention delivery, with most studies collecting samples 1-2 years post-intervention (Table 1). Sample types included source and stored drinking water, child and mother hand rinses, soil from the courtyard, household and latrine areas, and flies caught in latrines and kitchens. Food samples were collected in one study[41] but were not included in our analysis because only 9 samples were positive for MST targets. The number of samples in individual studies varied from 60[42] to 2107[36]. Our pooled dataset included 12,184 samples, with 40,156 observations for pathogen or MST marker prevalence.

The studies measured a range of bacterial, viral, protozoan, and helminthic pathogens, including pathogenic *E. coli, Vibrio cholerae, Shigella, Campylobacter jejuni/coli, Salmonella, Yersinia, Clostridium difficile*, rotavirus, norovirus, sapovirus, adenovirus, astrovirus, enterovirus, *Cryptosporidium, Giardia, Entamoeba histolytica, Ascaris lumbricoides and Trichuris trichiura* (Tables S5-S8). The MST markers included human (HumM2, HF183, BacHum, *M. smithii*), animal (BacCan, BacCow), ruminant (BacR) and avian (GFD) fecal markers (Tables S5-S8). Most studies used quantitative polymerase chain reaction (qPCR) or reverse-transcriptase (RT)-qPCR (Table 1). One study used slide agglutination serotyping to detect *V. cholerae* and *Shigella*.[43] One study detected *Cryptosporidium* oocysts and *Giardia* cysts using direct fluorescent antibody microscopy.[42] Two studies enumerated STH eggs by microscopy.[44,45]

Many targets had low or no variation. Out of 267 unique combinations of study, sample type, and target, 18 had no positive values, 41 had less than ten positive values and two had less than ten negative values. Therefore, 206/267 sample-target combinations had sufficient variability to estimate a PR and be individually included in our IPD analysis. Among these, pathogen prevalence ranged from 1·4% for *Giardia* on mothers' hands[46] to 62·1% for *Ascaris* in soil[44] and the prevalence of MST markers ranged from 2·4% for HumM2 on child hands [40] to 97·5% for BacCow on mothers' hands.[46]

Effects on the prevalence of any pathogens and any MST markers

Interventions decreased the prevalence of any pathogen in any sample type in most individual studies but confidence intervals for PRs often crossed the null (Figure 1). Among individual sample types, pathogen prevalence was significantly reduced in flies (adjusted PR [aPR]=0·37 (95% CI: 0·16, 0·85), Figure 1). Study-specific estimates were largely homogeneous, with no significant Cochran's Q-tests of homogeneity. Pooled across studies, there was a small reduction in the prevalence of any pathogen detected in any sample type (pooled aPR=0·94 (95% CI: 0·90, 0·99), Figure 1). Most studies showed no effect on the detection of any MST marker (Figure 2). There was a reduction in any MST marker in water samples in Boehm et al. 2016 (0·69 (95% CI: 0·50, 0·95)) but an increase in any sample type in Capone et al. 2022 (aPR=1·16 (95% CI: 1·02, 1·32), Figure 2). When pooled, interventions had no effects on the prevalence of any MST marker in any sample type (pooled aPR= 1·01 (95% CI: 0·98, 1·04)) or within specific sample types (Figure 2).

Effects on the prevalence of pathogen classes and specific pathogens

Interventions reduced the prevalence of any bacterial pathogens in any sample type (pooled aPR=0·92 (95% CI: 0·85, 0·99)), though intervention effects were not significant in any individual study (Figure 1). Interventions did not significantly reduce virus prevalence in any sample type (pooled aPR= 0·90 (95% CI: 0·62, 1·33) or within specific sample types (Figure 1). Intervention effects in individual studies were generally in the protective direction for protozoa and helminths but we did not have sufficient studies to pool estimates. Among specific pathogens, interventions reduced the prevalence of adenovirus (aPR=0·21 (95% CI: 0·06, 0·68)) and *Shigella* (aPR=0·28 (95% CI: 0·10, 0·78)) in any sample type in Capone et al. 2021, driven by significant reductions in soil around latrines (Figure S2).

Effects on the prevalence of MST marker types and individual markers

Interventions effects were inconsistent and largely null for both human markers (pooled aPR: 1·00 (95% CI: 0·88, 1·13)) and animal markers (pooled aPR: 1·00 (95% CI: 0·97, 1·03)) (Figure 2). There was a

reduction in any animal marker in stored water in Boehm et al. 2016 (aPR=0·69 (95% CI: 0·50, 0·95), Figure 2), driven by a significant reduction in the ruminant BacR marker (aPR= 0·62 (95% CI: 0·43, 0·90), Figure S3). There was a reduction in the human marker HF183 in any sample type (aPR=0·67 (95% CI: 0·48, 0·95), Figure S2) in Holcomb et al 2020, but not individual sample types. There were no other intervention effects on individual MST markers (Figure S3).

Effects on the abundance of specific pathogens and MST markers

Of all observations, 20% had abundances quantified, including STH egg counts in Steinbaum et al. 2019 and Kwong et al. 2021 and gene copies of enteropathogens and MST targets in Boehm et al. 2016, Fuhrmeister et al. 2020, and Capone et al. 2022 (Figure S4). Of these, 18% were below the specific study-reported LOD, 24% below the study-reported LOQ, and 58% within the study-reported ROQ. Of targets enumerated within specific sample types, only 18% had >50% of samples within the ROQ and were therefore included in our analysis. The abundance of the BacCow animal marker was lower in mothers' hand rinses in the sanitation intervention arm in Fuhrmeister et al. 2020, with an adjusted $\log_{10}$-transformed difference of -0·28 (95% CI: -0·49, -0·07) per pair of hands (Table 2). The interventions did not have significant effects on the abundance of any other MST target, nor STH egg counts (Table 2).

Subgroup and adjusted analyses

Intervention effects differed by season, but the direction of effects was inconsistent (Figure S5). There was no significant effect of any interventions on any pathogen or any MST marker prevalence when households were stratified by animal presence (Figure S6). There were no differences in intervention effects on pathogens with possible zoonotic transmission versus only human hosts (Figure S7). In Wald tests, there were no significant differences in pooled estimates between the one urban study (MapSan) and the four rural studies (p-value: 0·25), between randomized and quasi-experimental studies (p-value: 0·43), between studies with ≤1 year and >1 year of follow-up (p-value: 0·51) or between the four studies with high latrine access among intervention recipients compared to the TSC trial with lower access (p-value: 0·57). Adjustment covariates were measured differently across studies (Table S9). Unadjusted and adjusted estimates were similar (Figures S8-S9).

Discussion

Our IPD analysis of five intervention studies, mostly focused on household- and community-level onsite sanitation improvements, indicates a small overall reduction in pathogen prevalence in the environment associated with the interventions. While individual studies were underpowered to detect effects on pathogen prevalence with precision, point estimates of intervention effects were consistently in the protective direction across studies, despite differences in setting, intervention design, and length of follow-up. There were no overall effects on human or animal faecal markers.

These findings add to a body of literature on the effectiveness of sanitation improvements in low-income countries in interrupting faecal-oral transmission. A previous systematic review found no effect of sanitation interventions on FIB in the environment.[7] The small pooled effect on pathogens in the environment in our analysis indicates that any reductions in pathogen transmission through environmental pathways was likely small. This can help explain the null findings of the parent trials on child diarrhoea.[3,4,6,35,36] Among the five included studies, only WASH Benefits Bangladesh found a significant reduction in diarrhoea[3] as well as a reduction in parasite infections[47,48] in the sanitation arm compared to controls.[49] Diarrhoea was reduced by 2·2 percentage points on the absolute scale, compatible with a small reduction in pathogen transmission. Taken together, these findings indicate that the sanitation interventions in the studies in our review did not sufficiently isolate faecal waste from the environment, despite most of them achieving high levels of latrine access and use by adults. Young children's feces are a dominant source of faecal contamination in the household environment[49], while animal feces make up the majority of global faecal waste[50] and are associated with increased domestic contamination.[51] Therefore, containment of adult human waste may be insufficient to reduce environmental contamination in settings with continued child open defecation and high exposure to animal waste.[52] Only the WASH Benefits trials included tools for child feces management (potties and scoops), and the scoops could also be used to dispose of animal feces but adoption of these tools was low.[39,53]. Notably, we found reduced prevalence of ruminant (BacR) markers in stored water and reduced abundance of animal markers (BacCow) on mothers' hands in two studies nested within WASH Benefits

Bangladesh. The reduction in animal faecal contamination can help explain the unique health impacts in this trial. In our analysis, only the MapSan study achieved a reduction in a human (HF183) marker. More comprehensive sanitation programs, such as safely managed sanitation services that include safe removal in addition to containment of faecal waste, and interventions targeting child and animal feces can potentially more effectively interrupt environmental pathogen transmission.[54]

It is possible that current environmental measures have limited ability to detect intervention effects on pathogen presence in the environment. Faecal contamination in the domestic environment varies spatially and temporally[55,56], and pathogen presence in the environment is intermittent, depending on the presence of infected individuals, shedding rates and pathogen fate and survival in environmental reservoirs[57]. Different pathogens have different predominant transmission pathways, and specific pathogens may cause illness through a particular pathway too infrequently to capture with cross-sectional grab samples. Additionally, pathogen prevalence and abundance in the environment is typically low[57], leading to low statistical power to detect intervention effects. Any reductions in pathogen presence might be more apparent with larger sample sizes and/or repeated sampling with high temporal and spatial resolution, which is costly for currently available pathogen detection methods, or by analyzing larger quantities of composite samples. In addition, human MST markers have low specificity and sensitivity in settings with widespread faecal contamination in the environment.[27,58] Also, molecular methods for pathogen detection do not provide information on viability, and the clinical implications of small amounts of pathogen DNA/RNA detected in a sample are unclear. While FIB have limitations in terms of low specificity to faecal sources and poor correlation with pathogens, culture-based FIB enumeration captures viable organisms, and large numbers of temporal/spatial samples can be analyzed at low cost. Therefore, studies evaluating the environmental impact of WASH interventions can benefit from combining molecular pathogen measurements with culture-based FIB measurements to leverage the respective strengths of these approaches. Pathogen-specific testing can supplement FIB data to identify the specific etiologies through which WASH interventions improve health or the effects of targeted interventions on specific pathogens. Advances in technology that reduce the costs of molecular diagnostics or increased funding for environmental testing within WASH trials may allow broader use of pathogen detection methods to more precisely estimate intervention effects on environmental contamination.

Our analysis had some strengths and limitations. The IPD meta-analysis allowed us to pool data with standardized estimation approaches across studies.[59] The individual studies in our review were designed and powered to detect effects on child health rather than infrequently detected pathogens or MST markers in the environment. Pooling increased our statistical power for rare outcomes to detect a small overall effect on pathogens that individual studies were underpowered to detect. However, pooling assumes that individual studies are sufficiently homogeneous despite implementing different interventions in different settings. Pooled estimates should therefore be interpreted in conjunction with estimates from individual studies. In our analysis, we detected no statistical heterogeneity between studies, and low-precision estimates from individual studies were qualitatively aligned with high-precision pooled estimates, suggesting that pooling data did not obscure any study-specific trends. Similarly, because studies measured different targets in different environmental matrices and many targets were detected infrequently, we relied on composite measures, such as detection of any pathogen/MST marker in any sample type, to pool data across studies. This highlights both a strength and limitation of pathogen detection in the environment. While measuring pathogens directly provides high specificity and avoids the false positives associated with FIB, the low prevalence of a given pathogen along a given pathway results in low statistical power. Combining data on different pathogens along different pathways to increase power can provide a general understanding of intervention impacts but obscures nuances on which specific pathogens along which pathways are influenced by interventions. Therefore, effects on these composite outcomes should be interpreted in tandem with pathogen-specific estimates. Standardized measurement and reporting of a harmonized panel of enteropathogens in a consistent set of environmental matrices can allow better comparability of pathogen-specific data for future IPD meta-analyses.[15]

Only a small number of studies met our inclusion criteria, limiting the generalizability of our findings. Four of the included studies focused on onsite sanitation and one evaluated a combined piped water and sanitation intervention. Therefore, we were unable to explore the effects of individual water supply/treatment and hygiene interventions, and more comprehensive sanitation modalities such safely

managed sanitation services and sewer connections. Also, while the studies tested a diverse set of sample types, including understudied reservoirs such as soil, not all pathways were captured. For example, contaminated food has been identified as a dominant pathogen transmission pathway[60] but only one study in our review sampled food and we could not include these data in our analysis as the target was infrequently detected.

The basic sanitation interventions in our review resulted in a small reduction in the environmental presence of enteropathogens, consistent with the previously reported limited health impacts. Our results suggest that these sanitation interventions failed to contain human waste and thus prevent exposure to enteropathogens in these populations. More comprehensive approaches are needed to catalyze major health gains. Countries which have achieved universal access to effective sanitation have seen dramatic improvements in health.[61] Public health programs in low-income countries should pursue "transformative WASH" approaches that encompass the full chain of excreta management including safe removal rather than mere containment and address child and animal feces to more effectively interrupt environmental pathogen transmission. Also, our review identified no water supply, water quality and hygiene trials that measured pathogens in the environment. Future studies should assess the effect of such interventions on environmental contamination, using a combination of pathogen measurements and FIB proxies, and including understudied pathways such as soil, food, and flies.

References

1 Troeger C, Blacker BF, Khalil IA, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: A systematic analysis for the Global Burden of Disease Study 2016. The Lancet Infectious Diseases 2018; 18: 1211–28.

2 Humphrey JH, Mbuya MNN, Ntozini R, 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 2019; 7: e132–47.

3 Luby SP, Rahman M, Arnold BF, 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 2018; 6: e302–15.

4 Null C, Stewart CP, Pickering AJ, 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 2018; 6: e316–29.

5 Knee J, Sumner T, Adriano Z, et al. Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-after trial. eLife 2021; 10: e62278.

6 Clasen T, Boisson S, Routray P, 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 2014; 2: e645–653.

7 Sclar GD, Penakalapati G, Amato HK, et al. Assessing the impact of sanitation on indicators of fecal exposure along principal transmission pathways: A systematic review. International Journal of Hygiene and Environmental Health 2016; 219: 709–23.

8 Arnold BF, Colford JM. Treating water with chlorine at point-of-use to improve water quality and reduce child diarrhea in developing countries: A systematic review and meta-analysis. The American journal of tropical medicine and hygiene 2007; 76: 354–64.

9 Luby SP, Agboatwalla M, Raza A, et al. Microbiologic effectiveness of hand washing with soap in an urban squatter settlement, Karachi, Pakistan. Epidemiology and Infection 2001; 127: 237–44.

10 Pickering AJ, Boehm AB, Mwanjali M, Davis J. Efficacy of waterless hand hygiene compared with handwashing with soap: A field study in Dar es Salaam, Tanzania. The American Journal of Tropical Medicine and Hygiene 2010; 82: 270–8.

11 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in household drinking water: Systematic review and meta-analysis. PloS One 2014; 9: e107429.

12 Devane ML, Moriarty E, Weaver L, Cookson A, Gilpin B. Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. Water Research 2020; 185: 116204.

13 Wu J, Long SC, Das D, Dorner SM. Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research. Journal of Water and Health 2011; 9: 265–78.

14 Hardina CM, Fujioka RS. Soil: The environmental source of Escherichia coli and enterococci in Hawaii's streams. Environmental Toxicology and Water Quality 1991; 6: 185–95.

15 Lappan R, Henry R, Chown SL, et al. Monitoring of diverse enteric pathogens across environmental and host reservoirs with TaqMan array cards and standard qPCR: A methodological comparison study. The Lancet Planetary Health 2021; 5: e297–308.

16 Liu J, Gratz J, Amour C, et al. A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. Journal of Clinical Microbiology 2013; 51: 472–80.

17 Stoeckel DM, Harwood VJ. Performance, Design, and Analysis in Microbial Source Tracking Studies. Applied and Environmental Microbiology 2007; 73: 2405–15.

18 Wolf J, Hunter PR, Freeman MC, et al. Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: Updated meta-analysis and meta-regression. Tropical Medicine & International Health 2018; 23: 508–25.

19 Wolf J, Prüss-Ustün A, Cumming O, et al. Systematic review: Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression. Tropical Medicine & International Health 2014; 19: 928–42.

20 Mertens A, 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. 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. (In preparation).

21 Wells S G. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute.

22 Zou G. A modified poisson regression approach to prospective studies with binary data. American Journal of Epidemiology 2004; 159: 702–6.

23 Freedman DA. On The So-Called 'Huber Sandwich Estimator' and 'Robust Standard Errors'. The American Statistician 2006; 60: 299–302.

24 Cochran WG. The Combination of Estimates from Different Experiments. Biometrics 1954; 10: 101–29.

25 Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software 2010; 36: 1–48.

26 Borchardt MA, Boehm AB, Salit M, Spencer SK, Wigginton KR, Noble RT. The Environmental Microbiology Minimum Information (EMMI) Guidelines: qPCR and dPCR Quality and Reporting for Environmental Microbiology. Environmental Science & Technology 2021; 55: 10210–23.

27 Boehm AB, Van De Werfhorst LC, Griffith JF, et al. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. Water Research 2013; 47: 6812–28.

28 d.o.o YMG. Weather Atlas | Weather forecast and Climate information for cities all over the Globe. Weather Atlas.

29 Delahoy MJ, Wodnik B, McAliley L, et al. Pathogens transmitted in animal feces in low- and middle-income countries. International Journal of Hygiene and Environmental Health 2018; 221: 661–76.

30 Betson M, Nejsum P, Bendall RP, Deb RM, Stothard JR. Molecular epidemiology of ascariasis: A global perspective on the transmission dynamics of Ascaris in people and pigs. The Journal of Infectious Diseases 2014; 210: 932–41.

31 Steinbaum L, Kwong LH, Ercumen A, et al. Detecting and enumerating soil-transmitted helminth eggs in soil: New method development and results from field testing in Kenya and Bangladesh. PLOS Neglected Tropical Diseases 2017; 11: e0005522.

32 VanderWeele TJ, Knol MJ. A Tutorial on Interaction. Epidemiologic Methods 2014; 3: 33–72.

33 Selvin S. Statistical Analysis of Epidemiologic Data, Third. New York: Oxford University Press, 2004 DOI:10.1093/acprof:oso/9780195172805.001.0001.

34 Arnold BF, Null C, Luby SP, 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 2013; 3: e003476.

35 Knee JS. Quantifying the impact of an urban onsite shared sanitation intervention on child health in Maputo, Mozambique: The Mapsan trial. 2019; published online July.

36 Reese H, Routray P, Torondel B, 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 2019; 48: 1757–67.

37 Burrowes V, Perin J, Monira S, et al. Risk Factors for Household Transmission of Vibrio cholerae in Dhaka, Bangladesh (CHoBI7 Trial). The American Journal of Tropical Medicine and Hygiene 2017; 96: 1382–7.

38 Bick S, Buxton H, Chase RP, et al. Using path analysis to test theory of change: A quantitative process evaluation of the MapSan trial. BMC Public Health 2021; 21: 1411.

39 Parvez SM, Azad R, Rahman M, et al. Achieving optimal technology and behavioral uptake of single and combined interventions of water, sanitation hygiene and nutrition, in an efficacy trial (WASH benefits) in rural Bangladesh. Trials 2018; 19: 358.

40 Boehm AB, Wang D, Ercumen A, et al. Occurrence of Host-Associated Fecal Markers on Child Hands, Household Soil, and Drinking Water in Rural Bangladeshi Households. Environmental Science & Technology Letters 2016; 3: 393–8.

41 Holcomb DA, Knee J, Capone D, et al. Impacts of an Urban Sanitation Intervention on Fecal Indicators and the Prevalence of Human Fecal Contamination in Mozambique. Environmental Science & Technology 2021; 55: 11667–79.

42 Odagiri M, Schriewer A, Daniels ME, et al. Human fecal and pathogen exposure pathways in rural Indian villages and the effect of increased latrine coverage. Water Research 2016; 100: 232–44.

43 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. 2017.

44 Kwong LH, Sen D, Islam S, et al. Effect of sanitation improvements on soil-transmitted helminth eggs in courtyard soil from rural Bangladesh: Evidence from a cluster-randomized controlled trial. PLOS Neglected Tropical Diseases 2021; 15: e0008815.

45 Steinbaum L, Mboya J, Mahoney R, Njenga SM, Null C, Pickering AJ. 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 2019; 13: e0007180.

46 Fuhrmeister ER, Ercumen A, Pickering AJ, et al. Effect of Sanitation Improvements on Pathogens and Microbial Source Tracking Markers in the Rural Bangladeshi Household Environment. Environmental Science & Technology 2020; 54: 4316–26.

47 Lin A, Ercumen A, Benjamin-Chung J, et al. Effects of Water, Sanitation, Handwashing, and Nutritional Interventions on Child Enteric Protozoan Infections in Rural Bangladesh: A Cluster-Randomized Controlled Trial. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 2018; 67: 1515–22.

48 Ercumen A, Benjamin-Chung J, Arnold BF, et al. Effects of water, sanitation, handwashing and nutritional interventions on soil-transmitted helminth infections in young children: A cluster-randomized controlled trial in rural Bangladesh. PLOS Neglected Tropical Diseases 2019; 13: e0007323.

49 Bauza V, Madadi V, Ocharo RM, Nguyen TH, Guest JS. Microbial source tracking using 16S rRNA amplicon sequencing identifies evidence of widespread contamination from young children's feces in an Urban Slum of Nairobi, Kenya. Environmental Science & Technology 2019; 53: 8271–81.

50 Berendes DM, Yang PJ, Lai A, Hu D, Brown J. Estimation of global recoverable human and animal faecal biomass. Nature Sustainability 2018; 1: 679–85.

51 Ercumen A, Pickering AJ, Kwong LH, et al. Animal Feces Contribute to Domestic Fecal Contamination: Evidence from E. Coli Measured in Water, Hands, Food, Flies, and Soil in Bangladesh. Environmental Science & Technology 2017; 51: 8725–34.

52 Pickering AJ, Null C, Winch PJ, et al. The WASH Benefits and SHINE trials: Interpretation of WASH intervention effects on linear growth and diarrhoea. The Lancet Global Health 2019; 7: e1139–46.

53 Islam M, Rahman M, et al. Assessing sustained uptake of latrine and child feces management interventions in rural Bangladesh. PloS One 2022; In review.

54 Amebelu A, Ban R, Bhagwan J, et al. The Lancet Commission on water, sanitation and hygiene, and health. The Lancet 2021; 398: 1469–70.

55 Levy K, Hubbard AE, Nelson KL, Eisenberg JNS. Drivers of Water Quality Variability in Northern Coastal Ecuador. Environmental science & technology 2009; 43: 1788–97.

56 Ram PK, Jahid I, Halder AK, et al. Variability in hand contamination based on serial measurements: Implications for assessment of hand-cleansing behavior and disease risk. The American Journal of Tropical Medicine and Hygiene 2011; 84: 510–6.

57 Goddard FGB, Pickering AJ, Ercumen A, Brown J, Chang HH, Clasen T. Faecal contamination of the environment and child health: A systematic review and individual participant data meta-analysis. The Lancet Planetary Health 2020; 4: e405–15.

58 Payment P, Locas A. Pathogens in water: Value and limits of correlation with microbial indicators. Ground Water 2011 Jan-Feb; 49: 4–11.

59 Holcomb DA, Stewart JR. Microbial Indicators of Fecal Pollution: Recent Progress and Challenges in Assessing Water Quality. Current Environmental Health Reports 2020; 7: 311–24.

60 Hernán MA. Causal analyses of existing databases: No power calculations required. Journal of Clinical Epidemiology 2021; : S0895-4356(21)00273-0.

61 Amin N, Rahman M, Raj S, et al. Quantitative assessment of fecal contamination in multiple environmental sample types in urban communities in Dhaka, Bangladesh using SaniPath microbial approach. PLOS ONE 2019; 14: e0221193.

62 Mackenbach JP. Sanitation: Pragmatism works. BMJ 2007; 334: s17–7.

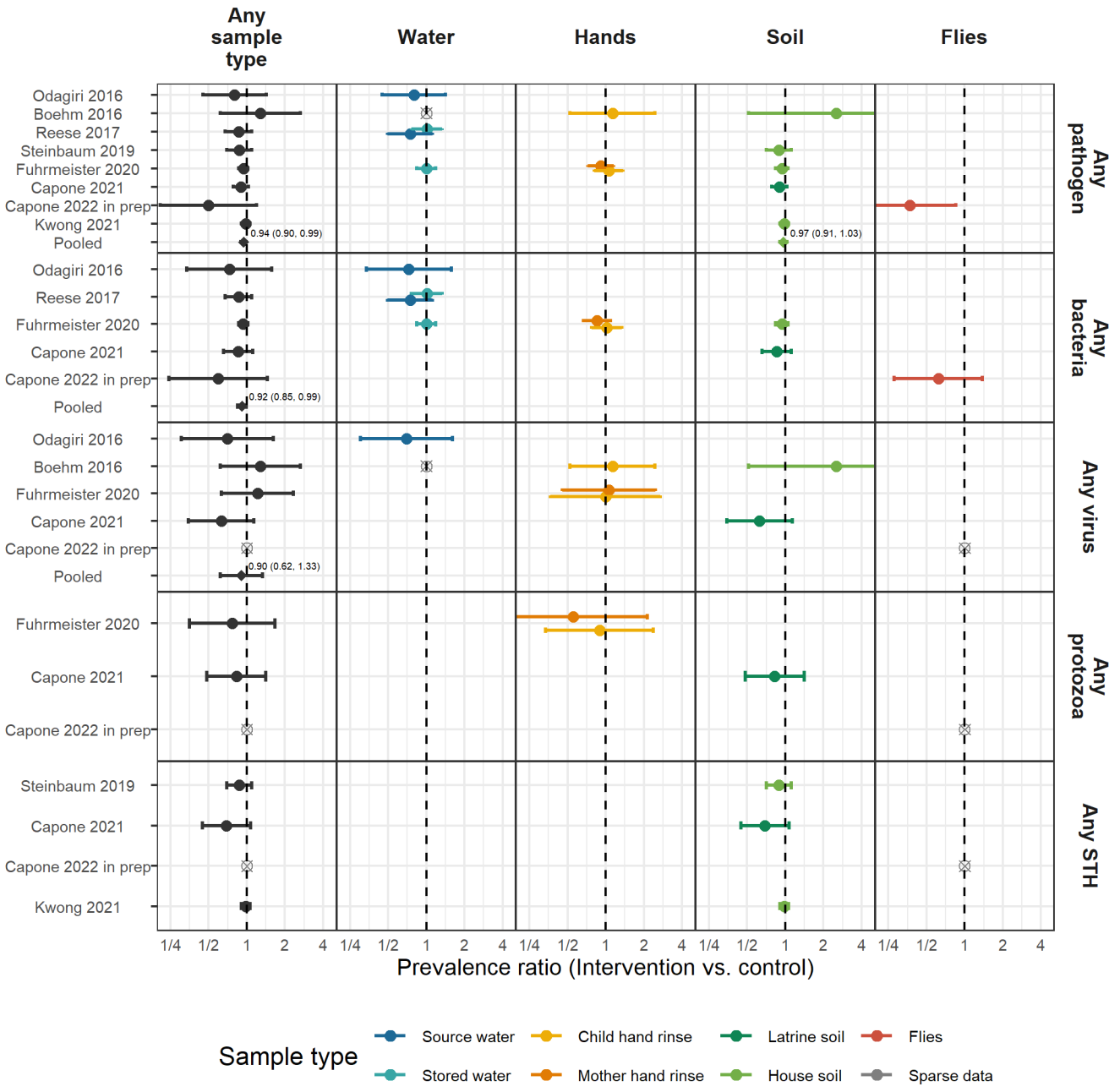


Figure 1. Forest plots of WASH intervention effects on the prevalence of any enteropathogen or type of enteropathogen (any bacteria, any virus, any protozoa and any STH) in different types of environmental samples. Pooled estimates are presented when there are four or more study-specific estimates for a specific sample type and target combination and are denoted with diamond-shaped points. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations). Samples of the same type from different locations (source vs. stored water, flies in kitchen vs. latrine, soil from courtyard vs. latrine) or different individuals (child vs. mother's hands) are plotted separately. Point estimates and confidence intervals are printed next to pooled estimates. All estimates are adjusted for potential confounders.

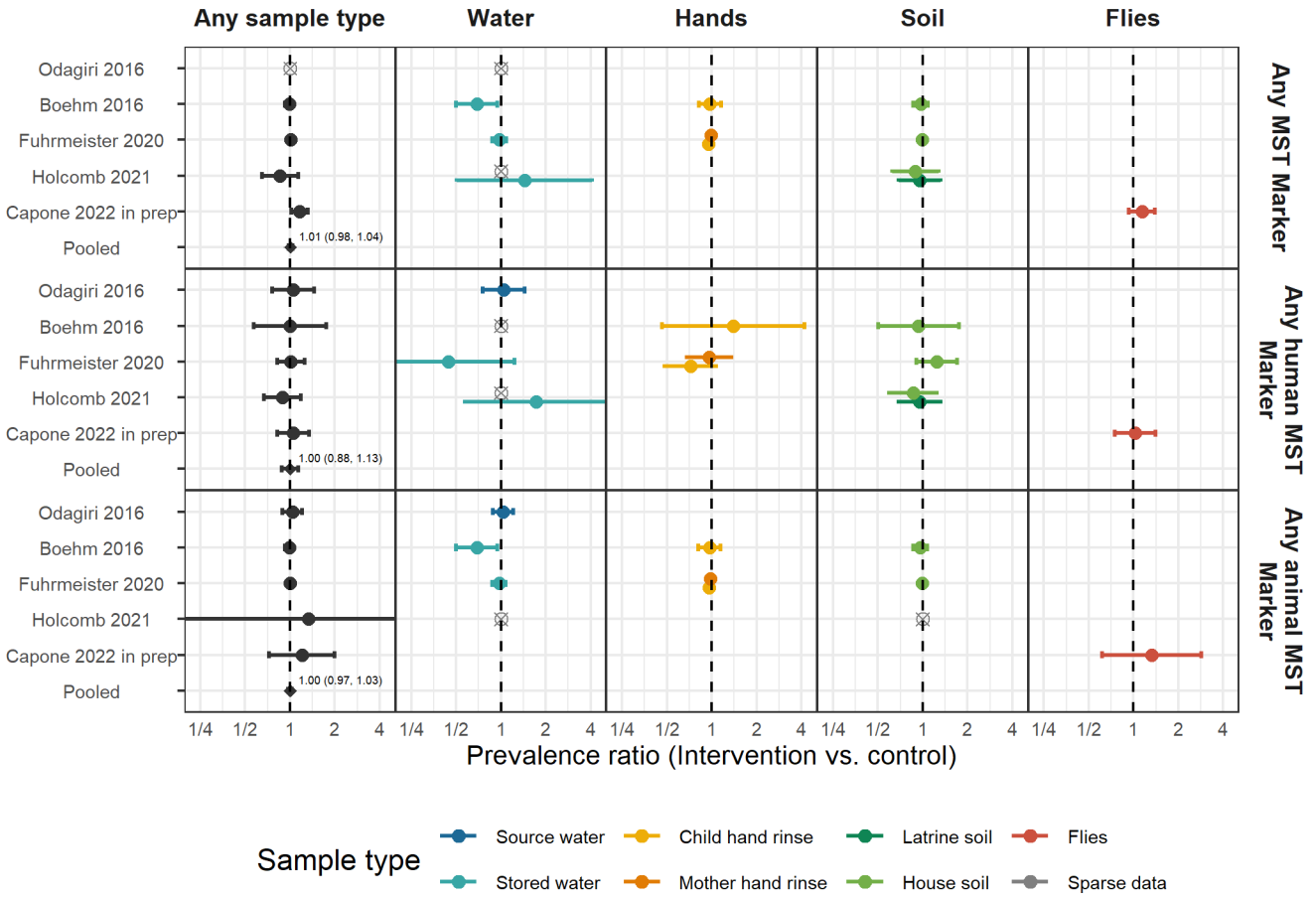


Figure 2. Forest plots of WASH intervention effects on the prevalence of any MST marker or type of MST marker (human or animal MST markers) in different types of environmental samples. Pooled estimates are presented when there are four or more study-specific estimates for a specific sample type and target combination and are denoted with diamond-shaped points. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations). Samples of the same type from different locations (source vs. stored water, flies in kitchen vs. latrine, soil from courtyard vs. latrine) or different individuals (child vs. mother's hands) are plotted separately. Point estimates and confidence intervals are printed next to pooled estimates. All estimates are adjusted for potential confounders.

Tables

Table 1. Characteristics of included publications

Parent study	Study design	Intervention	Time between intervention and environmental sampling	Location	Author/ year	Sample types	Targets	Analytic method	Number of samples
WASH Benefits Bangladesh	Cluster-randomized trial	Latrine upgrades, child potties, scoops for feces disposal	4 months	Rural Bangladesh	Boehm et al. 2016	Stored drinking water, child hands, soil	Rotavirus, General, human, avian and ruminant fecal markers	qPCR	1,482
-	-	-	16-35 months	-	Fuhrmeister et al. 2020	Stored drinking water, child and mother hands, soil	Pathogenic E. coli, norovirus, Giardia	qPCR	2,601
-	-	-	~2 years	-	Kwong et al. 2021	Courtyard soil	Soil-transmitted helminths	Microscopy	1,396
WASH Benefits Kenya	Cluster-randomized trial	Latrine upgrades, child potties, scoops for feces disposal	~2 years	Rural Kenya	Steinbaum et al. 2019	Courtyard soil	Soil-transmitted helminths	Microscopy	2,149
MapSan	Controlled before-and-after study	Latrine upgrades	~1 year	Urban Mozambique	Holcomb et al. 2020	Source and stored water, household and latrine soil, food	General, human and avian fecal MST markers	qPCR	353
-	-	-	~1 year	-	Capone et al. 2021	Household and latrine soil	Panel of 18 enteric pathogens	qPCR	88
-	-	-	~2 years	-	Capone et al. 2022 in prep.	Flies caught in latrine and kitchen	Panel of 16 enteric pathogens and MST markers	qPCR	86
Gram Vikas	Matched cohort study	Latrine upgrades, piped water	~6-10 years	Rural India	Reese et al. 2017	Source and stored water	V. cholerae, Shigella	Slide agglutination serotyping	3,452
Total Sanitation Campaign	Cluster-randomized trial	Latrine upgrades	~1 year	Rural India	Odagiri et al. 2016	Source water	V. cholerae, rotavirus, adenovirus,general, human, and animal fecal markers	qPCR, microscopy	60

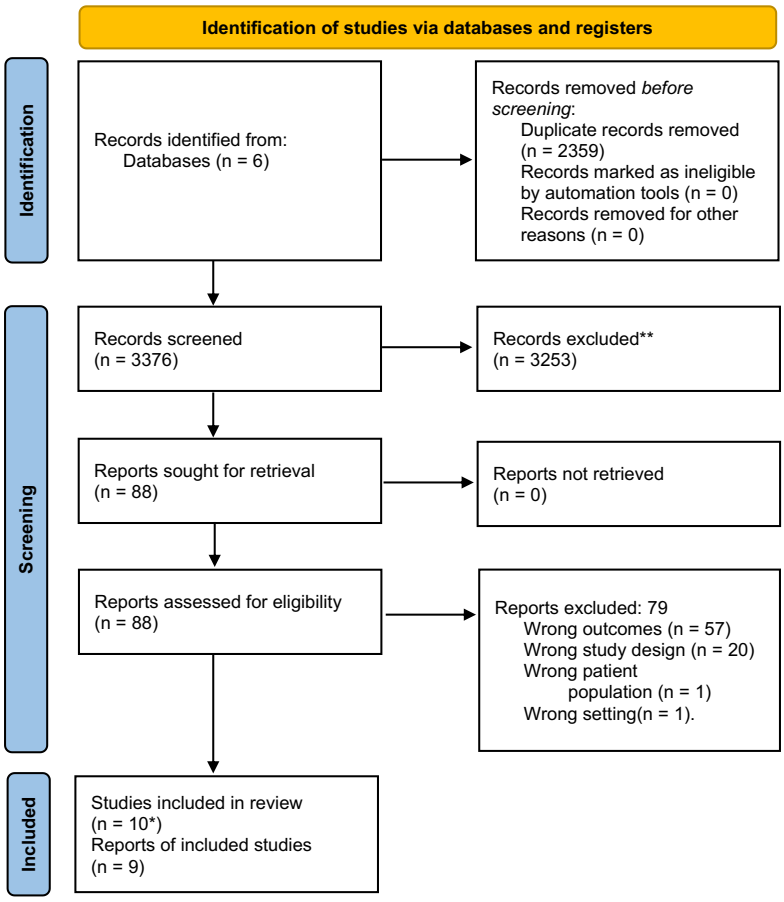
Table 2. Mean (SD) abundances of enteropathogen and MST targets by study arm. Means are log10-transformed gene copies for MST markers and mean egg counts for soil transmitted helminths (Ascaris and Trichuris). Intervention effects are shown as adjusted differences in log10-transformed gene copies and ratios of helminth egg counts between the intervention and control arms.

Study	Sample	Target	N	% in ROQ	Control mean, median (SD)	Intervention mean, median (SD)	Intervention effect (95% CI)	P value	Wilcoxon P value
Fuhrmeister 2020	Child hand rinse	Animal (BacCow)	365	75·9	3·6, 3·9 (1·4)	3·4, 3·8 (1·4)	-0·17 (-0·47 0·12)	0·25	0·17
-	Mother's hand rinse	Animal (BacCow)	725	66·5	3·3, 3·8 (1·4)	3, 3·7 (1·5)	-0·28 (-0·49 - 0·07)	0·01	0·01
Holcomb 2021	Latrine soil	Human (M. smithii)	113	51·3	6·7, 6·5 (0·6)	6·5, 6·3 (0·5)	-0·14 (-0·38 0·11)	0·27	0·58
Capone 2022 in prep		Human (BacHum)	173	77·5	3·8, 3·8 (1·3)	4, 4·2 (0·9)	0·14 (-0·19 0·47)	0·41	0·07
Steinbaum 2019	House soil	Ascaris	2,101	100·0	2·2, 0 (18·8)	1·4, 0 (9·3)	0·65 (0·33 1·28)[a]	0·21	0·33
-	-	Trichuris	2,102	100·0	0·2, 0 (1·8)	0·2, 0 (1)	0·73 (0·36 1·48)[a]	0·38	0·39
Kwong 2021	House soil	Ascaris	1,426	100·0	2·3, 0·7 (6·7)	2·2, 0·6 (6·9)	0·97 (0·68 1·38)[a]	0·85	0·54
-	-	Trichuris	1,426	100·0	1·6, 0·4 (5)	2, 0·4 (5)	1·22 (0·87 1·71)[a]	0·26	0·17

ROQ: Range of quantification; SD: Standard deviation; CI: Confidence interval; Wilcoxon P-value: Non-parametric Wilcoxon rank sum test P-value.

[a] Marks ratio estimates from negative binomial model

SUPPLEMENTARY FIGURES AND TABLES



*One unpublished and shared by authors of another included study

The systematic review was conducted on 1/19/2021.

Figure S1. PRISMA Flowchart.

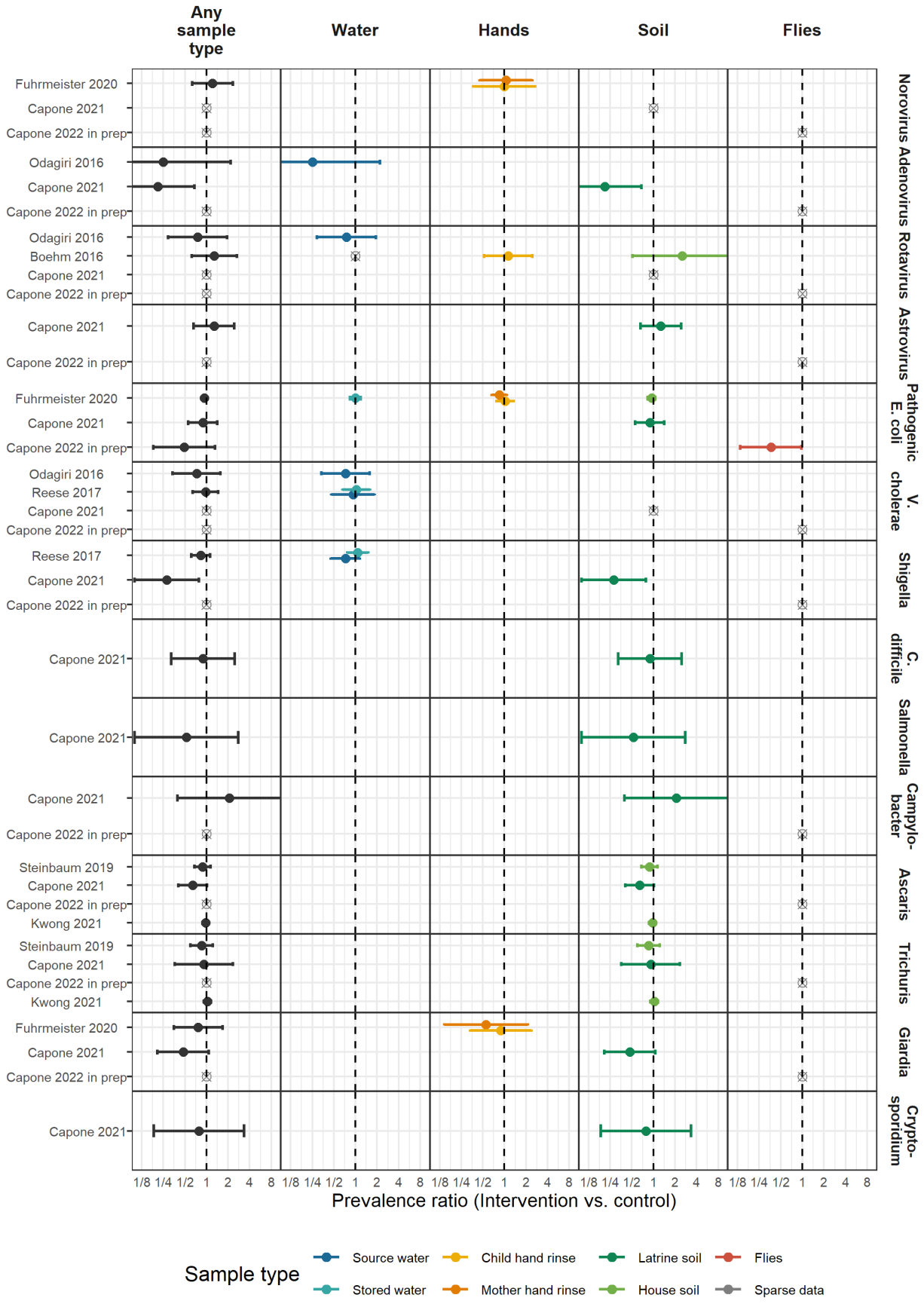


Figure S2. Forest plots of intervention effects on the prevalence of specific pathogens.

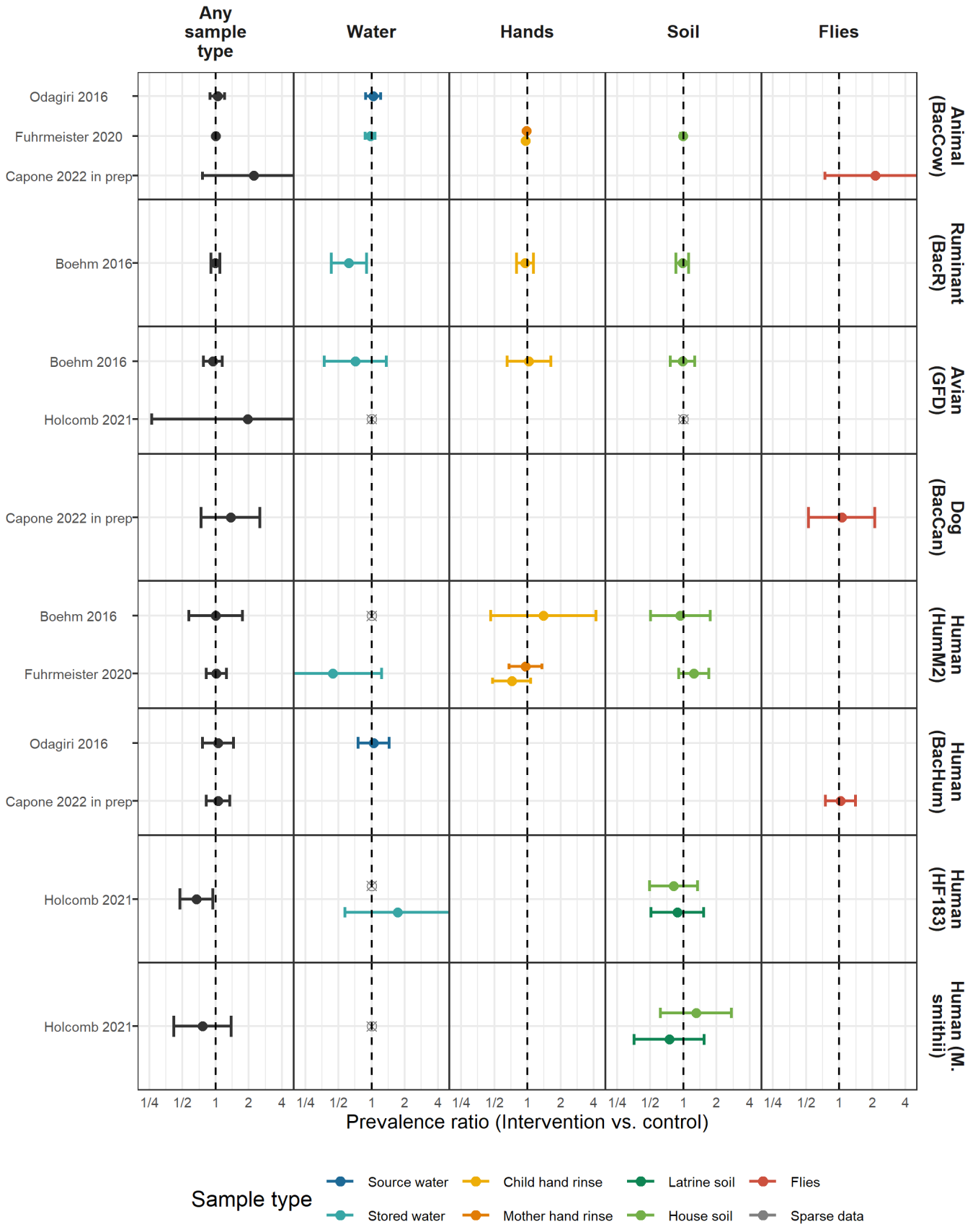
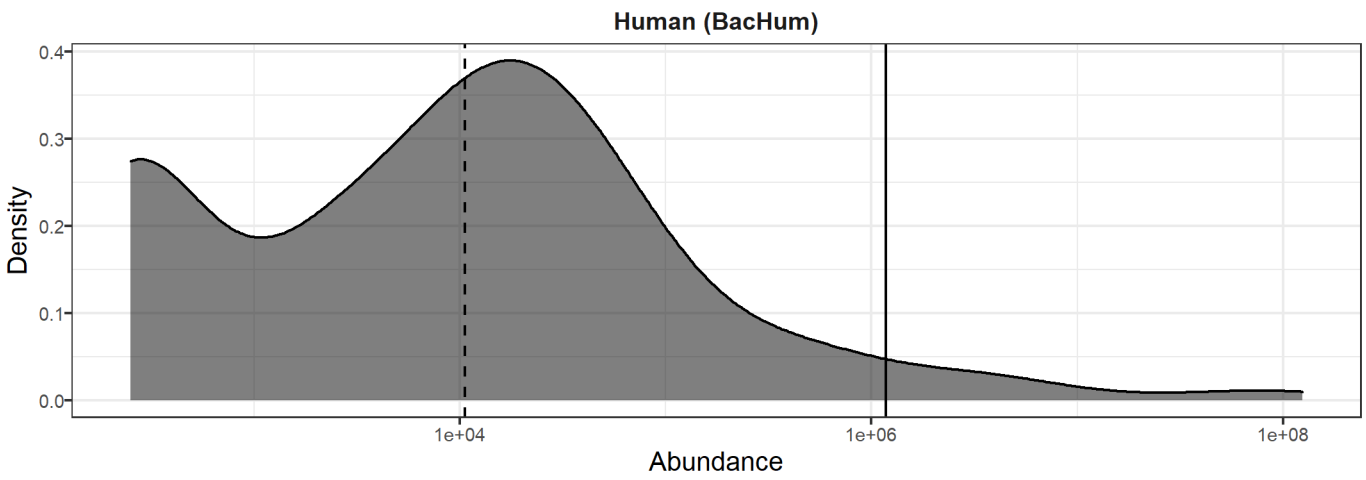
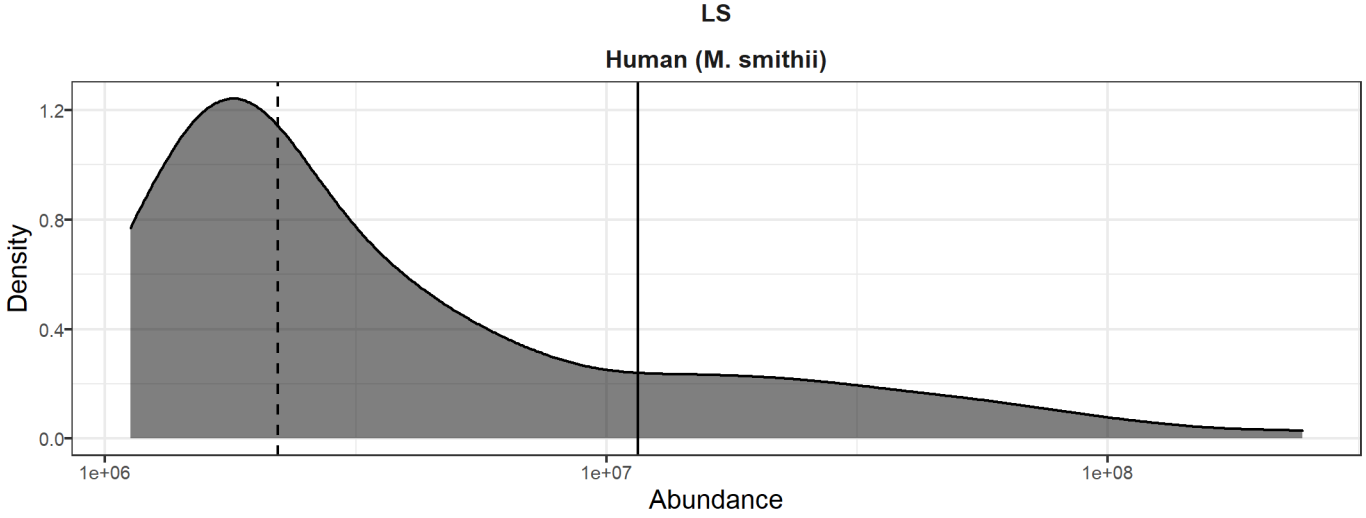


Figure S3. Forest plots of intervention effects on the prevalence of specific MST markers.

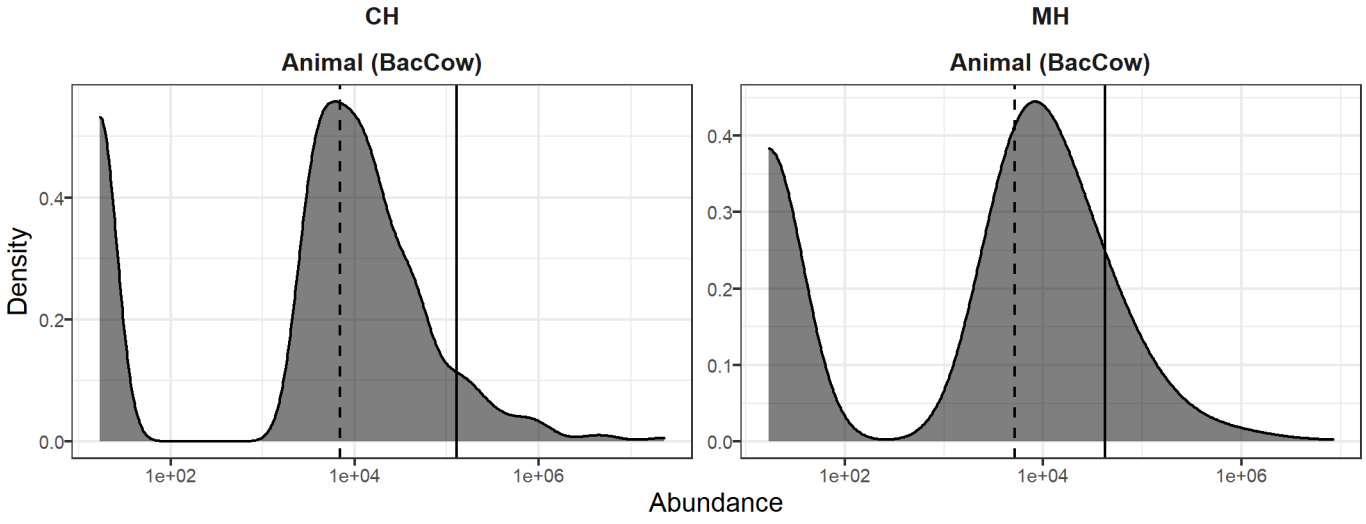
Capone et al. in prep 2022



Holcomb et al. 2020



Fuhrmeister et al. 2020



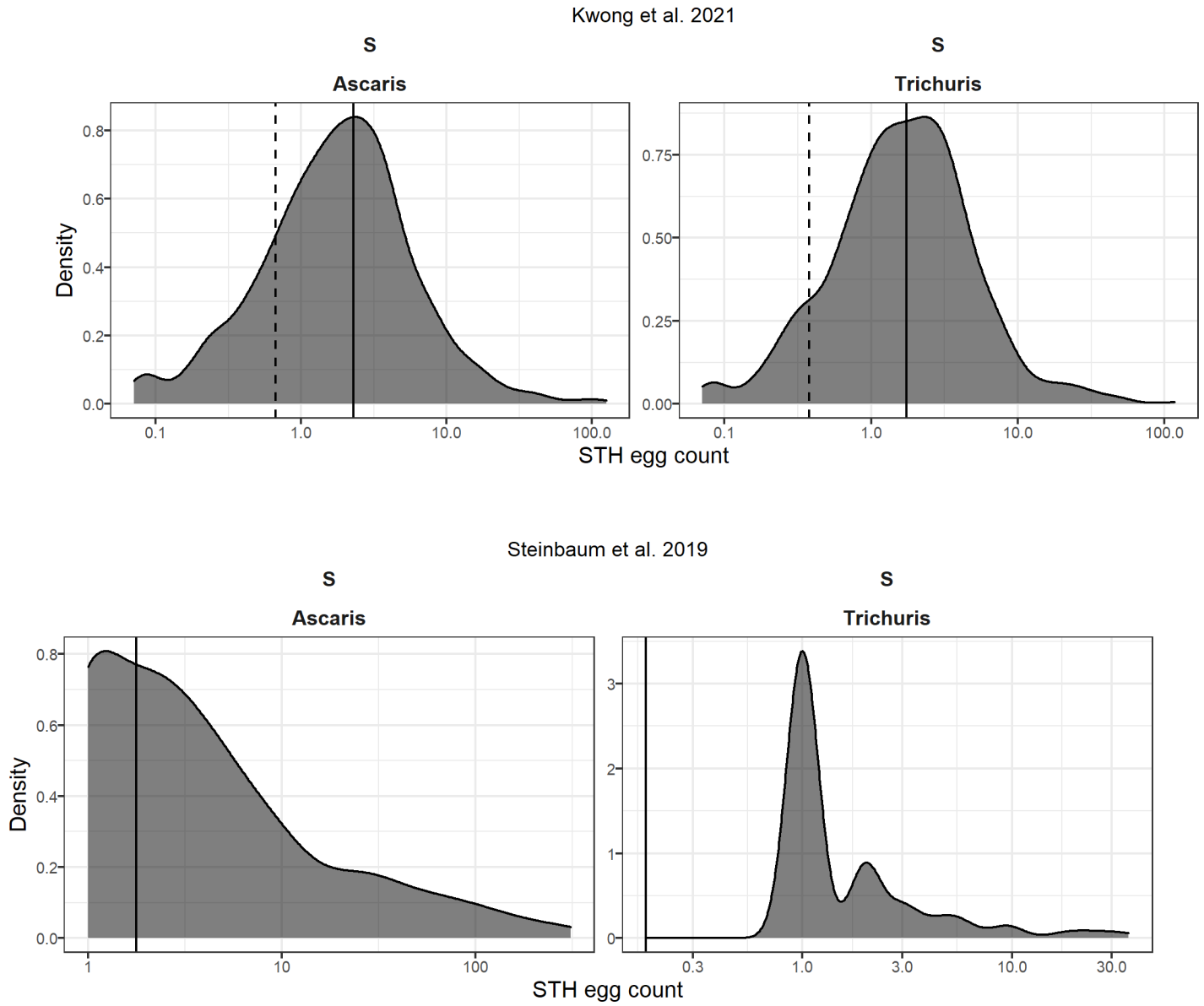


Figure S4. Distributions of abundance outcomes. The X-axes are displayed on the log-10 scale. Black vertical lines mark the means, and dashed lines mark the medians. Values below the limit of detection were imputed with with half the limit of detection and values below the limit of quantification were imputed with the midpoint between the limits of detections and quantification, leading to some bimodal distributions.

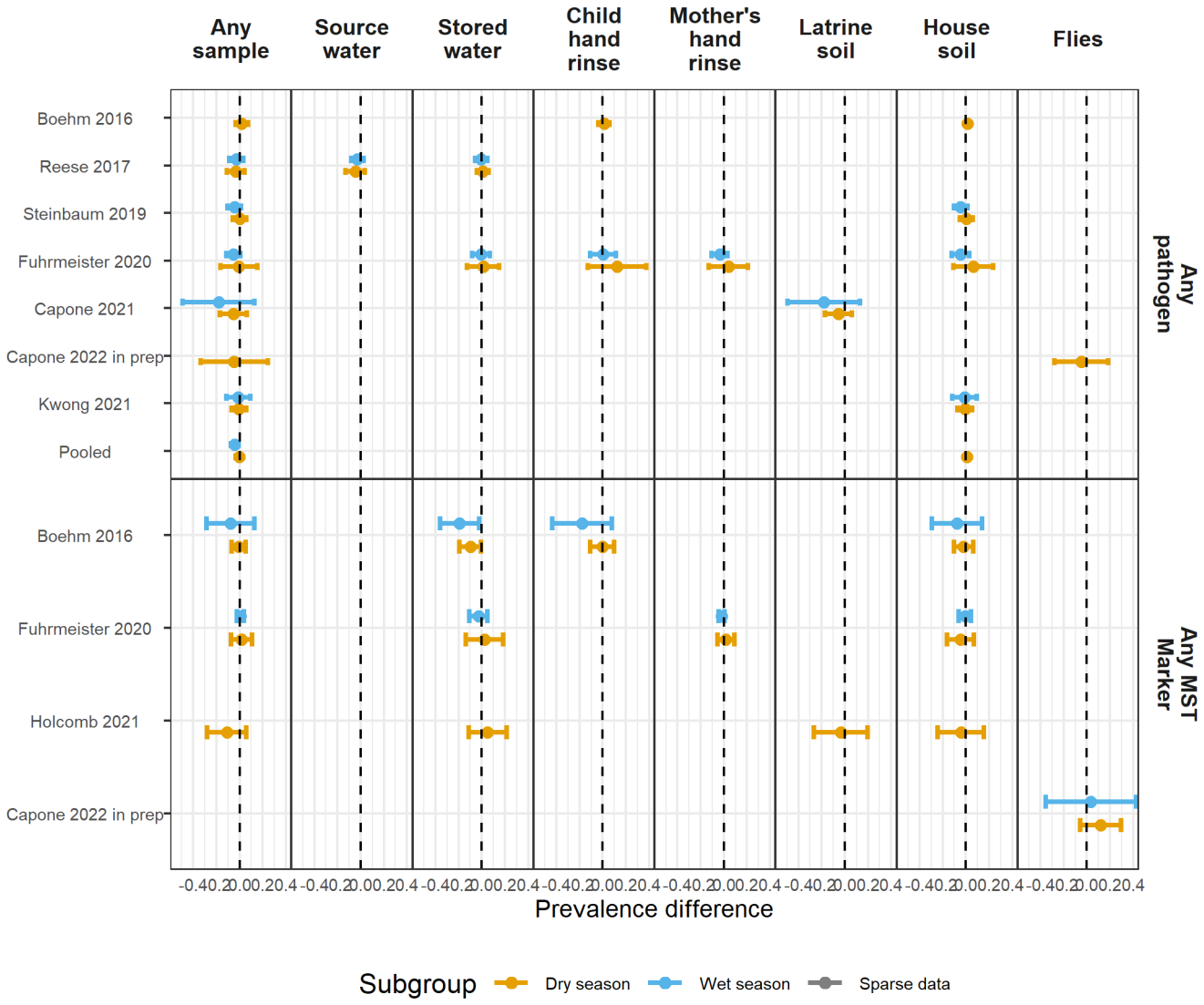


Figure S5. Forest plots of any enteropathogen prevalence differences or any MST prevalence differences between intervention and control arms, stratified by whether the sample was collected during the wet versus dry season (defined by the 6 months of highest average rainfall). Significant effect modification, as determined by the p-values on the regression model interaction term, is marked above points with asterisks (P < 0.05 = "*", P < 0.01 = "**", P < 0.001 = "***"). Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).

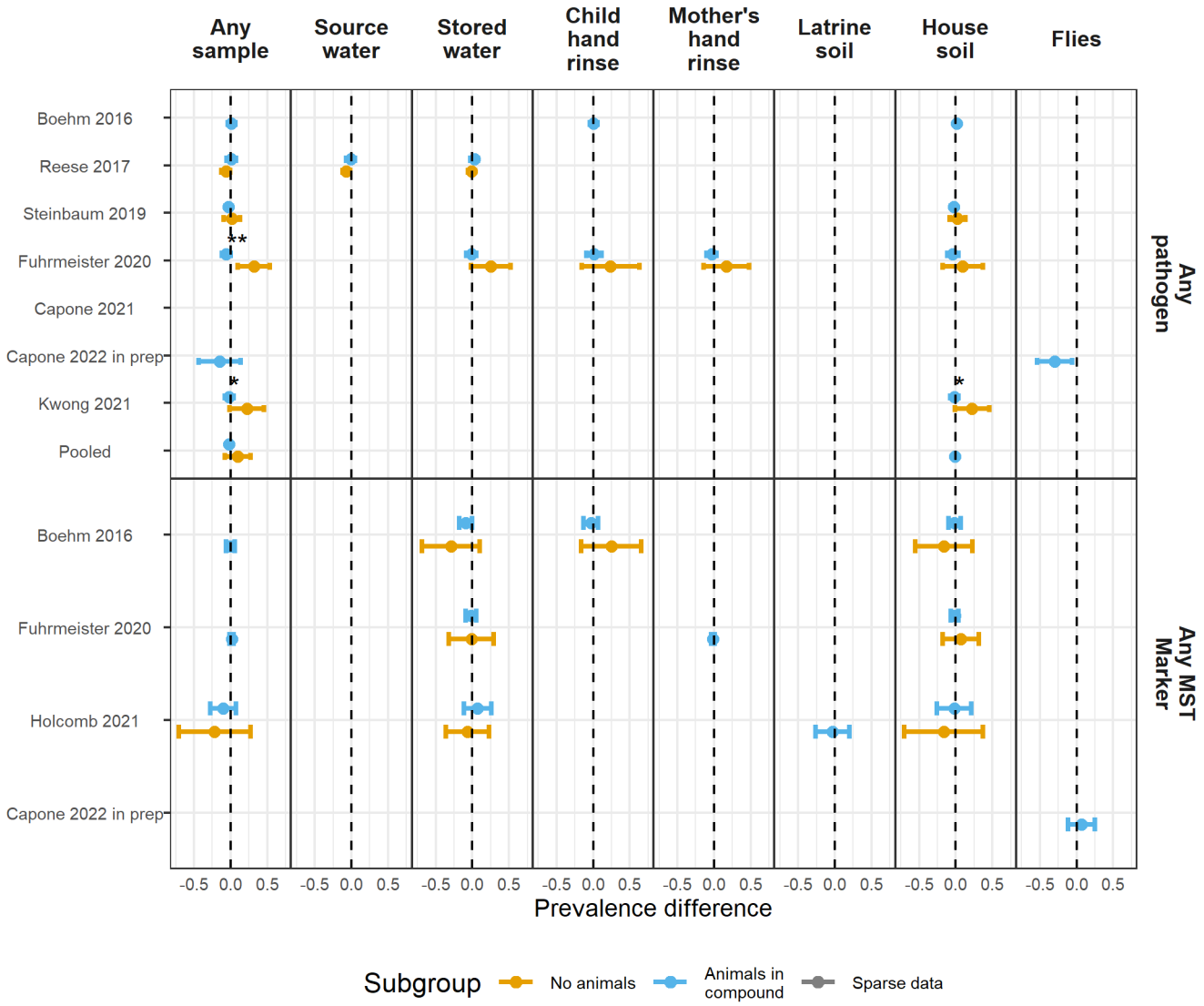


Figure S6. Forest plots of any enteropathogen prevalence differences or any MST prevalence differences between intervention and control arms, stratified by whether any animals were present in the compound. Significant effect modification, as determined by the p-values on the regression model interaction term, is marked above points with asterisks (P < 0.05 = "*", P < 0.01 = "**", P < 0.001 = "***"). Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).

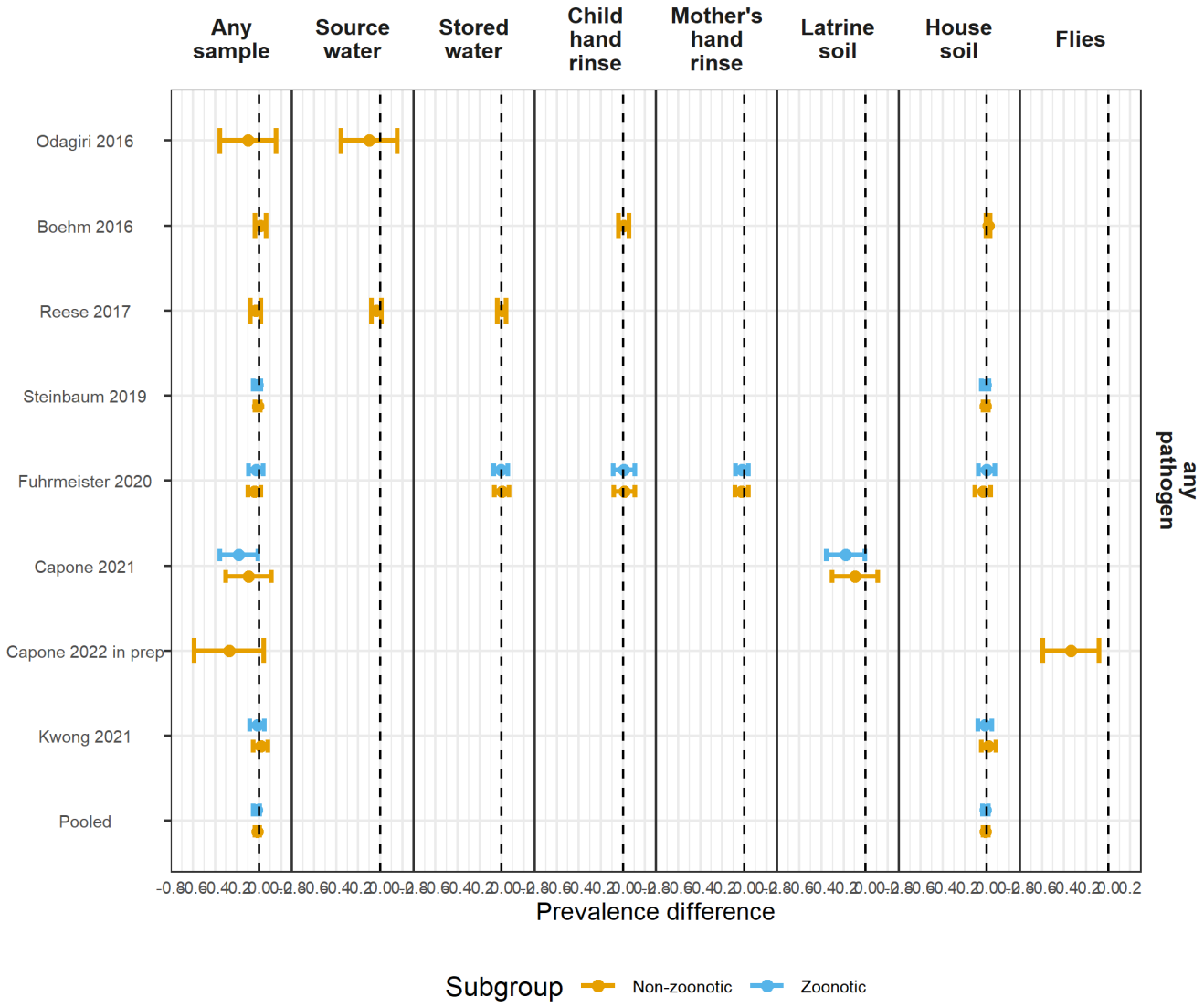


Figure S7. Forest plots of any enteropathogen prevalence differences or any MST prevalence differences between intervention and control arms, stratified by whether the pathogen is zoonotically transmitted. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations). Significant effect modification, as determined by the p-values on the regression model interaction term, is marked above points with asterisks (P < 0.05 = "*", P < 0.01 = "**", P < 0.001 = "***"). Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).

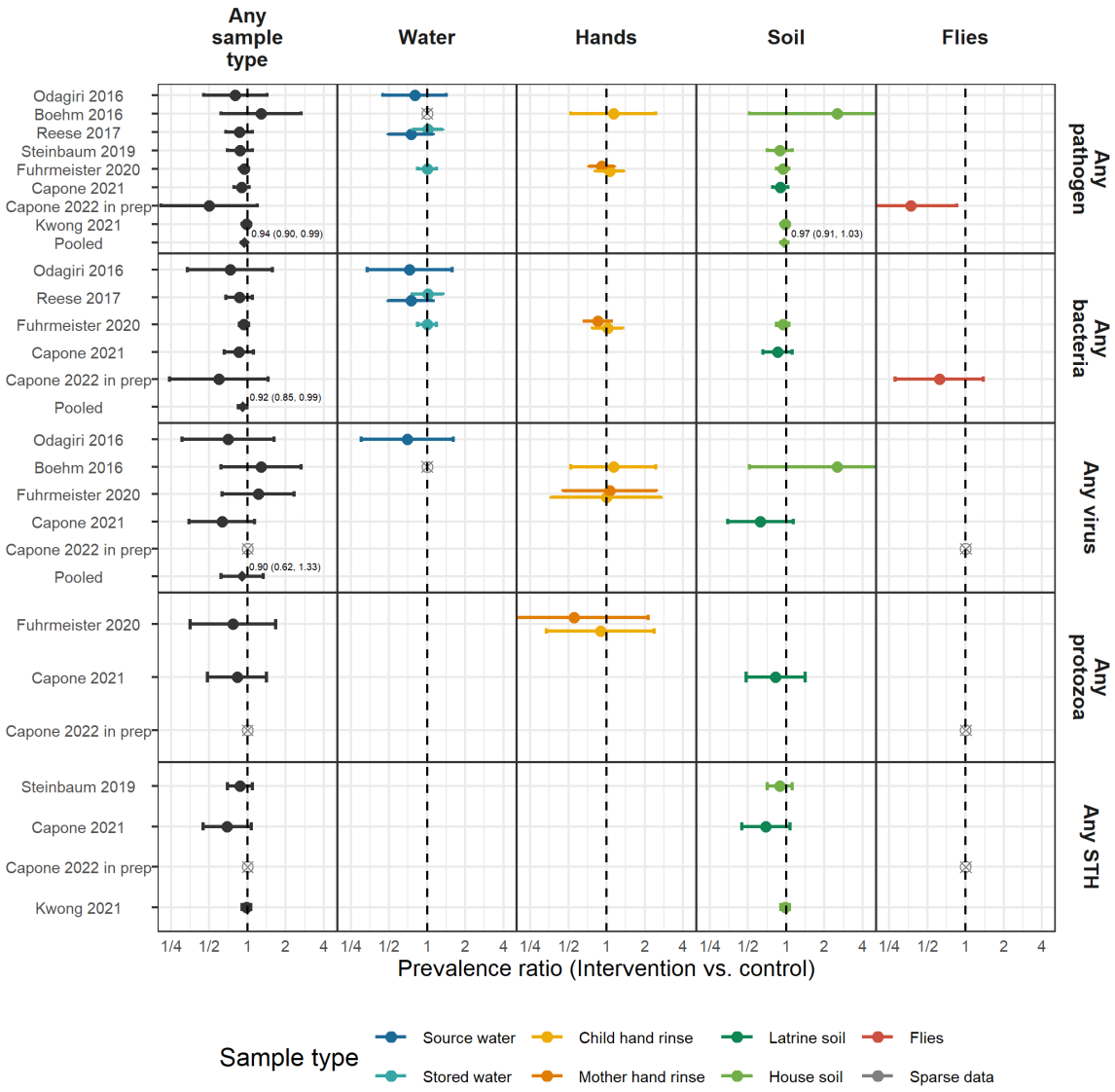


Figure S8. Forest plots of unadjusted intervention effects on the prevalence of any enteropathogen or type of enteropathogen (any bacteria, any virus, any protozoa and any STH) in different types of environmental samples. Point estimates and confidence intervals are printed next to pooled estimates. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).

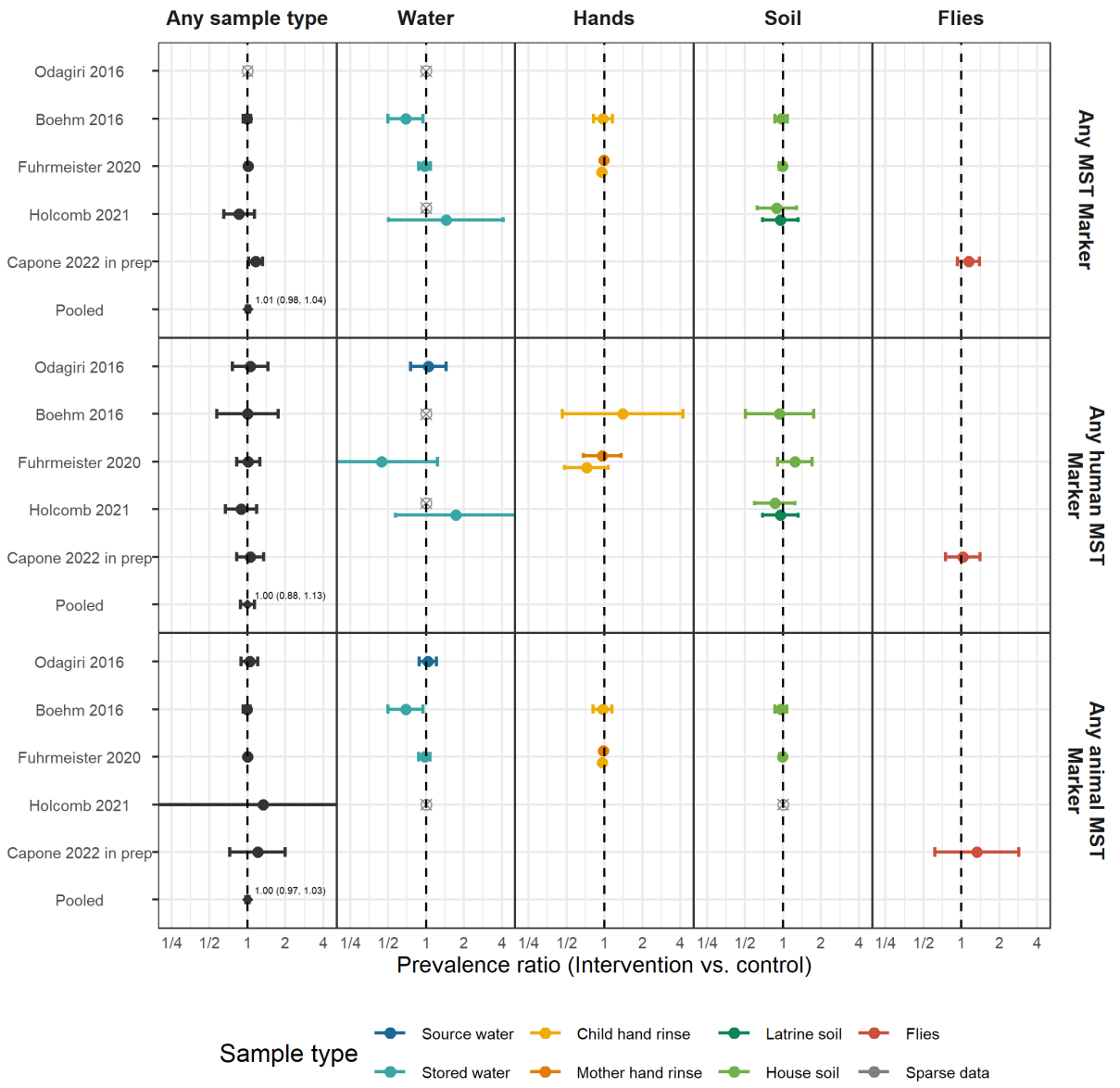


Figure S9. Forest plots of unadjusted intervention effects on the prevalence of any MST marker or type of MST marker (human or animal MST markers) in different types of environmental samples. Point estimates and confidence intervals are printed next to pooled estimates. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).

Table S1. Systematic review search terms

Search terms were combined with "OR" within columns and with "AND" across columns.

Study design	WASH	Environmental markers	Child health
matched, trial, RCT, experiment, intervention, randomized, randomised, quasi-randomized, quasi-randomised, quasi-experimental, pseudo-randomized, pseudo-randomised, non-randomized controlled trials	Water, Sanitation, Hygiene, Handwashing, WSH, Sanitation, Water Supply, Sanitary Drainage, Toilet Facilities, Drinking Water, Hand Hygiene, Water Purification, Waste Water, disinfection	molecular source tracking, microbial source tracking, microbial transmission, diarrheal pathogen, diarrheal pathogens, diarrhoeal pathogen, diarrhoeal pathogens, fecal-oral, faecal-oral, entericpathogen, entericpathogens, ruminant, avian, Feces, Faeces, Fecal, Faecal, Fecally, Faecally	Entericinfection, Soil-transmitted helminth, Protozoan, Seroconversion, Fecal microbiology, Faecal microbiology, Fecal biomarker, Faecal biomarker, Intestinal Diseases, Parasitic, Seroconversion, Enteritis, Helminthiasis, Helminthiases, Intestinal infection, Viral infection, Bacterial infection, Parasite infection, Parasitic infection, Helminth infection, Fecal sampling, Faecal sampling, Stool sampling, Stool collection, Diarrhea, Dysentery, Child growth faltering, Growth faltering, Child development, Length-for-age, Height-for-age, Weight-for-age, Head circumference, Waist circumference, Stunting, Stunted, Wasting, Wasted, Linear growth, Anthropometric measurement, Malnutrition, Undernourished, Undernutrition, Underweight, Growth Disorders, Childnutrition disorder, Wasting syndrome, Thinness, Growth velocity

Table S2. Pubmed search string

[MH] are mesh headers and [TW] are text words.

((matched [tw]) OR (trial [tw]) OR (RCT [tw]) OR (experiment [tw]) OR (intervention [tw]) OR (randomized [tw]) OR (randomised [tw]) OR (quasi-randomized [tw]) OR (quasi-randomised [tw]) OR (quasi-experimental [tw]) OR (pseudo-randomized [tw]) OR (pseudo-randomised [tw]) OR ("non-randomized controlled trials as topic" [mh])) AND ((Water [tw]) OR (Sanitation [tw]) OR (Hygiene [tw]) OR (Handwashing [tw]) OR (WSH [tw]) OR ("Sanitation" [mh]) OR ("Water Supply" [mh]) OR ("Drainage, Sanitary" [mh]) OR (Sanitary Drainage [tw]) OR ("Toilet Facilities" [mh]) OR ("Drinking Water" [mh]) OR ("Hand Hygiene" [mh]) OR ("Water Purification" [mh]) OR ("Waste Water" [mh]) OR (disinfect* [tw])) AND ((molecular source tracking [tw]) OR (microbial source tracking [tw]) OR (microbial transmission [tw]) OR (diarrheal pathogen [tw]) OR (diarrheal pathogens [tw]) OR (diarrhoeal pathogen [tw]) OR (diarrhoeal pathogens [tw]) OR (fecal-oral [tw]) OR (faecal-oral [tw]) OR (enteric pathogen [tw]) OR (enteric pathogens [tw]) OR (ruminant* [tw]) OR (avian* [tw]) OR ("Feces" [mh]) OR (Feces [tw]) OR (Faeces [tw]) OR (Fecal [tw]) OR (Faecal [tw]) OR (Fecally [tw]) OR (Faecally [tw])) AND (((Enteric infection* [tw]) OR (Soil-transmitted helminth* [tw]) OR (Protozoan* [tw]) OR (Seroconversion [tw]) OR (Fecal microbio* [tw]) OR (Faecal microbio* [tw]) OR (Fecal biomarker* [tw]) OR (Faecal biomarker* [tw]) OR ("Intestinal Diseases, Parasitic/epidemiology" [mh]) OR ("Seroconversion" [mh]) OR (Seroconversion [tw]) OR ("Enteritis/epidemiology" [mh]) OR ("Helminthiasis/complications" [mh]) OR (Helminthiasis [tw]) OR (Helminthiases)OR ("Helminthiasis/epidemiology" [mh]) OR ("Helminthiasis/prevention and control" [mh]) OR (Intestinal infection* [tw]) OR (Viral infection* [tw]) OR (Bacterial infection* [tw]) OR (Parasite infection* [tw]) OR (Parasitic infection* [tw]) OR (Helminth infection* [tw]) OR (Fecal sampling [tw]) OR (Faecal sampling [tw]) OR (Bacterial infection* [tw]) OR (Parasite infection* [tw]) OR (Parasitic infection* [tw]) OR (Helminth infection* [tw]) OR (Fecal sampling [tw]) OR (Faecal sampling [tw]) OR (Stool sampling [tw]) OR (Stool collection [tw])) OR ((Diarrh* [tw]) OR (Dysentery [tw]) OR ("Diarrhea/epidemiology" [mh]) OR ("Diarrhea/etiology" [mh]) OR ("Diarrhea/prevention and control" [mh]) OR ("Diarrhea, Infantile" [mh]) OR ("Dysentery" [mh])) OR (Child growth faltering [tw]) OR (Growth faltering [tw])OR (Child development [tw]) OR (Length-for-age [tw]) OR (Height-for-age [tw]) OR (Weight-for-age [tw]) OR (Head circumference [tw]) OR (Waist circumference [tw]) OR (Stunt* [tw]) OR (Wasting [tw]) OR (Wasted [tw]) OR (Linear growth [tw]) OR (Anthropometric measurement* [tw]) OR (Maln* [tw]) OR (Undernourish* [tw]) OR (Undernutrition [tw]) OR (Underweight [tw]) OR ("Growth Disorders" [mh]) OR (Growth Disorders [tw]) OR ("Child nutrition disorders" [mh]) OR (Child nutrition disorder* [tw]) OR ("Malnutrition" [mh]) OR ("Wasting Syndrome" [mh]) OR (Wasting syndrome [tw]) OR ("Thinness" [mh]) OR (Thinness [tw]) OR (Growth velocity [tw]))

Table S3. PRISMA Checklist

Topic	No.	Item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA for Abstracts checklist below	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Line 101-112
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Line 116-119
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Line 123-133
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Line 122-123, Fig. S1
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Tables S1-S2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Line 130-133

Topic	No.	Item	Location where item is reported
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Line 135-138
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Line 135-151
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Line 135-151
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Line 132-133
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Line 152-203
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	Line 139-151

Topic	No.	Item	Location where item is reported
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Line 139-173
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Figure captions
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Line 161-193
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Line 180-200
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Line 326-327
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Not applicable
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Not applicable
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Figure S1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Line 212-214

Topic	No.	Item	Location where item is reported
Study characteristics	17	Cite each included study and present its characteristics.	Line 219-230, Table 1
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Table S4
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Figures1 ,2 S2-S3, S5-S8, Tables 2, S6-S9
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Not applicable
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Figures1 ,2 S2-S3, S5-S8, Tables 2, S6-S9
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Line 318-327
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Line 326-327
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Not applicable
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Figures1 ,2 S2-S3, S5-S8, Tables 2, S6-S9
DISCUSSION			

Topic	No.	Item	Location where item is reported
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Line 330-359
	23b	Discuss any limitations of the evidence included in the review.	Line 384-405
	23c	Discuss any limitations of the review processes used.	Line 406-410
	23d	Discuss implications of the results for practice, policy, and future research.	Line 415-424
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	https://osf.io/8sgzn/
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	https://osf.io/8sgzn/
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not applicable
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Line 36
Competing interests	26	Declare any competing interests of review authors.	Not applicable
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	https://github.com/amertens/wash-ipd

Topic	No.	Item	Reported?
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	No
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesize results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	Yes
Registration	12	Provide the register name and registration number.	Yes

Table S4. Risk of bias based on modified Newcastle-Ottawa scale

Stars are given for low risk of bias in each category, up to a total of nine stars. Scoring details are in the footnotes.

Reference	Selection bias	Response bias	Follow-up bias	Misclassification bias	Outcome assessment	Outcome measurement	Bias in analysis	Total
	Is there evidence of selection bias, which refers to systematic differences between baseline characteristics of the groups that are compared?[a]	Is there evidence of response bias?[b]	Is there evidence of bias due to missing follow-up data?[c]	Is there risk of households not receiving the intervention being misclassified as having received it, or vice versa?[d]	Is there evidence of bias arising from how the outcome was assessed?[e]	Is there evidence of ascertainment bias?[f]	Is there evidence that analysis was not appropriately adjusted for clustering and/or confounding, if appropriate?[g]	Total number of stars (x/9 possible stars).
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. Lancet Glob Health. 2014.	*	* no, laboratory assessed and blinded	possible (86% of possible weeks are reported weeks)	* household-level interventions	**	*	** adjusted for clustering	8
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 2018	*	* no, laboratory assessed and blinded	* 94% complete FU	* household-level interventions	**	*	**	9
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 2018	*	* no, laboratory assessed and blinded	* <1% loss to FU	* household-level interventions	**	*	**	9

Reference	Selection bias	Response bias	Follow-up bias	Misclassification bias	Outcome assessment	Outcome measurement	Bias in analysis	Total
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 2019	selection bias is possible, as the study is not randomized and there are some baseline differences between intervention and control group	* no, laboratory assessed and blinded	substantial loss to FU	* household-level interventions	**	*	**	7
Knee, J. et al. Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-after trial. eLife 2011	selection bias is possible, as the study is not randomized, but intervention and control groups were mostly balanced at baseline. Control households were more likely to have covered floors and higher quality walls and intervention groups had more people per household.	* no, laboratory assessed and blinded	substantial loss to FU	* household-level interventions	**	*	**	7

[a] RCTs receive 1 star, unless evidence of selection bias (e.g. randomisation procedures not followed). Meaningful differences between groups at baseline in RCTs receive 0 stars. Rates of declining to participate >10% receive 0 stars. Non- or quasi-randomised studies receive 0 stars.

[b] If intervention recipient was not blinded to intervention status, 0 stars.

[c] <10% receives 1 star, greater than or equal to 10% receives 0 stars.

[d] Interventions delivered at the household/individual level receive 1 star. Interventions delivered at the community level that missed a substantial, i.e. greater than or equal to 10%, proportion of the target population receive 0 stars, including when there is insufficient information to verify whether this is the case. Interventions with substantial risk of contamination (control households receiving intervention) receive 0 stars.

[e] Parent / person recall (=0 stars). Fieldworker assessed (=1 star). Physician/microbiologically assessed (=2 stars)

[f] If outcome measurement staff were not blinded to intervention status, 0 stars.

[g] Scoring is based on losing stars (max. 2). Individual RCTs with baseline balance on covariates are unlikely to require adjustment (=2 stars). Cluster-RCTs and non-randomised trials may require adjustment for clustering (-1 star if not done). RCTs or cRCTs may require adjustment for covariates, with justification (-1 star if not done). Non-randomised studies require adjustment for covariates (-1 star if not done), but also adequate justification for covariate selection (-1 star if not included), and there can be too few or too many covariates.

Table S5. Prevalence of pathogens by sample type tested in each study

Study	Sample	Target	Percent positive (n/N)	PR (95% CI)
Odagiri 2016	Source water	V. cholerae	31.7% (19/60)	0.73 (0.34, 1.57)
-	-	Adenovirus	8.3% (5/60)	0.25 (0.03, 2.19)
-	-	Rotavirus	23.3% (14/60)	0.75 (0.29, 1.93)
Boehm 2016	Stored water	Rotavirus	0.6% (3/493)	-
-	Child hand rinse	Rotavirus	6.1% (30/493)	-
-	House soil	Rotavirus	1.4% (7/496)	2.52 (0.51, 12.42)
Reese 2017	Source water	Shigella	10.7% (161/1499)	0.73 (0.46, 1.15)
-	-	V. cholerae	13% (36/276)	0.93 (0.46, 1.85)
-	Stored water	Shigella	10.1% (190/1874)	1.08 (0.77, 1.51)
-	-	V. cholerae	23.7% (100/422)	1.03 (0.66, 1.6)
Steinbaum 2019	House soil	Ascaris	13% (273/2107)	0.88 (0.68, 1.13)
-	-	Trichuris	6.9% (146/2107)	0.86 (0.6, 1.23)
Fuhrmeister 2020	Stored water	Pathogenic E. coli	38.6% (286/741)	1 (0.84, 1.19)
-	Child hand rinse	Pathogenic E. coli	34% (127/373)	-
-	-	Giardia	4.8% (15/311)	-
-	-	Norovirus	4.2% (14/337)	-
-	Mother's hand rinse	Pathogenic E. coli	24% (177/737)	-
-	-	Giardia	2.3% (14/602)	-
-	-	Norovirus	3.1% (21/684)	-
-	House soil	Pathogenic E. coli	61.3% (453/739)	0.94 (0.84, 1.06)
Capone 2021	Latrine soil	C. difficile	14.8% (13/88)	0.9 (0.32, 2.48)
-	-	Campylobacter	6.8% (6/88)	2.09 (0.4, 11.05)
-	-	Pathogenic E. coli	56.8% (50/88)	0.89 (0.56, 1.42)
-	-	Salmonella	6.8% (6/88)	0.52 (0.1, 2.76)
-	-	Shigella	21.6% (19/88)	0.28 (0.1, 0.78)
-	-	V. cholerae	0% (0/88)	-
-	-	Yersinia	4.5% (4/88)	-
-	-	Ascaris	60.2% (53/88)	0.65 (0.41, 1.02)
-	-	Trichuris	17% (15/88)	0.92 (0.36, 2.33)
-	-	Cryptosporidium	8% (7/88)	0.78 (0.18, 3.36)
-	-	Entamoeba histolytica	1.1% (1/88)	-

Study	Sample	Target	Percent positive (n/N)	PR (95% CI)
-	-	Giardia	31.8% (28/88)	0.47 (0.21, 1.07)
-	-	Adenovirus	20.5% (18/88)	0.21 (0.06, 0.68)
-	-	Astrovirus	29.5% (26/88)	1.27 (0.67, 2.43)
-	-	Norovirus	2.3% (2/88)	-
-	-	Rotavirus	4.5% (4/88)	-
-	-	Sapovirus	0% (0/88)	-
Capone 2022 in prep	Flies	Campylobacter	1.2% (1/86)	-
-	-	Pathogenic E. coli	30.2% (26/86)	-
-	-	Shigella	2.3% (2/86)	-
-	-	V. cholerae	2.3% (2/86)	-
-	-	Ascaris	0% (0/86)	-
-	-	Trichuris	3.5% (3/86)	-
-	-	Giardia	4.7% (4/86)	-
-	-	Adenovirus	4.7% (4/86)	-
-	-	Astrovirus	0% (0/86)	-
-	-	Norovirus	2.3% (2/86)	-
-	-	Pan enterovirus	0% (0/86)	-
-	-	Rotavirus	1.2% (1/86)	-
-	-	Sapovirus	0% (0/86)	-
Kwong 2021	House soil	Ascaris	62.1% (886/1426)	0.97 (0.87, 1.08)
-	-	Trichuris	56% (798/1426)	1.03 (0.91, 1.15)

Table S6. Prevalence of microbial source tracking markers by sample type tested in each study

Study	Sample	Target	Percent positive (n/N)	PR (95% CI)
Odagiri 2016	Source water	Animal (BacCow)	91.7% (55/60)	1.04 (0.89, 1.21)
-	-	Human (BacHum)	71.7% (43/60)	1.05 (0.76, 1.45)
Boehm 2016	Stored water	Avian (GFD)	9.3% (46/493)	0.71 (0.37, 1.36)
-	-	Ruminant (BacR)	21.9% (108/493)	0.62 (0.43, 0.9)
-	-	Human (HumM2)	0% (0/493)	-
-	Child hand rinse	Avian (GFD)	16.2% (80/493)	-
-	-	Ruminant (BacR)	54.2% (267/493)	-
-	-	Human (HumM2)	2.4% (12/493)	-
-	House soil	Avian (GFD)	33.3% (165/496)	0.98 (0.76, 1.27)
-	-	Ruminant (BacR)	66.7% (331/496)	0.98 (0.85, 1.12)
-	-	Human (HumM2)	8.9% (44/496)	0.94 (0.5, 1.75)
Fuhrmeister 2020	Stored water	Animal (BacCow)	68.5% (482/704)	0.97 (0.87, 1.08)
-	-	Human (HumM2)	2.6% (17/651)	0.44 (0.16, 1.23)
-	Child hand rinse	Animal (BacCow)	97.5% (356/365)	-
-	-	Human (HumM2)	21.9% (74/338)	-
-	Mother's hand rinse	Animal (BacCow)	96.7% (702/726)	-
-	-	Human (HumM2)	18.1% (118/651)	-
-	House soil	Animal (BacCow)	90.6% (572/631)	0.99 (0.94, 1.04)
-	-	Human (HumM2)	20.1% (127/631)	1.24 (0.91, 1.7)
Holcomb 2021	Source water	Avian (GFD)	0% (0/41)	-
-	-	Human (HF183)	2.4% (1/41)	-
-	-	Human (M. smithii)	0% (0/41)	-
-	Stored water	Avian (GFD)	1.1% (1/94)	-
-	-	Human (HF183)	14.9% (14/94)	1.72 (0.57, 5.18)
-	-	Human (M. smithii)	0% (0/94)	-
-	Latrine soil	Avian (GFD)	3.3% (2/60)	-
-	-	Human (HF183)	50% (30/60)	0.88 (0.51, 1.52)
-	-	Human (M. smithii)	45% (27/60)	0.74 (0.36, 1.55)
-	House soil	Avian (GFD)	3.6% (3/83)	-
-	-	Human (HF183)	42.2% (35/83)	0.81 (0.49, 1.34)
-	-	Human (M. smithii)	24.1% (20/83)	1.3 (0.62, 2.73)

Study	Sample	Target	Percent positive (n/N)	PR (95% CI)
Capone 2022 in prep	Flies	Animal (BacCow)	12.8% (11/86)	-
-	-	Dog (BacCan)	30.2% (26/86)	-
-	-	Human (BacHum)	72.1% (62/86)	-

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any pathogen	Any sample	7	13	20	17	57	PR=0.65 (95% CI: 0.33, 1.28)	0.21	PR=0.5 (95% CI: 0.21, 1.19)
Any pathogen	Any sample	37	6	43	2	88	PR=0.9 (95% CI: 0.78, 1.03)	0.13	PR=0.9 (95% CI: 0.78, 1.03)
Any pathogen	Any sample	314	136	348	123	921	PR=0.94 (95% CI: 0.87, 1.02)	0.17	PR=0.94 (95% CI: 0.87, 1.02)
Any pathogen	Any sample	206	979	173	707	2,065	PR=0.88 (95% CI: 0.7, 1.11)	0.29	PR=0.87 (95% CI: 0.7, 1.09)
Any pathogen	Any sample	185	792	238	825	2,040	PR=0.85 (95% CI: 0.66, 1.08)	0.18	PR=0.86 (95% CI: 0.68, 1.09)
Any pathogen	Any sample	19	229	15	234	497	PR=1.27 (95% CI: 0.6, 2.68)	0.53	PR=1.28 (95% CI: 0.62, 2.66)
Any pathogen	Any sample	12	18	15	15	60	PR=0.8 (95% CI: 0.45, 1.42)	0.45	
Any pathogen	Source water	68	588	122	747	1,525	PR=0.74 (95% CI: 0.49, 1.12)	0.15	PR=0.74 (95% CI: 0.5, 1.12)
Any pathogen	Source water	12	18	15	15	60	PR=0.8 (95% CI: 0.45, 1.42)	0.45	
Any pathogen	Stored water	138	218	148	237	741	PR=1.01 (95% CI: 0.85, 1.2)	0.93	PR=1 (95% CI: 0.84, 1.19)
Any pathogen	Stored water	134	786	147	860	1,927	PR=1 (95% CI: 0.75, 1.32)	0.99	PR=1.01 (95% CI: 0.77, 1.34)
Any pathogen	Stored water	2	243	1	245	491	Not estimated		Not estimated
Any pathogen	House soil	363	125	687	221	1,396	PR=0.98 (95% CI: 0.91, 1.06)	0.67	PR=0.98 (95% CI: 0.91, 1.06)
Any pathogen	House soil	217	144	236	142	739	PR=0.96 (95% CI: 0.86, 1.08)	0.53	PR=0.94 (95% CI: 0.84, 1.06)
Any pathogen	House soil	209	1,000	173	725	2,107	PR=0.9 (95% CI: 0.72, 1.13)	0.35	PR=0.89 (95% CI: 0.71, 1.11)
Any pathogen	House soil	5	242	2	247	496	PR=2.52 (95% CI: 0.51, 12.42)	0.26	PR=2.52 (95% CI: 0.51, 12.42)
Any pathogen	Latrine soil	37	6	43	2	88	PR=0.9 (95% CI: 0.78, 1.03)	0.13	PR=0.9 (95% CI: 0.78, 1.03)
Any pathogen		8	23	25	30	86	PR=0.57 (95% CI: 0.28, 1.15)	0.12	PR=0.37 (95% CI: 0.16, 0.85)
Any pathogen		75	113	72	116	376	PR=1.04 (95% CI: 0.8, 1.35)	0.76	PR=1.05 (95% CI: 0.81, 1.37)
Any pathogen		96	266	110	267	739	PR=0.91 (95% CI: 0.72, 1.15)	0.43	PR=0.92 (95% CI: 0.72, 1.16)

Table S7.

Unadjusted and adjusted results by study, sample type, and aggregated variables for pathogen targets (any pathogen, any bacteria, any viruses, any protozoa, any STH).

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any pathogen		16	231	14	232	493	PR=1.14 (95% CI: 0.52, 2.48)	0.75	PR=1.13 (95% CI: 0.52, 2.44)
Any bacteria	Any sample	7	13	17	20	57	PR=0.76 (95% CI: 0.38, 1.54)	0.45	PR=0.6 (95% CI: 0.24, 1.46)
Any bacteria	Any sample	28	15	35	10	88	PR=0.84 (95% CI: 0.64, 1.1)	0.2	PR=0.85 (95% CI: 0.65, 1.11)
Any bacteria	Any sample	306	144	340	131	921	PR=0.94 (95% CI: 0.86, 1.03)	0.18	PR=0.94 (95% CI: 0.86, 1.02)
Any bacteria	Any sample	185	792	238	825	2,040	PR=0.85 (95% CI: 0.66, 1.08)	0.18	PR=0.86 (95% CI: 0.68, 1.09)
Any bacteria	Any sample	8	22	11	19	60	PR=0.73 (95% CI: 0.34, 1.57)	0.42	
Any bacteria	Source water	68	588	122	747	1,525	PR=0.74 (95% CI: 0.49, 1.12)	0.15	PR=0.74 (95% CI: 0.5, 1.12)
Any bacteria	Source water	8	22	11	19	60	PR=0.73 (95% CI: 0.34, 1.57)	0.42	
Any bacteria	Stored water	138	218	148	237	741	PR=1.01 (95% CI: 0.85, 1.2)	0.93	PR=1 (95% CI: 0.84, 1.19)
Any bacteria	Stored water	134	786	147	860	1,927	PR=1 (95% CI: 0.75, 1.32)	0.99	PR=1.01 (95% CI: 0.77, 1.34)
Any bacteria	House soil	217	144	236	142	739	PR=0.96 (95% CI: 0.86, 1.08)	0.53	PR=0.94 (95% CI: 0.84, 1.06)
Any bacteria	Latrine soil	28	15	35	10	88	PR=0.84 (95% CI: 0.64, 1.1)	0.2	PR=0.85 (95% CI: 0.65, 1.11)
Any bacteria		8	23	21	34	86	PR=0.68 (95% CI: 0.32, 1.41)	0.3	PR=0.62 (95% CI: 0.28, 1.38)
Any bacteria		64	122	63	124	373	PR=1.02 (95% CI: 0.78, 1.35)	0.88	PR=1.02 (95% CI: 0.78, 1.35)
Any bacteria		81	281	96	279	737	PR=0.87 (95% CI: 0.68, 1.13)	0.3	PR=0.85 (95% CI: 0.67, 1.09)
Any virus	Any sample	0	20	4	33	57	Not estimated		Not estimated
Any virus	Any sample	16	27	22	23	88	PR=0.76 (95% CI: 0.46, 1.25)	0.28	PR=0.63 (95% CI: 0.35, 1.14)
Any virus	Any sample	17	330	14	338	699	PR=1.23 (95% CI: 0.63, 2.4)	0.54	PR=1.22 (95% CI: 0.63, 2.34)
Any virus	Any sample	19	229	15	234	497	PR=1.27 (95% CI: 0.6, 2.68)	0.53	PR=1.28 (95% CI: 0.62, 2.66)
Any virus	Any sample	7	23	10	20	60	PR=0.7 (95% CI: 0.3, 1.62)	0.4	
Any virus	Source water	7	23	10	20	60	PR=0.7 (95% CI: 0.3, 1.62)	0.4	

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any virus	Stored water	2	243	1	245	491	Not estimated		Not estimated
Any virus	House soil	5	242	2	247	496	PR=2.52 (95% CI: 0.51, 12.42)	0.26	PR=2.52 (95% CI: 0.51, 12.42)
Any virus	Latrine soil	16	27	22	23	88	PR=0.76 (95% CI: 0.46, 1.25)	0.28	PR=0.63 (95% CI: 0.35, 1.14)
Any virus		0	31	5	50	86	PR=0 (95% CI: 0, 0)	0	PR=0 (95% CI: 0, 0)
Any virus		7	162	7	161	337	PR=0.99 (95% CI: 0.37, 2.69)	0.99	PR=0.99 (95% CI: 0.37, 2.69)
Any virus		11	331	10	332	684	PR=1.1 (95% CI: 0.47, 2.57)	0.83	PR=1.06 (95% CI: 0.45, 2.46)
Any virus		16	231	14	232	493	PR=1.14 (95% CI: 0.52, 2.48)	0.75	PR=1.13 (95% CI: 0.52, 2.44)
Any protozoa	Any sample	0	20	3	34	57	Not estimated		Not estimated
Any protozoa	Any sample	15	28	19	26	88	PR=0.83 (95% CI: 0.48, 1.42)	0.49	PR=0.83 (95% CI: 0.48, 1.42)
Any protozoa	Any sample	12	293	16	291	612	PR=0.75 (95% CI: 0.35, 1.65)	0.48	PR=0.77 (95% CI: 0.35, 1.67)
Any protozoa	Latrine soil	15	28	19	26	88	PR=0.83 (95% CI: 0.48, 1.42)	0.49	PR=0.83 (95% CI: 0.48, 1.42)
Any protozoa		0	31	4	51	86	Not estimated		Not estimated
Any protozoa		7	147	8	149	311	PR=0.89 (95% CI: 0.33, 2.38)	0.82	PR=0.89 (95% CI: 0.33, 2.38)
Any protozoa		5	296	9	292	602	PR=0.56 (95% CI: 0.14, 2.13)	0.39	PR=0.56 (95% CI: 0.14, 2.13)
Any STH	Any sample	0	20	3	34	57	Not estimated		Not estimated
Any STH	Any sample	20	23	34	11	88	PR=0.62 (95% CI: 0.43, 0.89)	0.01	PR=0.69 (95% CI: 0.45, 1.07)
Any STH	Any sample	206	979	173	707	2,065	PR=0.88 (95% CI: 0.7, 1.11)	0.29	PR=0.87 (95% CI: 0.7, 1.09)
Any STH	House soil	363	125	687	221	1,396	PR=0.98 (95% CI: 0.91, 1.06)	0.67	PR=0.98 (95% CI: 0.91, 1.06)
Any STH	House soil	209	1,000	173	725	2,107	PR=0.9 (95% CI: 0.72, 1.13)	0.35	PR=0.89 (95% CI: 0.71, 1.11)
Any STH	Latrine soil	20	23	34	11	88	PR=0.62 (95% CI: 0.43, 0.89)	0.01	PR=0.69 (95% CI: 0.45, 1.07)

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any STH		0	31	3	52	86	Not estimated		Not estimated

Table S8.

Unadjusted and adjusted results by study, sample type, and aggregated variables for MST targets (any MST, any general MST, any human MST, any animal MST).

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any MST Marker	Any sample	20	0	32	5	57	PR=1.16 (95% CI: 1.02, 1.32)	0.03	PR=1.16 (95% CI: 1.02, 1.32)
Any MST Marker	Any sample	41	28	44	17	130	PR=0.82 (95% CI: 0.62, 1.09)	0.18	PR=0.86 (95% CI: 0.65, 1.13)
Any MST Marker	Any sample	421	26	438	29	914	PR=1 (95% CI: 0.97, 1.04)	0.8	PR=1.01 (95% CI: 0.97, 1.04)
Any MST Marker	Any sample	220	28	222	27	497	PR=0.99 (95% CI: 0.93, 1.06)	0.88	PR=0.99 (95% CI: 0.93, 1.06)
Any MST Marker	Any sample	30	0	28	2	60	Not estimated		
Any MST Marker	Source water	1	21	0	19	41	Not estimated		Not estimated
Any MST Marker	Source water	30	0	28	2	60	Not estimated		
Any MST Marker	Stored water	9	39	6	40	94	PR=1.44 (95% CI: 0.51, 4.08)	0.5	PR=1.44 (95% CI: 0.51, 4.08)
Any MST Marker	Stored water	230	119	256	119	724	PR=0.97 (95% CI: 0.87, 1.07)	0.52	PR=0.97 (95% CI: 0.88, 1.08)
Any MST Marker	Stored water	57	188	82	164	491	PR=0.7 (95% CI: 0.51, 0.96)	0.03	PR=0.69 (95% CI: 0.5, 0.95)
Any MST Marker	House soil	21	18	26	18	83	PR=0.91 (95% CI: 0.6, 1.38)	0.66	PR=0.89 (95% CI: 0.62, 1.28)
Any MST Marker	House soil	283	38	297	36	654	PR=0.99 (95% CI: 0.93, 1.05)	0.7	PR=0.99 (95% CI: 0.93, 1.05)
Any MST Marker	House soil	180	67	187	62	496	PR=0.97 (95% CI: 0.87, 1.08)	0.59	PR=0.97 (95% CI: 0.87, 1.08)
Any MST Marker	Latrine soil	21	9	22	8	60	PR=0.95 (95% CI: 0.69, 1.32)	0.78	PR=0.95 (95% CI: 0.69, 1.32)
Any MST Marker		27	4	42	13	86	PR=1.14 (95% CI: 0.93, 1.39)	0.2	PR=1.14 (95% CI: 0.93, 1.39)

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any MST Marker		174	11	182	1	368	PR=0.95 (95% CI: 0.91, 0.98)	0.01	PR=0.95 (95% CI: 0.91, 0.98)
Any MST Marker		346	14	359	9	728	PR=0.99 (95% CI: 0.96, 1.01)	0.26	PR=0.99 (95% CI: 0.96, 1.01)
Any MST Marker		145	102	148	98	493	PR=0.98 (95% CI: 0.82, 1.16)	0.78	PR=0.97 (95% CI: 0.82, 1.15)
Any human MST Marker	Any sample	17	3	30	7	57	PR=1.05 (95% CI: 0.82, 1.34)	0.71	PR=1.05 (95% CI: 0.82, 1.34)
Any human MST Marker	Any sample	41	28	43	18	130	PR=0.84 (95% CI: 0.63, 1.12)	0.24	PR=0.89 (95% CI: 0.67, 1.18)
Any human MST Marker	Any sample	124	313	133	330	900	PR=0.99 (95% CI: 0.8, 1.22)	0.91	PR=1.01 (95% CI: 0.82, 1.25)
Any human MST Marker	Any sample	26	222	26	223	497	PR=1 (95% CI: 0.57, 1.75)	0.99	PR=1 (95% CI: 0.57, 1.76)
Any human MST Marker	Any sample	22	8	21	9	60	PR=1.05 (95% CI: 0.76, 1.45)	0.78	
Any human MST Marker	Source water	1	21	0	19	41	Not estimated		Not estimated
Any human MST Marker	Source water	22	8	21	9	60	PR=1.05 (95% CI: 0.76, 1.45)	0.78	
Any human MST Marker	Stored water	9	39	5	41	94	PR=1.72 (95% CI: 0.57, 5.18)	0.33	PR=1.72 (95% CI: 0.57, 5.18)
Any human MST Marker	Stored water	5	310	12	324	651	PR=0.44 (95% CI: 0.16, 1.23)	0.12	PR=0.44 (95% CI: 0.16, 1.23)
Any human MST Marker	Stored water	0	245	0	246	491	Not estimated		Not estimated
Any human MST Marker	House soil	20	19	26	18	83	PR=0.87 (95% CI: 0.57, 1.32)	0.5	PR=0.86 (95% CI: 0.6, 1.24)
Any human MST Marker	House soil	68	243	59	261	631	PR=1.19 (95% CI: 0.87, 1.61)	0.28	PR=1.24 (95% CI: 0.91, 1.7)

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any human MST Marker	House soil	21	226	23	226	496	PR=0.92 (95% CI: 0.5, 1.71)	0.79	PR=0.94 (95% CI: 0.5, 1.75)
Any human MST Marker	Latrine soil	21	9	22	8	60	PR=0.95 (95% CI: 0.69, 1.32)	0.78	PR=0.95 (95% CI: 0.69, 1.32)
Any human MST Marker		24	7	38	17	86	PR=1.12 (95% CI: 0.83, 1.51)	0.46	PR=1.02 (95% CI: 0.75, 1.41)
Any human MST Marker		30	142	44	122	338	PR=0.66 (95% CI: 0.44, 0.99)	0.04	PR=0.72 (95% CI: 0.48, 1.07)
Any human MST Marker		58	268	60	265	651	PR=0.96 (95% CI: 0.68, 1.37)	0.84	PR=0.96 (95% CI: 0.68, 1.35)
Any human MST Marker		7	240	5	241	493	PR=1.39 (95% CI: 0.46, 4.2)	0.56	PR=1.39 (95% CI: 0.46, 4.2)
Any animal MST Marker	Any sample	12	8	17	20	57	PR=1.31 (95% CI: 0.78, 2.17)	0.3	PR=1.2 (95% CI: 0.72, 1.99)
Any animal MST Marker	Any sample	3	66	2	59	130	PR=1.33 (95% CI: 0.18, 9.59)	0.78	PR=1.33 (95% CI: 0.18, 9.59)
Any animal MST Marker	Any sample	419	26	437	28	910	PR=1 (95% CI: 0.97, 1.04)	0.91	PR=1 (95% CI: 0.97, 1.04)
Any animal MST Marker	Any sample	219	29	221	28	497	PR=0.99 (95% CI: 0.93, 1.06)	0.88	PR=0.99 (95% CI: 0.93, 1.06)
Any animal MST Marker	Any sample	28	2	27	3	60	PR=1.04 (95% CI: 0.89, 1.21)	0.65	
Any animal MST Marker	Source water	0	22	0	19	41	Not estimated		Not estimated
Any animal MST Marker	Source water	28	2	27	3	60	PR=1.04 (95% CI: 0.89, 1.21)	0.65	
Any animal MST Marker	Stored water	0	48	1	45	94	Not estimated		Not estimated
Any animal MST Marker	Stored water	229	113	253	109	704	PR=0.96 (95% CI: 0.86, 1.07)	0.43	PR=0.97 (95% CI: 0.87, 1.08)

Target	Sample	Positive, Intervention	Negative, Intervention	Positive, Control	Negative, Control	Total observations	Unadjusted Prevalence Ratio	Unadjusted p-value	Adjusted Prevalence Ratio
Any animal MST Marker	Stored water	57	188	82	164	491	PR=0.7 (95% CI: 0.51, 0.96)	0.03	PR=0.69 (95% CI: 0.5, 0.95)
Any animal MST Marker	House soil	2	37	1	43	83	Not estimated		Not estimated
Any animal MST Marker	House soil	281	30	291	29	631	PR=0.99 (95% CI: 0.94, 1.05)	0.82	PR=0.99 (95% CI: 0.94, 1.04)
Any animal MST Marker	House soil	178	69	186	63	496	PR=0.96 (95% CI: 0.86, 1.08)	0.53	PR=0.96 (95% CI: 0.86, 1.08)
Any animal MST Marker	Latrine soil	2	28	0	30	60	Not estimated		Not estimated
Any animal MST Marker		12	19	18	37	86	PR=1.18 (95% CI: 0.7, 2)	0.53	PR=1.33 (95% CI: 0.62, 2.86)
Any animal MST Marker		174	8	182	1	365	PR=0.96 (95% CI: 0.93, 1)	0.03	PR=0.96 (95% CI: 0.93, 1)
Any animal MST Marker		344	15	358	9	726	PR=0.98 (95% CI: 0.96, 1.01)	0.17	PR=0.98 (95% CI: 0.96, 1.01)
Any animal MST Marker		144	103	147	99	493	PR=0.98 (95% CI: 0.82, 1.16)	0.78	PR=0.97 (95% CI: 0.82, 1.15)

.	Boehm 2016	Reese 2017	Steinbaum 2019	Fuhrmeister 2020	Holcomb 2021	Capone 2021	Capone 2022 in prep.	Kwong 2021
5-8	199 (40.0%)	171 (70.7%)	1149 (54.5%)	224 (37.5%)	44 (27.0%)	7 (8.0%)	3 (5.3%)	528 (37.8%)
>8	27 (5.4%)	54 (22.3%)	245 (11.6%)	38 (6.4%)	81 (49.7%)	81 (92.0%)	54 (94.7%)	85 (6.1%)
Missing	0 (0%)	0 (0%)	101 (4.8%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Number of rooms in the household								
1-2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	98 (60.1%)	61 (69.3%)	41 (71.9%)	0 (0%)
>3	0 (0%)	0 (0%)	0 (0%)	0 (0%)	65 (39.9%)	27 (30.7%)	16 (28.1%)	0 (0%)
Missing	497 (100%)	242 (100%)	2107 (100%)	597 (100%)	0 (0%)	0 (0%)	0 (0%)	1396 (100%)
Improved roof								
0	8 (1.6%)	0 (0%)	693 (32.9%)	8 (1.3%)	0 (0%)	0 (0%)	0 (0%)	23 (1.6%)
1	489 (98.4%)	0 (0%)	1414 (67.1%)	589 (98.7%)	0 (0%)	0 (0%)	0 (0%)	1373 (98.4%)
Missing	0 (0%)	242 (100%)	0 (0%)	0 (0%)	163 (100%)	88 (100%)	57 (100%)	0 (0%)
Father in agriculture								
0	332 (66.8%)	126 (52.1%)	0 (0%)	419 (70.2%)	0 (0%)	0 (0%)	0 (0%)	952 (68.2%)
1	165 (33.2%)	89 (36.8%)	0 (0%)	178 (29.8%)	0 (0%)	0 (0%)	0 (0%)	444 (31.8%)
Missing	0 (0%)	27 (11.2%)	2107 (100%)	0 (0%)	163 (100%)	88 (100%)	57 (100%)	0 (0%)
Land owned								
0	0 (0%)	97 (40.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
1	0 (0%)	117 (48.3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Missing	497 (100%)	28 (11.6%)	2107 (100%)	597 (100%)	163 (100%)	88 (100%)	57 (100%)	1396 (100%)
Acres of land owned								
Mean (SD)	0.110 (0.128)	NA (NA)	NA (NA)	0.150 (0.206)	NA (NA)	NA (NA)	NA (NA)	0.142 (0.212)
Median [Min, Max]	0.0700 [0.0100, 1.23]	NA [NA, NA]	NA [NA, NA]	0.0800 [0.0100, 2.10]	NA [NA, NA]	NA [NA, NA]	NA [NA, NA]	0.0800 [0.0100, 3.15]

.	Boehm 2016	Reese 2017	Steinbaum 2019	Fuhrmeister 2020	Holcomb 2021	Capone 2021	Capone 2022 in prep.	Kwong 2021
Missing	13 (2.6%)	242 (100%)	2107 (100%)	21 (3.5%)	163 (100%)	88 (100%)	57 (100%)	62 (4.4%)
Maternal education								
No education	85 (17.1%)	0 (0%)	0 (0%)	86 (14.4%)	6 (3.7%)	0 (0%)	0 (0%)	207 (14.8%)
Incomplete Primary	0 (0%)	83 (34.3%)	1095 (52.0%)	0 (0%)	38 (23.3%)	0 (0%)	0 (0%)	0 (0%)
Primary	180 (36.2%)	30 (12.4%)	511 (24.3%)	183 (30.7%)	14 (8.6%)	0 (0%)	0 (0%)	449 (32.2%)
Secondary	232 (46.7%)	70 (28.9%)	499 (23.7%)	328 (54.9%)	41 (25.2%)	0 (0%)	0 (0%)	740 (53.0%)
More than secondary	0 (0%)	11 (4.5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Missing	0 (0%)	48 (19.8%)	2 (0.1%)	0 (0%)	64 (39.3%)	88 (100%)	57 (100%)	0 (0%)
Maternal age								
Mean (SD)	23.7 (5.18)	NA (NA)	26.4 (6.32)	23.7 (5.08)	NA (NA)	NA (NA)	NA (NA)	24.0 (5.03)
Median [Min, Max]	23.0 [15.0, 42.0]	NA [NA, NA]	25.5 [14.9, 47.9]	23.0 [15.0, 41.0]	NA [NA, NA]	NA [NA, NA]	NA [NA, NA]	24.0 [15.0, 43.0]
Missing	0 (0%)	242 (100%)	11 (0.5%)	0 (0%)	163 (100%)	88 (100%)	57 (100%)	2 (0.1%)
Improved wall								
0	78 (15.7%)	0 (0%)	2019 (95.8%)	197 (33.0%)	41 (25.2%)	16 (18.2%)	10 (17.5%)	369 (26.4%)
1	419 (84.3%)	0 (0%)	88 (4.2%)	400 (67.0%)	122 (74.8%)	72 (81.8%)	47 (82.5%)	1027 (73.6%)
Missing	0 (0%)	242 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Improved floor								
0	461 (92.8%)	0 (0%)	1999 (94.9%)	524 (87.8%)	4 (2.5%)	1 (1.1%)	1 (1.8%)	1253 (89.8%)
1	36 (7.2%)	0 (0%)	108 (5.1%)	73 (12.2%)	159 (97.5%)	87 (98.9%)	56 (98.2%)	143 (10.2%)
Missing	0 (0%)	242 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Electricity								
0	234 (47.1%)	34 (14.0%)	1958 (92.9%)	246 (41.2%)	3 (1.8%)	4 (4.5%)	2 (3.5%)	584 (41.8%)

.	Boehm 2016	Reese 2017	Steinbaum 2019	Fuhrmeister 2020	Holcomb 2021	Capone 2021	Capone 2022 in prep.	Kwong 2021
1	263 (52.9%)	202 (83.5%)	147 (7.0%)	351 (58.8%)	160 (98.2%)	84 (95.5%)	55 (96.5%)	812 (58.2%)
Missing	0 (0%)	6 (2.5%)	2 (0.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)