Environmental Science & Technology

pubs.acs.org/est

Article

Impacts of an Urban Sanitation Intervention on Fecal Indicators and the Prevalence of Human Fecal Contamination in Mozambique

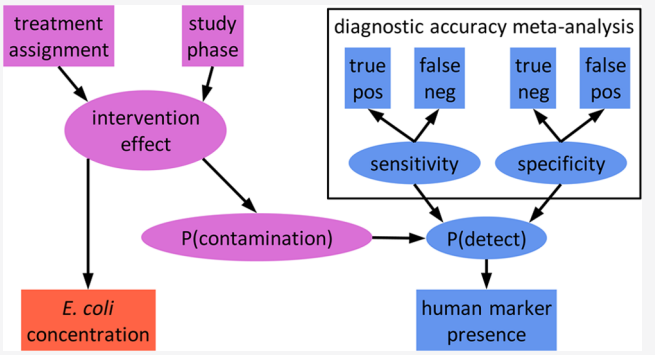
David A. Holcomb, Jackie Knee, Drew Capone, Trent Sumner, Zaida Adriano, Rassul Nalá, Oliver Cumming, Joe Brown, and Jill R. Stewart*

Cite This: https://doi.org/10.1021/acs.est.1c01538

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Fecal source tracking (FST) may be useful to assess pathways of fecal contamination in domestic environments and to estimate the impacts of water, sanitation, and hygiene (WASH) interventions in low-income settings. We measured two nonspecific and two human-associated fecal indicators in water, soil, and surfaces before and after a shared latrine intervention from low-income households in Maputo, Mozambique, participating in the Maputo Sanitation (MapSan) trial. Up to a quarter of households were impacted by human fecal contamination, but trends were unaffected by improvements to shared sanitation facilities. The intervention reduced *Escherichia coli* gene concentrations in soil but did not impact culturable *E. coli* or the prevalence of human FST markers in a difference-in-differences analysis. Using a novel Bayesian hierarchical modeling approach to account for human marker diagnostic sensitivity and specificity, we revealed a high amount of uncertainty associated with human FST measurements and intervention effect estimates. The field of microbial source tracking would benefit from adding measures of diagnostic accuracy to better interpret findings, particularly when FST analyses convey insufficient information for robust inference. With improved measures, FST could help identify dominant pathways of human and animal fecal contamination in communities and guide the implementation of effective interventions to safeguard health.



KEYWORDS: *diagnostic accuracy, water, sanitation, and hygiene, shared sanitation, microbial source tracking, fecal indicator, qPCR, Bayesian hierarchical model*

INTRODUCTION

Water, sanitation, and hygiene (WASH) interventions aim to improve health by preventing exposure to enteric pathogens, which are introduced to the environment in the feces of infected human and animal hosts.[1] Environmental pathways of pathogen exposure include contaminated environmental compartments like water, soil, and surfaces as well as hands, flies, food, and fomites that have been in contact with contaminated environments.[2−4] Recent evaluations of a range of WASH interventions found inconsistent and largely negligible impacts on child diarrhea, growth, and enteric infection.[5−12] Notably, combined interventions did not provide greater protection than their constituent interventions alone, suggesting that key sources of pathogens and pathways of exposure are inadequately addressed by conventional WASH strategies.[6,7,9,13−15]

Characterizing fecal contamination in potential exposure pathways may help explain why specific interventions do or do not improve health by identifying which pathways the intervention interrupts and which remain unaffected. Fecal contamination is typically assessed by measuring fecal indicator organisms, microbes abundant in feces used to infer the presence of fecal contamination and therefore the likely presence of enteric pathogens, which are challenging to measure directly due to their diversity and low environmental concentrations.[15,16] Indicator organisms can also be used for fecal source tracking (FST) by targeting microbes specific to the feces of a particular host. Animals are important sources of fecal contamination in both domestic and public environments, but traditional efforts have focused on preventing exposure to human feces; differentiating between human and various animal feces could inform more appropriate intervention approaches.[4,17−22]

Fecal indicator approaches have increasingly been applied to domestic environments in low-income settings with high burdens of enteric disease.[3,15−18,23−26] Occurrence of non-specific indicators like *Escherichia coli* is challenging to interpret

Received: March 8, 2021
Revised: July 26, 2021
Accepted: July 27, 2021

in these settings due to elevated and highly variable ambient concentrations, possibly from naturalized sources, which are typically assessed in limited numbers of (cross-sectional) observations from each environmental compartment.[16,27−30] Other than ruminant FST markers, host-associated fecal indicators have demonstrated poor diagnostic accuracy in domestic settings.[16,17,26,31,32] The use of multiple FST markers has been proposed to help address the limited accuracy of individual indicators.[33,34] Several studies have calculated the conditional probability of contamination by a specific fecal source given the detection of one or more source-associated indicators in one sample.[31,34−36] Such analyses provide valuable intuition about the uncertainty associated with individual measurements, which can be particularly important in decision-making contexts like beach closures. To our knowledge, diagnostic performance has not been similarly accounted for when FST has been used to infer patterns and predictors of source-specific fecal contamination in domestic environments, likely overstating the confidence of such estimates.[4,17,18,26,37−39]

In this study, we analyze two nonspecific and two human-associated fecal indicators in water, soil, and surfaces from low-income households in Maputo, Mozambique, before and after a shared sanitation intervention. We explore the conditional probability of human fecal contamination in individual samples under different prevalence and indicator detection scenarios and develop a Bayesian hierarchical modeling approach that accounts for the diagnostic accuracy of multiple markers to estimate the prevalence of source-specific fecal contamination. Finally, we implement these models using both human markers to estimate intervention effects on the prevalence of human fecal contamination in multiple exposure pathways.

MATERIALS AND METHODS

Study Setting and Intervention. We characterized fecal contamination of households with children participating in the Maputo Sanitation (MapSan) study (clinicaltrials.gov NCT02362932), a prospective, controlled before and after health impact trial of an urban, onsite sanitation intervention.[40] The intervention was delivered to compounds (self-defined clusters of households sharing outdoor space) in low-income neighborhoods of Maputo, Mozambique, areas with high burdens of enteric disease and predominantly onsite sanitation infrastructure.[41,42] Similar compounds that did not receive the intervention were recruited to serve as control sites. At baseline, both intervention and control compounds shared sanitation facilities in poor condition.[26] The existing shared latrines in intervention compounds were replaced with pour-flush latrines that discharged aqueous effluent to infiltration pits and had sturdy, private superstructures. Intervention latrines were constructed between 2015 and 2016 by the nongovernmental organization (NGO) Water and Sanitation for the Urban Poor (WSUP), which selected intervention sites according to engineering and demand criteria (Table S1).[40]

Data Collection. The intervention impact on fecal contamination was evaluated using a controlled before-and-after (CBA) study design.[5,43] Intervention compounds were enrolled immediately before the new latrine was opened for use, with a concurrent enrollment of control compounds conducted at a similar frequency (Table S1). Follow-up visits to each compound were conducted approximately 12 months following baseline enrollment. We administered compound-, household-, and child-level surveys during both baseline and follow-up visits, as described elsewhere.[5,42] Concurrent with survey administration during May−August 2015, we opportunistically collected environmental samples at a subset of MapSan study compounds from the shared outdoor space and from each household with children participating in the health study [see the Supporting Information (SI)]. During the 12 month follow-up phase in June−September 2016, we revisited the original subset of compounds and collected environmental samples from additional study compounds not sampled at baseline, as time permitted.

Detailed descriptions of environmental sample collection, processing, and analysis have been published previously.[26] Briefly, we assessed fecal indicators in five environmental compartments: compound source water, household stored water, latrine entrance soil, household entrance soil, and household food preparation surfaces (see the SI). Source water and latrine soil were sampled once per compound on each visit, while stored water, food preparation surfaces, and household soil were collected from each household with children enrolled in the health impacts study. Samples were processed by membrane filtration, preceded by manual elution for soil and swab samples, and the sample filters were analyzed for microbial indicators of fecal contamination using both culture- and molecular-based detection.[25,26,44] We enumerated culturable *E. coli* (cEC) from filters on modified mTEC broth (Hi-Media, Mumbai, India) and immediately archived additional filters at −80 °C for molecular analysis.[16,45] Archived filters were analyzed by three locally validated real-time polymerase chain reaction (qPCR) assays targeting fecal microbe genes. The EC23S857 (EC23S) assay targets *E. coli* and served as an indicator of nonspecific fecal contamination, while HF183/BacR287 (HF183) and Mnif both target microbes specific to human feces and served as indicators of human-source fecal contamination.[46−48] Limits of detection for each assay were previously determined using receiver operating characteristic (ROC) analysis to identify optimal quantification cycle (Cq) cutoff values (see SI).[26,49]

DNA was isolated from soil and surface sample filters using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) and from water sample filters with the DNA-EZ ST01 Kit (GeneRite, North Brunswick, NJ), with a positive control (PC) and negative extraction control (NEC) included in each batch of up to 22 sample filters. PCs consisted of a clean filter spiked with 2×10^8 copies of each composite DNA standard (Table S4).[26] Filters were treated with 3 μg salmon testes DNA (MilliporeSigma, Burlington, MA) immediately before extraction as a specimen processing control (SPC) to assess PCR inhibition.[50,51] We tested each extract with four qPCR assays using a CFX96 Touch thermocycler (Bio-Rad, Hercules, CA), three targeting fecal microbes and Sketa22 targeting the salmon DNA SPC, with 10% of each sample type analyzed in duplicate for all microbial targets.[52] Each reaction consisted of 12.5 μL TaqMan Environmental Master Mix 2.0, 2.5 μL 10× primers/probe mix, 5 μL nuclease-free water (NFW), and 5 μL DNA template, for 25 μL total reaction volume. After an initial 10 min, 95 °C incubation period, cycling conditions specified by the original developers were followed for each assay (Table S3). Samples with Sketa22 quantification cycle (Cq) values > 3 above the mean Cq of extraction controls (NEC and PC) were considered inhibited and diluted 1:5 for further analysis. Each plate included three no-template controls (NTCs) and five-point, 10-fold dilution series of three extracted PCs, corresponding to triplicate reactions with 10^5−10^1 or 10^6−

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

10^2 target gene copies (gc). Target concentrations were estimated from calibration curves fit to the standard dilution series using multilevel Bayesian regression with varying slopes and intercepts by extraction batch and instrument run (see the SI).[53] Fecal indicator concentrations were $\log_{10}$ transformed and expressed as $\log_{10}$ colony forming units (cfu) or gc per 100 mL water, 100 cm^2 surface, or 1 dry gram soil.

Estimating Intervention Effects. We used a difference-in-differences (DID) approach to estimate the effect of the intervention on fecal indicator occurrence. DID enables unbiased estimation of the treatment effect in the absence of randomization, including when different samples of each group are observed pre- and post-treatment, under the "parallel trend" assumption that all unmeasured time-varying covariates related to the outcome are constant across treatment groups and that unmeasured covariates varying between treatment groups are constant through time.[43,54,55] Although we estimated gene copy concentrations for all fecal indicators assessed by qPCR, we treated the human markers as binary, diagnostic tests of the presence or absence of human fecal contamination due to their relatively low baseline detection frequency (and limited availability of concentration data as a result).[26] By contrast, *E. coli* was detected in the large majority of baseline samples by both culture and qPCR approaches; treating such outcomes as presence/absence would discard a great deal of information conveyed by the *E. coli* concentration measurements, producing a binary outcome with very little variation. Direct DID estimates for the mean concentration of nonspecific indicators and the prevalence of human-associated indicators were obtained using a bootstrap approach with 2000 samples. We calculated the mean concentration or prevalence in each of the four design strata (pretreatment intervention compounds, post-treatment intervention compounds, pretreatment control compounds, and post-treatment controls) by sample type, from which the DID was calculated directly (see the SI). Bootstrap 95% compatibility intervals (CI) were obtained as the 2.5 and 97.5 percentile values of the bootstrap samples.[56]

We also conducted regression analyses incorporating potential confounding variables to obtain conditional DID estimates. We used the product-term representation of the DID estimator, in which binary indicators of treatment group, study phase, and their product (interaction) were included as linear predictors. The coefficient on the product term provides the conditional DID estimate.[54,57] Separate models were fit for each combination of fecal indicator and sample type using Bayesian multilevel models with compound-varying intercepts. Censored linear regression was used to estimate the intervention impact on the $\log_{10}$ concentration of nonspecific indicators and the effect of the intervention on human-associated indicator prevalence was estimated using logistic regression and the prevalence odds ratio (POR) as the measure of effect.[58,59] Models were fit with the package **brms** in **R** version 4.0.2 using 1500 warmup and 1000 sampling iterations on four chains (see the SI for prior distributions).[58,60] Estimates of the intervention effect were summarized by the mean and central 95% CI of the resulting 4000 posterior draws.

Adjusted models included variables for selected compound, household, meteorological, and sample characteristics. Compound population density, presence of domestic animals, and asset-based household wealth scores were derived from household and compound surveys administered during each study phase.[42,61] Previous-day mean temperature and 7-day antecedent rainfall were drawn from daily summary records for a local weather station. For stored water samples, we considered whether the storage container was covered and if the mouth was wide enough to admit hands. The surface material was considered for food surface swabs, and for soil samples, we accounted for sun exposure and visibly wet soil surfaces. Covariate data sources and processing have been described previously.[26,42]

Conditional Probability Analysis. Both HF183 and Mnif were previously found to frequently misdiagnose human feces in our study area.[26] An indicator's diagnostic accuracy is described by its sensitivity (Se), the probability of detecting the indicator when contamination is present, and specificity (Sp), the probability of not detecting the indicator when contamination is absent. The probability that a positive sample is contaminated depends on the marker sensitivity and specificity and the prevalence of human fecal contamination. This marginal probability of contamination can be approximated as the frequency of indicator detection among all samples to explore indicator reliability in a specific study.[31] We assessed the probability that human feces were present in an environmental sample in which HF183 or Mnif was detected using Bayes' Theorem and the local sensitivity and specificity of the two markers (see the SI).[34−36] We calculated the conditional probability of contamination for HF183 and Mnif separately and for each combination of the two indicators by sample type. The marginal probability of contamination was approximated as the detection frequency of HF183 among all samples of a given type.

Accounting for Diagnostic Accuracy. Fecal indicator measurements are used as proxies for unobserved fecal contamination to estimate its prevalence and associations of interest, such as the effects of mitigation practices. This approach is vulnerable to measurement error, illustrated by the limited diagnostic accuracy of many host-associated fecal indicators.[16] Bias due to inaccurate diagnostic tests can be mitigated by incorporating external information on the sensitivity and specificity of the test.[62] The expected detection frequency, p, of a test with sensitivity Se and specificity Sp is given by

$$p = \mathrm{Se} \times \pi + (1 - \mathrm{Sp})(1 - \pi) \quad (1)$$

for an underlying condition with prevalence π.[62,63] We adapted the approach of Gelman and Carpenter to estimate the intervention effect on human fecal contamination prevalence from observations of human-associated fecal indicators by incorporating external information on indicator performance within a Bayesian hierarchical framework.[63] We included the product-term representation of the DID estimator and other covariates as linear predictors of the prevalence log-odds. Assuming indicator detection in the ith of n samples, y_i, was Bernoulli distributed with probability p_i, where p_i was related to the prevalence as shown in eq 1, the accuracy-adjusted prevalence model was

$$\begin{aligned} y_i &\sim \mathrm{Bernoulli}(p_i) \\ p_i &= \mathrm{Se} \times \pi_i + (1 - \mathrm{Sp})(1 - \pi_i) \\ \mathrm{logit}(\pi_i) &= \beta^0 + \beta^P P_i + \beta^T T_i + \beta^{\mathrm{DID}} P_i \times T_i + X_i \boldsymbol{\gamma} \end{aligned} \quad (2)$$

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Characteristics of Maputo Sanitation Study Compounds and Households Selected for Environmental Sampling, Samples Collected, and Sampling Dates, Stratified by Study Phase and Treatment Arm

			before				after			
			control		intervention		control		intervention	
characteristic	level	metric	*N*	summary	*N*	summary	*N*	summary	*N*	summary
animals present	compound	*n* (%)	32	15 (47)	25	17 (68)	30	24 (80)	34	30 (88)
population density (persons/100 m^2)	compound	median (IQR[a])	29	5.5 (3.5)	23	8.1 (5.9)	28	5.9 (4.8)	33	6.7 (4.6)
wealth index (0−100)	household	median (IQR)	51	43 (12)	40	43 (12)	55	45 (19)	52	44 (14)
previous-day mean temperature (°C)	date	median (IQR)	19	21 (2)	16	20 (2)	17	20 (1)	17	21 (3)
seven-day cumulative precipitation (mm)	date	median (IQR)	19	9 (3)	16	14 (3)	17	13 (39)	17	7 (0)
water container covered	sample	*n* (%)	44	25 (57)	28	21 (75)	38	21 (55)	47	30 (64)
narrow-mouth water container	sample	*n* (%)	44	13 (30)	28	10 (36)	38	13 (34)	47	14 (30)
plastic food surface material	sample	*n* (%)	34	30 (88)	23	18 (78)	29	26 (90)	36	29 (81)
shaded latrine soil	sample	*n* (%)	32	24 (75)	17	12 (71)	30	25 (83)	30	22 (73)
shaded household soil	sample	*n* (%)	42	31 (74)	28	24 (86)	35	32 (91)	39	31 (79)
wet latrine soil surface	sample	*n* (%)	32	20 (62)	17	13 (76)	30	18 (60)	30	21 (70)
wet household soil surface	sample	*n* (%)	42	24 (57)	27	13 (48)	35	18 (51)	39	13 (33)
latrine soil moisture (%)	sample	median (IQR)	33	9.8 (9.8)	23	8.4 (7.2)	30	10.0 (7.9)	30	8.7 (8.3)
household soil moisture (%)	sample	median (IQR)	49	9.9 (8.6)	35	6.9 (6.1)	47	7.8 (5.4)	43	5.4 (5.9)

[a]Interquartile range.

where β^0 is the intercept; β^P, β^T, and β^{DID} are the parameters corresponding to indicators for study phase (P), treatment group (T), and their product, respectively; and $\boldsymbol{\gamma}$ is a $p \times 1$ vector of regression coefficients corresponding to p additional covariates in the $n \times p$ matrix $\boldsymbol{X}$.

We fit three models that differed by definition of Se and Sp. In the simplest case (model 1), we assumed a perfectly accurate test with Se = Sp = 1, thus $p = \pi$. The second model (model 2) incorporated observations from the local validation analysis by assuming

$$y^{Se} \sim \text{binomial}(n^{Se}, \text{Se})$$
$$y^{Sp} \sim \text{binomial}(n^{Sp}, \text{Sp}) \quad (3)$$

for y^{Se} positive results in n^{Se} human fecal samples and y^{Sp} negative results in n^{Sp} nonhuman fecal samples. Because our validation sample set was small and performance estimates vary widely between studies, we fit a third model (model 3) featuring a meta-analysis of indicator sensitivity and specificity (see the SI). We assumed that the log-odds of the sensitivity in the kth study, $\text{Se}_{[k]}$, were normally distributed with mean μ^{Se} and SD σ^{Se}, such that

$$y^{Se}_{[k]} \sim \text{binomial}(n^{Se}_{[k]}, \text{Se}_{[k]})$$
$$\text{logit}(\text{Se}_{[k]}) \sim \text{normal}(\mu^{Se}, \sigma^{Se}) \quad (4)$$

with an equivalent structure for the specificity. We assigned $k = 1$ to our local validation study, using $\text{Se}_{[1]}$ and $\text{Sp}_{[1]}$ as the values of Se and Sp in eq 2.[26,63] This emphasized the local performance data while allowing information from other settings to influence the estimates through partial pooling, with the extent of pooling learned from the data (expressed through σ^{Se} and σ^{Sp}).[59]

Modeling Latent Human Fecal Contamination. Fecal contamination can be understood as a latent environmental condition for which fecal indicators serve as imperfect diagnostic tests.[64,65] Information from multiple fecal indicators may be utilized by modeling each as arising from the same underlying contamination to potentially improve inference. We extended the meta-analytic model (model 3) to include observations of both HF183 and Mnif in the same samples (model 4), with separate detection probabilities, $^{hf}_{pi}$ and $^{mn}_{pi}$, obtained from indicator-specific sensitivity and specificity estimates applied to the same underlying prevalence, π_i. As in previous models, the DID estimator and other predictor variables were included in a linear model on the log-odds of π_i, assuming that intervention effects and other covariates acted directly on the latent prevalence.

As environmental compartments from the same compound share sources of fecal exposure, we extended the previous model to simultaneously consider observations of latrine soil, household soil, and stored water in each compound (model 5). Sample type-specific prevalence variables, $\pi_i^{[type]}$, were modeled as linear deviations from a latent compound-level prevalence π_j on the log-odds scale

$$\text{logit}(\pi_i^{[type]}) = \alpha^{[type]} + X_i^{[type]}\boldsymbol{\gamma}^{[type]} + \text{logit}(\pi_{[j]}^{comp})$$
$$\text{logit}(\pi_{[j]}^{comp}) = \alpha_{[j]}^{comp} + \beta^P P_{[j]} + \beta^T T_{[j]} + \beta^{DID} P_{[j]} \times T_{[j]} + X_{[j]}^{comp}\boldsymbol{\gamma}^{comp}$$
$$\alpha_{[j]}^{comp} \sim \text{normal}(\mu^{comp}, \sigma^{comp})$$
$$\alpha^{[type]} \sim \text{normal}(0, \sigma^{type}) \quad (5)$$

for sample i of a given type (latrine soil, household soil, or stored water) in compound j, where $\alpha_{[j]}^{comp}$ is a compound-varying intercept and $\alpha^{[type]}$ is a varying intercept by sample type. Compound-level predictors, including the DID estimator terms, were placed on the compound-prevalence log-odds.[63,66] Parameters for sample-level and meteorological predictors in $\boldsymbol{X}_i^{[type]}$ were estimated separately for each sample type.

We coded each model in the probabilistic programming language **Stan** and fit the models using the **RStan** interface with four chains of 1000 warmup and 1000 sampling iterations each, for a total of 4000 posterior samples (see the SI for Stan code and discussion of prior distributions).[67,68] Models 1−3 were fit separately for HF183 and Mnif in each sample type (latrine entrance soil, household entrance soil, and stored water), model 4 was fit separately to each sample type, and a single Model 5 fit was produced incorporating both indicators and all sample types. In addition to the DID POR given by the product-term parameter, we used the posterior predictive

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

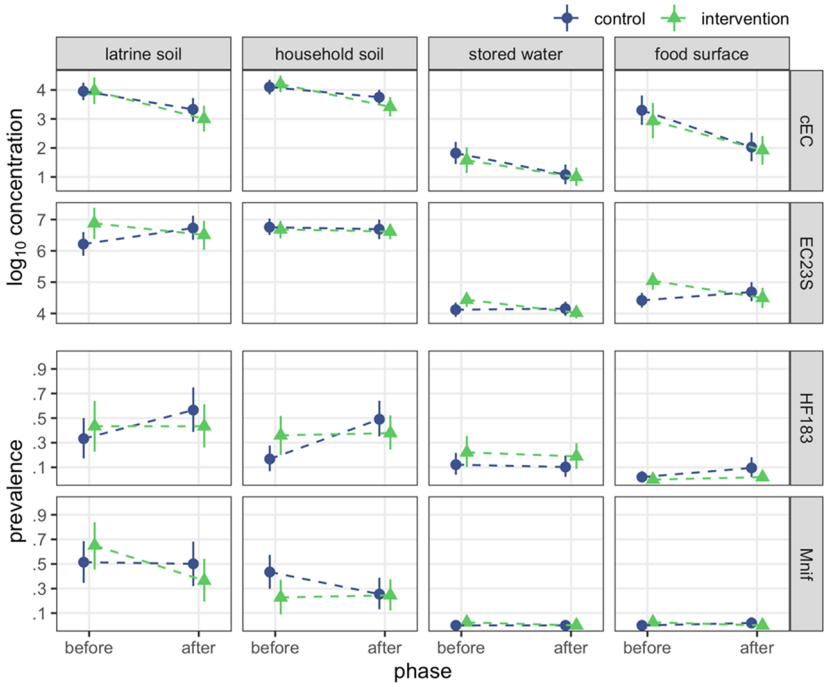


Figure 1. Bootstrap estimates of fecal indicator occurrence by study phase and treatment arm. Points indicate mean $\log_{10}$ concentration for *E. coli* indicators and prevalence of human-associated indicators, with bars presenting bootstrap 95% CIs.

distribution to estimate the prevalence of human fecal contamination in each stratum and to directly calculate DID on the probability scale.[59,69] Models were adjusted for the same covariates as DID regression models.

Ethical Approval. This study was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill (IRB no. 15-0963), and the associated health study was approved by the Comité Nacional de Bioética para a Saúde (CNBS), Ministério da Saúde, Republic of Mozambique (333/CNBS/14), the Ethics Committee of the London School of Hygiene and Tropical Medicine (reference no. 8345), and the Institutional Review Board of the Georgia Institute of Technology (protocol no. H15160). Environmental samples were only collected from households with enrolled children for whom written, informed parental or guardian consent had been given.

RESULTS

Sample Characteristics. We collected a total of 770 environmental samples from 507 unique locations at 139 households in 71 compounds. Samples were collected both pre- and post-intervention at 263 locations (52%), for a total of 526 paired samples and 244 unpaired samples (Table S2). Characteristics expected to confound the relationship between sanitation and fecal contamination were largely similar between treatment arms during each study phase (Table 1). Cumulative precipitation was higher on average in intervention compounds at baseline and in control compounds at follow-up. Water storage containers were also more frequently covered in intervention (75%) than control households (57%) at baseline, though the majority of containers were covered in all strata. Soil surfaces were more often visibly wet in control households (51%) than intervention (33%) at follow-up, both of which were lower than at baseline (57 and 48%, respectively). Most food preparation surfaces were plastic, though more often so in control households during both study phases. A higher percentage of compounds from both treatment arms reported owning domestic animals at follow-up (80−88%) than baseline (47−68%), which may be related to differences in the questionnaire between survey phases. Median household wealth was 40−45 on a 100-point index, with higher variance among controls at follow-up. Median compound population density ranged from 5.5 to 8.1 residents/100 m^2.

Fecal Indicator Occurrence. At least one fecal indicator was detected in 94% of samples (720/770) and *E. coli* was detected in 718 samples: by culture in 81% (611/755) and by qPCR in 86% (655/763). Mean cEC concentrations were lower at follow-up for all sample types in both treatment arms, a pattern not observed for EC23S concentrations (Figure 1). Of the 763 samples tested for human-associated indicators, 28% (217) were positive for at least one human marker. Human-associated indicators were common in soils (23−65% prevalence, across treatment groups and study phases), but only HF183 was regularly detected in stored water (10−22%), and both indicators were rare on food surfaces (0−9%). qPCR calibration curves (Table S5), detection limits (Table S6), and the results of laboratory quality controls are presented in the SI.

Bootstrap DID estimates suggest the intervention reduced EC23S concentrations on food preparation surfaces and HF183 prevalence in household soil but minimally impacted fecal indicator occurrence in other sample types (Table S7). Notably, HF183 prevalence in household soil was similar among intervention households in both study phases but increased among control compounds at follow-up. By contrast, model-based DID estimates, adjusted for potential confounding, were consistent with no intervention effect on food preparation surface EC23S concentration or household soil HF183 prevalence (Table S8). Adjusted models instead indicate the intervention reduced latrine soil concentrations of EC23S [mean difference: −1.2 (95% CI: −2.1, −0.30) $\log_{10}$ gc/dry g]. Although several sample characteristics were

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

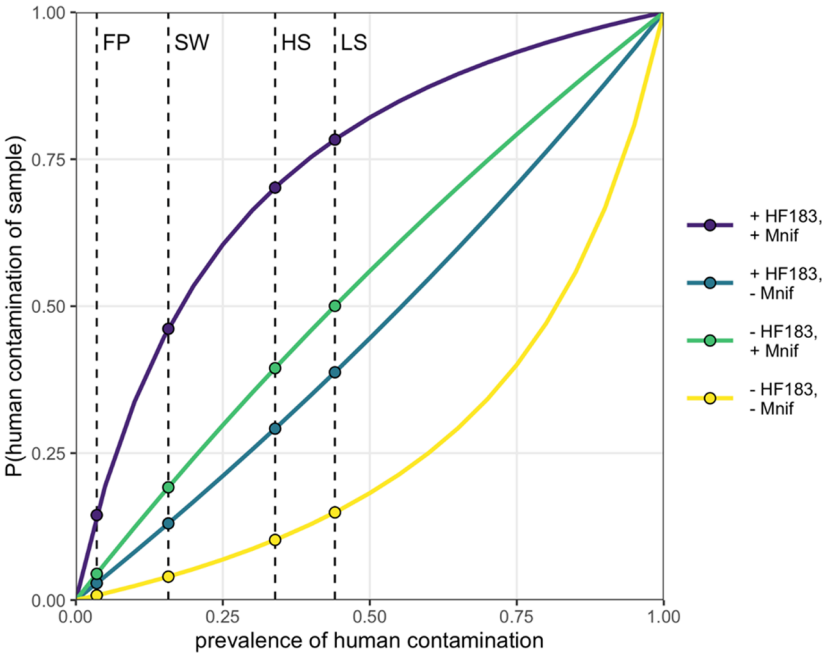


Figure 2. Conditional probability of sample contamination with human feces given detection status of both HF183 and Mnif for all values of human contamination prevalence. Values of sensitivity and specificity were obtained using human and animal feces from the study area, and are 64 and 67%, respectively, for HF183 and 71 and 70% for Mnif. The dashed vertical lines indicate the HF183 detection frequency for each sample type to illustrate relevant human contamination probabilities. FP: food preparation surfaces; SW: stored water; HS: household entrance soil; and LS: latrine entrance soil.

imbalanced between treatment arms and study phases (Table 1), estimates from models that adjusted for these variables were largely similar to the unadjusted models, with adjusted estimates marginally closer to the null in most cases (Table S8). EC23S concentrations in latrine soil were again the exception, with a substantially larger reduction obtained under the adjusted model than the unadjusted estimate of −0.84 (95% CI: −1.6, −0.02) $\log_{10}$ gc/dry g. Due to low detection frequency, models were not fit for either human marker on food surfaces or for Mnif in stored water; source water samples were excluded from all analyses.[26]

Conditional Probability of Human Fecal Contamination. The probability that a sample is contaminated with human feces given the detection of a human indicator is a function of the indicator's sensitivity and specificity (Table S9) and the prevalence of human contamination in the study environment. At 15% prevalence (approximately the detection frequency of HF183 in stored water), the probability of human contamination given a positive test was 26% for HF183 and 30% for Mnif. Only with prevalence above 30−35% was detecting either indicator more likely than not to correctly diagnose human fecal contamination. Combining test results from both indicators improved identification of human contamination, increasing the probability of contamination to 45% when both markers were positive and the prevalence was 15% (Figure 2). However, the two human markers frequently disagreed when assessed in the same sample, conflicting in 44% of household soil, 43% of latrine soil, and 15% of stored water samples. Furthermore, at 44% prevalence (the highest detection frequency for HF183, observed in latrine soils), there remained a >20% chance that a sample positive for both indicators was not contaminated. Among lower-prevalence sample types, the conditional probability never reached 50%. Unless the background prevalence in the study area was about 45% or greater, it is unlikely that the use of HF183 and Mnif reliably identified human contamination in individual samples, particularly given the frequent disagreement between the two markers.

Prevalence of Human Fecal Contamination. Posterior predictions from each of the five accuracy-adjusted models were used to estimate stratum-specific prevalence of human fecal contamination. To compare treatment assignments and study phases, we predicted prevalence for compounds with no animals or antecedent precipitation and the sample mean population density (7 persons/100 m^2), wealth score (46), and previous-day temperature (20.4 °C), in which soil surfaces were dry and shaded and water storage containers possessed wide, uncovered mouths. The prevalence estimates were notably imprecise; the 95% CI of the HF183 prevalence in post-treatment latrine soil ranged from 3 to 92% for model 2 (Table 2). The 95% CI widths were similar for model 1 and the bootstrap estimates but were substantially wider for the other four models, which accounted for FST marker sensitivity and specificity (see the SI). The intervals narrowed somewhat when both indicators were considered (model 4) and narrowed further when all sample types were incorporated (model 5) but were still wider than the estimates that did not account for diagnostic accuracy.

Although we did not formally assess the pairwise differences between prevalence estimates, the wide and largely overlapping posterior predictive CIs indicate a limited ability to distinguish between prevalence estimates between different strata or models. The DID estimates on the probability scale were strongly consistent with no effect for all model specifications, which further suggests that the available data were insufficient to assess prevalence differences between strata. The corresponding prevalence odds ratio estimates obtained directly from the DID product term were likewise imprecise (Figure

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 2. Bootstrap and Adjusted Model-Based Estimates of Human Marker Sensitivity and Specificity, Prevalence of Human Fecal Contamination Stratified by Treatment Arm and Study Phase, and Effect of the Sanitation Intervention on Human Fecal Contamination Prevalence in Soil and Water from Mapsan Study Compounds

					prevalence estimate (95% CI)[a]				
					control		intervention		
	marker	sensitivity (95% CI)	specificity (95% CI)	*N*	before	after	before	after	prevalence DID[b] (95% CI)
					Latrine Soil				
bootstrap	HF183	1	1	116	0.33 (0.17, 0.50)	0.57 (0.39, 0.75)	0.43 (0.23, 0.64)	0.43 (0.26, 0.61)	−0.23 (−0.60, 0.14)
	Mnif	1	1	116	0.51 (0.35, 0.69)	0.50 (0.32, 0.68)	0.65 (0.45, 0.84)	0.36 (0.19, 0.54)	−0.27 (−0.63, 0.08)
model 1[c]	HF183	1	1	98	0.32 (0.17, 0.49)	0.42 (0.24, 0.60)	0.32 (0.15, 0.52)	0.37 (0.20, 0.57)	−0.04 (−0.22, 0.13)
	Mnif	1	1	98	0.44 (0.27, 0.63)	0.37 (0.20, 0.55)	0.43 (0.24, 0.65)	0.27 (0.13, 0.45)	−0.09 (−0.27, 0.07)
model 2[d]	HF183	0.60 (0.42, 0.79)	0.66 (0.53, 0.80)	98	0.38 (0.05, 0.88)	0.40 (0.05, 0.90)	0.38 (0.05, 0.89)	0.39 (0.03, 0.92)	−0.01 (−0.19, 0.18)
	Mnif	0.64 (0.47, 0.82)	0.66 (0.51, 0.81)	98	0.48 (0.09, 0.90)	0.44 (0.07, 0.90)	0.47 (0.07, 0.90)	0.39 (0.05, 0.92)	−0.04 (−0.25, 0.15)
model 3[e]	HF183	0.65 (0.45, 0.85)	0.68 (0.55, 0.82)	98	0.34 (0.05, 0.83)	0.37 (0.05, 0.85)	0.34 (0.04, 0.85)	0.36 (0.04, 0.88)	−0.01 (−0.19, 0.18)
	Mnif	0.70 (0.56, 0.83)	0.72 (0.58, 0.85)	98	0.49 (0.14, 0.84)	0.43 (0.11, 0.83)	0.47 (0.13, 0.84)	0.35 (0.07, 0.82)	−0.06 (−0.27, 0.13)
model 4[f]	HF183	0.64 (0.47, 0.82)	0.71 (0.57, 0.84)	98	0.39 (0.11, 0.73)	0.37 (0.10, 0.73)	0.37 (0.10, 0.74)	0.29 (0.07, 0.68)	−0.06 (−0.25, 0.11)
	Mnif	0.71 (0.58, 0.84)	0.71 (0.57, 0.84)						
model 5[g]	HF183	0.72 (0.57, 0.87)	0.85 (0.78, 0.91)	98	0.34 (0.12, 0.65)	0.35 (0.13, 0.65)	0.29 (0.08, 0.63)	0.28 (0.08, 0.60)	−0.02 (−0.17, 0.14)
	Mnif	0.71 (0.59, 0.83)	0.78 (0.68, 0.86)						
					Household Soil				
bootstrap	HF183	1	1	176	0.17 (0.07, 0.28)	0.49 (0.35, 0.64)	0.36 (0.20, 0.52)	0.38 (0.24, 0.52)	−0.30 (−0.57, −0.01)
	Mnif	1	1	175	0.43 (0.30, 0.57)	0.25 (0.13, 0.39)	0.23 (0.09, 0.37)	0.24 (0.12, 0.38)	0.20 (−0.07, 0.46)
model 1	HF183	1	1	147	0.26 (0.15, 0.41)	0.43 (0.27, 0.58)	0.29 (0.15, 0.46)	0.41 (0.26, 0.58)	−0.04 (−0.21, 0.12)
	Mnif	1	1	146	0.37 (0.23, 0.52)	0.27 (0.15, 0.42)	0.27 (0.14, 0.43)	0.18 (0.09, 0.31)	0.01 (−0.13, 0.14)
model 2	HF183	0.60 (0.38, 0.80)	0.72 (0.61, 0.83)	147	0.28 (0.04, 0.73)	0.34 (0.03, 0.80)	0.27 (0.03, 0.74)	0.34 (0.02, 0.83)	0.00 (−0.18, 0.19)
	Mnif	0.57 (0.34, 0.80)	0.73 (0.63, 0.84)	146	0.30 (0.03, 0.78)	0.25 (0.02, 0.76)	0.25 (0.02, 0.77)	0.19 (0.01, 0.77)	−0.01 (−0.18, 0.14)
model 3	HF183	0.66 (0.43, 0.85)	0.74 (0.63, 0.85)	147	0.25 (0.04, 0.63)	0.33 (0.04, 0.74)	0.25 (0.03, 0.69)	0.33 (0.03, 0.80)	0.00 (−0.18, 0.20)
	Mnif	0.68 (0.50, 0.82)	0.76 (0.67, 0.86)	146	0.26 (0.03, 0.60)	0.20 (0.03, 0.52)	0.20 (0.02, 0.50)	0.13 (0.02, 0.40)	−0.01 (−0.16, 0.11)
model 4	HF183	0.69 (0.47, 0.87)	0.73 (0.63, 0.83)	146	0.20 (0.04, 0.44)	0.23 (0.03, 0.50)	0.15 (0.03, 0.37)	0.16 (0.02, 0.40)	−0.02 (−0.16, 0.11)
	Mnif	0.68 (0.51, 0.82)	0.75 (0.66, 0.84)						
model 5	HF183	0.72 (0.57, 0.87)	0.85 (0.78, 0.91)	146	0.26 (0.09, 0.49)	0.27 (0.10, 0.51)	0.22 (0.06, 0.47)	0.22 (0.06, 0.45)	−0.01 (−0.16, 0.12)
	Mnif	0.71 (0.59, 0.83)	0.78 (0.68, 0.86)						
					Stored Water				
bootstrap	HF183	1	1	193	0.12 (0.04, 0.22)	0.10 (0.02, 0.20)	0.22 (0.10, 0.35)	0.19 (0.09, 0.30)	−0.01 (−0.21, 0.19)
model 1	HF183	1	1	170	0.23 (0.11, 0.38)	0.19 (0.09, 0.34)	0.28 (0.13, 0.48)	0.24 (0.11, 0.42)	0.00 (−0.14, 0.14)

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 2. continued

					prevalence estimate (95% CI)[a]				
					control		intervention		
	marker	sensitivity (95% CI)	specificity (95% CI)	N	before	after	before	after	prevalence DID[b] (95% CI)
model 2	HF183	0.60	0.85	170	0.15	0.14	0.17	0.16	0.00
		(0.38, 0.81)	(0.78, 0.91)		(0.02, 0.40)	(0.02, 0.38)	(0.02, 0.47)	(0.01, 0.47)	(−0.13, 0.14)
model 3	HF183	0.67	0.86	170	0.15	0.13	0.17	0.16	0.00
		(0.43, 0.85)	(0.79, 0.92)		(0.02, 0.38)	(0.02, 0.36)	(0.02, 0.45)	(0.02, 0.44)	(−0.13, 0.15)
model 5	HF183	0.72	0.85	169	0.19	0.20	0.16	0.16	−0.01
		(0.57, 0.87)	(0.78, 0.91)		(0.04, 0.43)	(0.03, 0.45)	(0.03, 0.40)	(0.02, 0.38)	(−0.14, 0.11)
				Latent Compound					
model 5	HF183	0.72	0.85	109	0.27	0.28	0.22	0.22	−0.01
		(0.57, 0.87)	(0.78, 0.91)		(0.09, 0.52)	(0.09, 0.53)	(0.06, 0.50)	(0.06, 0.47)	(−0.16, 0.13)
	Mnif	0.71	0.78						
		(0.59, 0.83)	(0.68, 0.86)						

[a]All models (excluding bootstrap estimates) were adjusted for population density, presence of animals, wealth score, temperature, antecedent precipitation, and sun exposure and surface wetness for soil samples and storage container mouth width and cover status for water samples. [b]Difference-in-differences. [c]Model 1: single sample type, single marker assuming perfect sensitivity and specificity. [d]Model 2: single sample type, single marker with sensitivity and specificity from a local validation study. [e]Model 3: single sample type, single marker with meta-analytic sensitivity and specificity. [f]Model 4: single sample type, two markers with meta-analytic sensitivity and specificity. [g]Model 5: three sample types, two markers with meta-analytic sensitivity and specificity.

S1). Nonetheless, the model-based prevalence estimates were consistently more similar between the study phase and treatment group than the corresponding bootstrap estimates. This trend was notable for model 5, which assumed that time and treatment effects acted directly on the compound-wide prevalence of human contamination, thus affecting all three sample types equally. The compound-level prevalence estimates were quite similar, particularly between study phases for the same treatment group: 27% (95% CI: 9−52%) at baseline and 28% (9−53%) at follow-up for control compounds and 22% (6−50%) at baseline and 22% (6−47%) at follow-up for intervention compounds. The corresponding estimates for household soil were nearly identical to the compound-level estimates, with somewhat higher estimates for latrine soil and lower for stored water. Although the physical interpretation of this compound-level construct is uncertain, these estimates suggest that about a quarter of compounds were measurably impacted by human fecal contamination, which was unaffected by improvements to shared sanitation facilities.

DISCUSSION

The provision of shared latrines reduced average soil concentrations of the molecular *E. coli* marker EC23S at latrine entrances by more than 1-$\log_{10}$ but did not have a comparable effect on culturable *E. coli*. EC23S latrine soil concentrations rose more in control compounds than they fell in intervention compounds, which under the parallel trend assumption is interpreted as a secular trend upwards that the intervention mitigated, for a much smaller absolute reduction than suggested by the DID estimate (Figure 1).[43] However, an opposite, downward trend was observed for all cEC concentrations. This discrepancy between two tests for the same organism complicates the interpretation of the relatively strong intervention effect estimated for EC23S. While the exact reasons for this discrepancy are yet to be determined, preliminary evidence from a related analysis suggests that the modified mTEC broth used for *E. coli* culture may have produced colonies of the same color and morphology for *Klebsiella* spp., which are commonly soil derived and not specific to feces.[70] By contrast, the developers of EC23S reported 95% specificity to *E. coli* and cross reactions only with other *Escherichia* species, not *Klebsiella*.[46] Accordingly, EC23S potentially better reflected trends in fecal contamination, while cEC may have been confounded by soil microbes more susceptible to environmental conditions, such as the 2016 drought in southern Mozambique.[71]

A cluster-randomized trial in rural Bangladesh likewise found scant evidence of reductions in culturable *E. coli* concentrations from sanitation improvements.[72,73] Latrine provision also did not reduce the prevalence of pathogenic *E. coli* genes in soil, meaning neither culture- nor molecular-based measurements of soil *E. coli* were affected.[39] Other recent trials have not assessed intervention impacts on fecal contamination of soil, but several have evaluated contamination of drinking water, with some also testing child hands, food, or fomites.[15] As with the present study, all found no effect of sanitation-only interventions on any environmental compartment; combined water, sanitation, and hygiene interventions improved drinking water quality in two studies.[13,14]

Measures of human-associated FST markers demonstrated that about a quarter of compounds were impacted by human fecal contamination, with compound-level prevalence estimates not statistically different at baseline and follow-up. Similarly, two cluster-randomized trials, in India and Bangladesh, found no effect of rural sanitation interventions on the prevalence of human-associated indicators in stored drinking water.[37,39] Both studies also assessed human markers in mother and child hand rinse samples, which were not collected in this study. No effect was observed for either hand type in India or on mother hands in Bangladesh, although the human marker prevalence may have been reduced on child hands.[39]

Accounting for the diagnostic accuracy of FST markers revealed far greater uncertainty about host-specific fecal contamination, both of individual samples and population averages, than indicated by raw indicator measurements. The relatively poor sensitivity and specificity of both human markers in this setting severely limited their ability to identify

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

specific samples contaminated with human feces, but even moderate improvements in accuracy could substantially increase FST marker utility. For example, a study in Singapore reported 75% sensitivity and 89% specificity for HF183,[74] corresponding to a 55% chance a positive sample is contaminated at 15% background prevalence and an 84% chance at 44% prevalence, compared with 26 and 60%, respectively, for detection of HF183 in our study. Correcting for indicator sensitivity and specificity to human-source contamination, coupled with the limited observations of each sample type, yielded imprecise prevalence estimates that were consistent with both near absence and almost omnipresence of contamination. While the reduced amplification efficiency of HF183 (82%) may have contributed to its low sensitivity, it produced similar accuracy-corrected estimates as Mnif, which was 95% efficient (Table S5). This imprecision inhibited detecting intervention effects. The point estimates for the intervention effect were relatively close to the null, but the full posterior distributions were consistent with both large reductions and substantial increases in prevalence attributable to the intervention. This analysis does not rule out the possibility that sanitation improvements reduced the prevalence of human fecal contamination. Rather, it strongly suggests that the tools used were inadequate, conveying too little information to address the research question with an acceptable degree of confidence.

These limitations highlight the importance of conducting local validation studies for any new FST application.[75] Accounting for diagnostic accuracy is unlikely to improve the strength or precision of estimates, but may help mitigate overconfidence and overinterpretation by revealing limitations of the available measurements. This practice could also be extended to account for indicator sensitivity and specificity to strictly fecal targets, rather than environmental microbes with nonfecal origins, although we lacked the appropriate data to implement such an analysis for our two nonspecific indicators, EC23S and cEC. As the diagnostic accuracy framework is currently limited to binary outcomes, analysis of such high-prevalence indicators would benefit from the development of analogous approaches for continuous outcomes. Given the intermingling in low-income settings of humans and animals, and their gut microbiomes, alternative FST targets such as mitochondrial DNA could prove more accurate.[76,77] Recent technological advances also present opportunities for new approaches that might bypass the limitations of the current FST paradigm, including portable, long-read sequencing platforms for metagenomic-based source tracking and parallel PCR platforms that render simultaneous analysis of multiple FST markers and comprehensive direct pathogen detection increasingly feasible.[20,78−82] These technologies will also need to overcome the substantial variability, limited analytical sensitivity, and matrix interference characteristic of environmental microbial assessments.[16]

The low signal typical of environmental measurements suggests that study designs—preferably longitudinal—that maximize observations on select pathways of greatest interest should be prioritized to support more robust inference, regardless of analytical approach.[83] A recent longitudinal analysis of *E. coli* concentrations in rural Bangladesh, collected at eight timepoints over 2.5 years from 720 households, demonstrates the advantages of maximizing the number of basic measurements across time. Although pooled estimates from certain sample types achieved statistical significance, the sheer quantity of information available convincingly demonstrated the lack of physically meaningful sanitation intervention impacts on ambient fecal contamination.[73]

Many have speculated that sanitation's apparent lack of effect may be due in part to animal fecal contamination.[12,22] Animal feces often contain pathogens capable of infecting humans and animal fecal biomass in domestic environments is estimated to far exceed that from humans.[22,84−86] Inadequate management of child feces and fecal sludge, contamination of food and water outside the home, and inadequate community-level drainage, solid waste, and sanitation services all present potential pathways of continued contamination despite household sanitation improvements.[24,87−92] Recognizing calls for "transformative" WASH to address these multifarious hazards, sustained progress may require high standards of housing and public services in addition to WASH improvements, necessitating multisectoral coordination and financing.[12,93−95] Even small treatment effects may translate to positive economic benefits.[12] Additionally, quality sanitation infrastructure can provide important benefits irrespective of preventing pathogen exposure, particularly in crowded urban settlements.[96,97] For example, previous research found users of MapSan intervention latrines and similar facilities in the same neighborhoods reported reduced disgust and embarrassment about unhygienic conditions and improved perceptions of security and privacy.[98] Based on the results of our study, we recommend future research to understand the etiology and ecology of fecal pathogens in domestic environments and beyond to help inform interventions needed to construct healthy environments and to protect children's health.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.1c01538.

Site selection criteria; samples collected; qPCR assay details; calibration curves; detection limits; laboratory quality control; conditional probability; difference-in-differences estimates; validation studies; diagnostic accuracy; accuracy-adjusted intervention effect estimates; human fecal contamination prevalence estimates; prior distributions; and model Stan code (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jill R. Stewart − *Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States;* orcid.org/0000-0002-3474-5233; Phone: (919) 966-7553; Email: jill.stewart@unc.edu

Authors

David A. Holcomb − *Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States; Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States;* orcid.org/0000-0003-4055-7164

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Jackie Knee − *School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, U.K.;* orcid.org/0000-0002-0834-8488
Drew Capone − *Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States; School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* orcid.org/0000-0002-2138-6382
Trent Sumner − *School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*
Zaida Adriano − *We Consult, Maputo, Mozambique*
Rassul Nalá − *Instituto Nacional de Saúde, Ministério da Saúde, Maputo, Mozambique*
Oliver Cumming − *Department of Disease Control, London School of Hygiene & Tropical Medicine, London WC1E 7HT, U.K.*
Joe Brown − *Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States;* orcid.org/0000-0002-5200-4148

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.1c01538

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was funded by the United States Agency for International Development under Translating Research into Action, Cooperative Agreement No. GHS-A-00-09-00015-00 and the National Institute of Environmental Health Sciences (T32ES007018). This work was supported, in whole or in part, by the Bill & Melinda Gates Foundation OPP1137224. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission. Deidentified data are publicly available on OSF (https://osf.io/53gwy/). The authors gratefully acknowledge their implementing partner, Water and Sanitation for the Urban Poor; the field research team at WE Consult; the laboratory staff at the Mozambique National Institute of Health; and Olimpio Zavale, Olivia Ginn, and Anna Stamatogiannakis for sampling support.

■ REFERENCES

(1) Julian, T. R. Environmental Transmission of Diarrheal Pathogens in Low and Middle Income Countries. *Environ. Sci.: Processes Impacts* **2016**, *18*, 944−955.
(2) Wagner, E. G.; Laniox, J. N. Excreta Disposal for Rural Areas and Small Communities. *Monogr. Ser. World Health Organ* **1958**, *39*, 1−182.
(3) Sclar, G. D.; Penakalapati, G.; Amato, H. K.; Garn, J. V.; Alexander, K.; Freeman, M. C.; Boisson, S.; Medlicott, K. O.; Clasen, T. Assessing the Impact of Sanitation on Indicators of Fecal Exposure along Principal Transmission Pathways: A Systematic Review. *Int. J. Hyg. Environ. Health* **2016**, *219*, 709−723.
(4) Fuhrmeister, E. R.; Ercumen, A.; Pickering, A. J.; Jeanis, K. M.; Ahmed, M.; Brown, S.; Arnold, B. F.; Hubbard, A. E.; Alam, M.; Sen, D.; Islam, S.; Kabir, M. H.; Kwong, L. H.; Islam, M.; Unicomb, L.; Rahman, M.; Boehm, A. B.; Luby, S. P.; Colford, J. M.; Nelson, K. L. Predictors of Enteric Pathogens in the Domestic Environment from Human and Animal Sources in Rural Bangladesh. *Environ. Sci. Technol.* **2019**, *53*, 10023−10033.
(5) Knee, J.; Sumner, T.; Adriano, Z.; Anderson, C.; Bush, F.; Capone, D.; Casmo, V.; Holcomb, D. A.; Kolsky, P.; MacDougall, A.; Molotkova, E.; Braga, J. M.; Russo, C.; Schmidt, W. P.; Stewart, J.; Zambrana, W.; Zuin, V.; Nalá, R.; Cumming, O.; Brown, J. Effects of an Urban Sanitation Intervention on Childhood Enteric Infection and Diarrhea in Maputo, Mozambique: A Controlled before-and-after Trial. *eLife* **2021**, *10*, No. e62278.
(6) Null, C.; Stewart, C. P.; Pickering, A. J.; Dentz, H. N.; Arnold, B. F.; Arnold, C. D.; Benjamin-Chung, J.; Clasen, T.; Dewey, K. G.; Fernald, L. C. H.; Hubbard, A. E.; Kariger, P.; Lin, A.; Luby, S. P.; Mertens, A.; Njenga, S. M.; Nyambane, G.; Ram, P. K.; Colford, J. M. Effects of Water Quality, Sanitation, Handwashing, and Nutritional Interventions on Diarrhoea and Child Growth in Rural Kenya: A Cluster-Randomised Controlled Trial. *Lancet Glob. Health* **2018**, *6*, e316−e329.
(7) Luby, S. P.; Rahman, M.; Arnold, B. F.; Unicomb, L.; Ashraf, S.; Winch, P. J.; Stewart, C. P.; Begum, F.; Hussain, F.; Benjamin-Chung, J.; Leontsini, E.; Naser, A. M.; Parvez, S. M.; Hubbard, A. E.; Lin, A.; Nizame, F. A.; Jannat, K.; Ercumen, A.; Ram, P. K.; Das, K. K.; Abedin, J.; Clasen, T. F.; Dewey, K. G.; Fernald, L. C.; Null, C.; Ahmed, T.; Colford, J. M. Effects of Water Quality, Sanitation, Handwashing, and Nutritional Interventions on Diarrhoea and Child Growth in Rural Bangladesh: A Cluster Randomised Controlled Trial. *Lancet Glob. Health* **2018**, *6*, e302−e315.
(8) Grembi, J. A.; Lin, A.; Karim, M. A.; Islam, M. O.; Miah, R.; Arnold, B. F.; McQuade, E. T. R.; Ali, S.; Rahman, M. Z.; Hussain, Z.; Shoab, A. K.; Famida, S. L.; Hossen, M. S.; Mutsuddi, P.; Rahman, M.; Unicomb, L.; Haque, R.; Taniuchi, M.; Liu, J.; Platts-Mills, J. A.; Holmes, S. P.; Stewart, C. P.; Benjamin-Chung, J.; Colford, J. M.; Houpt, E. R.; Luby, S. P. Effect of Water, Sanitation, Handwashing and Nutrition Interventions on Enteropathogens in Children 14 Months Old: A Cluster-Randomized Controlled Trial in Rural Bangladesh. *J. Infect. Dis.* **2020**, *18*, No. jiaa549.
(9) Humphrey, J. H.; Mbuya, M. N. N.; Ntozini, R.; Moulton, L. H.; Stoltzfus, R. J.; Tavengwa, N. V.; Mutasa, K.; Majo, F.; Mutasa, B.; Mangwadu, G.; Chasokela, C. M.; Chigumira, A.; Chasekwa, B.; Smith, L. E.; Tielsch, J. M.; Jones, A. D.; Manges, A. R.; Maluccio, J. A.; Prendergast, A. J. 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. *Lancet Glob. Health* **2019**, *7*, e132−e147.
(10) Pickering, A. J.; Crider, Y.; Sultana, S.; Swarthout, J.; Goddard, F. G.; Anjerul Islam, S.; Sen, S.; Ayyagari, R.; Luby, S. P. Effect of In-Line Drinking Water Chlorination at the Point of Collection on Child Diarrhoea in Urban Bangladesh: A Double-Blind, Cluster-Randomised Controlled Trial. *Lancet Glob. Health* **2019**, *7*, e1247−e1256.
(11) Reese, H.; Routray, P.; Torondel, B.; Sinharoy, S. S.; Mishra, S.; Freeman, M. C.; Chang, H. H.; Clasen, T. 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. *Int. J. Epidemiol.* **2019**, *48*, 1757−1767.
(12) Whittington, D.; Radin, M.; Jeuland, M. Evidence-Based Policy Analysis? The Strange Case of the Randomized Controlled Trials of Community-Led Total Sanitation. *Oxford Rev. Econ. Policy* **2020**, *36*, 191−221.
(13) Ercumen, A.; Mertens, A.; Arnold, B. F.; Benjamin-Chung, J.; Hubbard, A. E.; Ahmed, M. A.; Kabir, M. H.; Rahman Khalil, Md. M.; Kumar, A.; Rahman, Md. S.; Parvez, S. M.; Unicomb, L.; Rahman, M.; Ram, P. K.; Clasen, T.; Luby, S. P.; Colford, J. M. Effects of Single and Combined Water, Sanitation and Handwashing Interventions on

Fecal Contamination in the Domestic Environment: A Cluster-Randomized Controlled Trial in Rural Bangladesh. *Environ. Sci. Technol.* **2018**, *52*, 12078−12088.
(14) Pickering, A. J.; Swarthout, J.; Mureithi, M.; Mboya, J.; Arnold, B. F.; Wolfe, M.; Dentz, H. N.; Lin, A.; Arnold, C. D.; Rao, G.; Stewart, C. P.; Ram, P. K.; Clasen, T.; Colford, J. M.; Null, C. Can Individual and Integrated Water, Sanitation, and Handwashing Interventions Reduce Fecal Contamination in the Household Environment? Evidence from the WASH Benefits Cluster-Randomized Trial in Rural Kenya. *bioRxiv* **2019**, No. 202022404.
(15) Goddard, F.; Ban, R.; Barr, D. B.; Brown, J.; Cannon, J.; Colford, J. M.; Eisenberg, J. N. S.; Ercumen, A.; Petach, H.; Freeman, M. C.; Levy, K.; Luby, S. P.; Moe, C.; Pickering, A. J.; Sarnat, J. A.; Stewart, J. R.; Thomas, E. A.; Taniuchi, M.; Clasen, T. F. Measuring Environmental Exposure to Enteric Pathogens in Low-Income Settings: Review and Recommendations of an Interdisciplinary Working Group. *Environ. Sci. Technol.* **2020**, *54*, No. 11673.
(16) Holcomb, D. A.; Stewart, J. R. Microbial Indicators of Fecal Pollution: Recent Progress and Challenges in Assessing Water Quality. *Curr. Environ. Health Rep.* **2020**, *7*, 311−324.
(17) Harris, A. R.; Pickering, A. J.; Harris, M.; Doza, S.; Islam, Md. S.; Unicomb, L.; Luby, S.; Davis, J.; Boehm, A. B. Ruminants Contribute Fecal Contamination to the Urban Household Environment in Dhaka, Bangladesh. *Environ. Sci. Technol.* **2016**, *50*, 4642−4649.
(18) Boehm, A. B.; Wang, D.; Ercumen, A.; Shea, M.; Harris, A. R.; Shanks, O. C.; Kelty, C.; Ahmed, A.; Mahmud, Z. H.; Arnold, B. F.; Chase, C.; Kullmann, C.; Colford, J. M.; Luby, S. P.; Pickering, A. J. Occurrence of Host-Associated Fecal Markers on Child Hands, Household Soil, and Drinking Water in Rural Bangladeshi Households. *Environ. Sci. Technol. Lett.* **2016**, *3*, No. 393.
(19) Ercumen, A.; Pickering, A. J.; Kwong, L. H.; Arnold, B. F.; Parvez, S. M.; Alam, M.; Sen, D.; Islam, S.; Kullmann, C.; Chase, C.; Ahmed, R.; Unicomb, L.; Luby, S. P.; Colford, J. M. Animal Feces Contribute to Domestic Fecal Contamination: Evidence from E. coli Measured in Water, Hands, Food, Flies, and Soil in Bangladesh. *Environ. Sci. Technol.* **2017**, *51*, 8725−8734.
(20) Baker, K. K.; Senesac, R.; Sewell, D.; Sen Gupta, A.; Cumming, O.; Mumma, J. Fecal Fingerprints of Enteric Pathogen Contamination in Public Environments of Kisumu, Kenya, Associated with Human Sanitation Conditions and Domestic Animals. *Environ. Sci. Technol.* **2018**, *52*, 10263−10274.
(21) Wardrop, N. A.; Hill, A. G.; Dzodzomenyo, M.; Aryeetey, G.; Wright, J. A. Livestock Ownership and Microbial Contamination of Drinking-Water: Evidence from Nationally Representative Household Surveys in Ghana, Nepal and Bangladesh. *Int. J. Hyg. Environ. Health* **2018**, *221*, 33−40.
(22) Prendergast, A. J.; Gharpure, R.; Mor, S.; Viney, M.; Dube, K.; Lello, J.; Berger, C.; Siwila, J.; Joyeux, M.; Hodobo, T.; Hurt, L.; Brown, T.; Hoto, P.; Tavengwa, N.; Mutasa, K.; Craddock, S.; Chasekwa, B.; Robertson, R. C.; Evans, C.; Chidhanguro, D.; Mutasa, B.; Majo, F.; Smith, L. E.; Hirai, M.; Ntozini, R.; Humphrey, J. H.; Berendes, D. Putting the "A" into WaSH: A Call for Integrated Management of Water, Animals, Sanitation, and Hygiene. *Lancet Planet. Health* **2019**, *3*, e336−e337.
(23) Navab-Daneshmand, T.; Friedrich, M. N. D.; Gächter, M.; Montealegre, M. C.; Mlambo, L. S.; Nhiwatiwa, T.; Mosler, H.-J.; Julian, T. R. *Escherichia coli* Contamination across Multiple Environmental Compartments (Soil, Hands, Drinking Water, and Handwashing Water) in Urban Harare: Correlations and Risk Factors. *Am. J. Trop. Med. Hyg.* **2018**, *98*, 803−813.
(24) Robb, K.; Null, C.; Armah, G.; Teunis, P.; Moe, C. L.; Yakubu, H. Assessment of Fecal Exposure Pathways in Low-Income Urban Neighborhoods in Accra, Ghana: Rationale, Design, Methods, and Key Findings of the SaniPath Study. *Am. J. Trop. Med. Hyg.* **2017**, *97*, 1−26.
(25) Pickering, A. J.; Julian, T. R.; Marks, S. J.; Mattioli, M. C.; Boehm, A. B.; Schwab, K. J.; Davis, J. Fecal Contamination and Diarrheal Pathogens on Surfaces and in Soils among Tanzanian Households with and without Improved Sanitation. *Environ. Sci. Technol.* **2012**, *46*, 5736−5743.
(26) Holcomb, D. A.; Knee, J.; Sumner, T.; Adriano, Z.; de Bruijn, E.; Nalá, R.; Cumming, O.; Brown, J.; Stewart, J. R. Human Fecal Contamination of Water, Soil, and Surfaces in Households Sharing Poor-Quality Sanitation Facilities in Maputo, Mozambique. *Int. J. Hyg. Environ. Health* **2020**, *226*, No. 113496.
(27) Ishii, S.; Sadowsky, M. J. *Escherichia coli* in the Environment: Implications for Water Quality and Human Health. *Microb. Environ.* **2008**, *23*, 101−108.
(28) Byappanahalli, M. N.; Nevers, M. B.; Korajkic, A.; Staley, Z. R.; Harwood, V. J. Enterococci in the Environment. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 685−706.
(29) Price, H. D.; Adams, E. A.; Nkwanda, P. D.; Mkandawire, T. W.; Quilliam, R. S. Daily Changes in Household Water Access and Quality in Urban Slums Undermine Global Safe Water Monitoring Programmes. *Int. J. Hyg. Environ. Health* **2021**, *231*, No. 113632.
(30) Wyer, M. D.; Kay, D.; Morgan, H.; Naylor, S.; Clark, S.; Watkins, J.; Davies, C. M.; Francis, C.; Osborn, H.; Bennett, S. Within-Day Variability in Microbial Concentrations at a UK Designated Bathing Water: Implications for Regulatory Monitoring and the Application of Predictive Modelling Based on Historical Compliance Data. *Water Res. X* **2018**, No. 100006.
(31) Jenkins, M. W.; Tiwari, S.; Lorente, M.; Gichaba, C. M.; Wuertz, S. Identifying Human and Livestock Sources of Fecal Contamination in Kenya with Host-Specific Bacteroidales Assays. *Water Res.* **2009**, *43*, 4956−4966.
(32) Odagiri, M.; Schriewer, A.; Hanley, K.; Wuertz, S.; Misra, P. R.; Panigrahi, P.; Jenkins, M. W. Validation of Bacteroidales Quantitative PCR Assays Targeting Human and Animal Fecal Contamination in the Public and Domestic Domains in India. *Sci. Total Environ.* **2015**, *502*, 462−470.
(33) Feng, S.; Ahmed, W.; McLellan, S. L. Ecological and Technical Mechanisms for Cross-Reaction of Human Fecal Indicators with Animal Hosts. *Appl. Environ. Microbiol.* **2020**, *86*, No. e02319-19.
(34) Johnston, C.; Byappanahalli, M. N.; Gibson, J. M.; Ufnar, J.; Whitman, R. L.; Stewart, J. R. Probabilistic Analysis Showing That a Combination of Bacteroides and Methanobrevibacter Source Tracking Markers Is Effective for Identifying Waters Contaminated by Human Fecal Pollution. *Environ. Sci. Technol.* **2013**, *47*, 13621−13628.
(35) Kildare, B. J.; Leutenegger, C. M.; McSwain, B. S.; Bambic, D. G.; Rajal, V. B.; Wuertz, S. 16S RRNA-Based Assays for Quantitative Detection of Universal, Human-, Cow-, and Dog-Specific Fecal Bacteroidales: A Bayesian Approach. *Water Res.* **2007**, *41*, 3701−3715.
(36) Curtis, K.; Gonzalez, R. A. Integrating Bayesian Analysis and Cumulative Probability Generates High Confidence Using a Single Microbial Source Tracking Marker. *Environ. Sci. Technol.* **2019**, *53*, No. 13929.
(37) Odagiri, M.; Schriewer, A.; Daniels, M. E.; Wuertz, S.; Smith, W. A.; Clasen, T.; Schmidt, W.; Jin, Y.; Torondel, B.; Misra, P. R.; Panigrahi, P.; Jenkins, M. W. Human Fecal and Pathogen Exposure Pathways in Rural Indian Villages and the Effect of Increased Latrine Coverage. *Water Res.* **2016**, *100*, 232−244.
(38) Schriewer, A.; Odagiri, M.; Wuertz, S.; Misra, P. R.; Panigrahi, P.; Clasen, T.; Jenkins, M. W. Human and Animal Fecal Contamination of Community Water Sources, Stored Drinking Water and Hands in Rural India Measured with Validated Microbial Source Tracking Assays. *Am. J. Trop. Med. Hyg.* **2015**, *93*, 509−516.
(39) Fuhrmeister, E. R.; Ercumen, A.; Pickering, A. J.; Jeanis, K. M.; Crider, Y.; Ahmed, M.; Brown, S.; Alam, M.; Sen, D.; Islam, S.; Kabir, M. H.; Islam, M.; Rahman, M.; Kwong, L. H.; Arnold, B. F.; Luby, S. P.; Colford, J. M.; Nelson, K. L. Effect of Sanitation Improvements on Pathogens and Microbial Source Tracking Markers in the Rural Bangladeshi Household Environment. *Environ. Sci. Technol.* **2020**, *54*, No. 4316.
(40) Brown, J.; Cumming, O.; Bartram, J.; Cairncross, S.; Ensink, J.; Holcomb, D.; Knee, J.; Kolsky, P.; Liang, K.; Liang, S.; Nala, R.;

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Norman, G.; Rheingans, R.; Stewart, J.; Zavale, O.; Zuin, V.; Schmidt, W.-P. A Controlled, before-and-after Trial of an Urban Sanitation Intervention to Reduce Enteric Infections in Children: Research Protocol for the Maputo Sanitation (MapSan) Study, Mozambique. *BMJ Open* **2015**, *5*, No. e008215.
(41) Devamani, C.; Norman, G.; Schmidt, W. A Simple Microbiological Tool to Evaluate the Effect of Environmental Health Interventions on Hand Contamination. *Int. J. Environ. Res. Public Health* **2014**, *11*, 11846−11859.
(42) Knee, J.; Sumner, T.; Adriano, Z.; Berendes, D.; de Bruijn, E.; Schmidt, W.; Nalá, R.; Cumming, O.; Brown, J. Risk Factors for Childhood Enteric Infection in Urban Maputo, Mozambique: A Cross-Sectional Study. *PLoS Negl. Trop. Dis.* **2018**, *12*, No. e0006956.
(43) Schmidt, W. Randomised and Non-Randomised Studies to Estimate the Effect of Community-Level Public Health Interventions: Definitions and Methodological Considerations. *Emerg. Themes Epidemiol.* **2017**, *14*, No. 9.
(44) Capone, D.; Adriano, Z.; Berendes, D.; Cumming, O.; Dreibelbis, R.; Holcomb, D. A.; Knee, J.; Ross, I.; Brown, J. A Localized Sanitation Status Index as a Proxy for Fecal Contamination in Urban Maputo, Mozambique. *PLoS One* **2019**, *14*, No. e0224333.
(45) USEPA. *Method 1603: Escherichia coli (E. coli) in Water by Membrane Filtration Using Modified Membrane-Thermotolerant Escherichia Coli Agar*, 2009, p 13.
(46) Chern, E. C.; Siefring, S.; Paar, J.; Doolittle, M.; Haugland, R. A. Comparison of Quantitative PCR Assays for *Escherichia coli* Targeting Ribosomal RNA and Single Copy Genes. *Lett. Appl. Microbiol.* **2011**, *52*, 298−306.
(47) Green, H. C.; Haugland, R. A.; Varma, M.; Millen, H. T.; Borchardt, M. A.; Field, K. G.; Walters, W. A.; Knight, R.; Sivaganesan, M.; Kelty, C. A.; Shanks, O. C. Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Appl. Environ. Microbiol.* **2014**, *80*, 3086−3094.
(48) Johnston, C.; Ufnar, J.; Griffith, J. F.; Gooch, J.; Stewart, J. R. A Real-Time QPCR Assay for the Detection of the NifH Gene of Methanobrevibacter Smithii, a Potential Indicator of Sewage Pollution. *J. Appl. Microbiol.* **2010**, *109*, 1946−1956.
(49) Nutz, S.; Döll, K.; Karlovsky, P. Determination of the LOQ in Real-Time PCR by Receiver Operating Characteristic Curve Analysis: Application to QPCR Assays for Fusarium Verticillioides and F. Proliferatum. *Anal. Bioanal. Chem.* **2011**, *401*, 717−726.
(50) Haugland, R. A.; Siefring, S. C.; Wymer, L. J.; Brenner, K. P.; Dufour, A. P. Comparison of Enterococcus Measurements in Freshwater at Two Recreational Beaches by Quantitative Polymerase Chain Reaction and Membrane Filter Culture Analysis. *Water Res.* **2005**, *39*, 559−568.
(51) Haugland, R. A.; Siefring, S.; Lavender, J.; Varma, M. Influences of Sample Interference and Interference Controls on Quantification of Enterococci Fecal Indicator Bacteria in Surface Water Samples by the QPCR Method. *Water Res.* **2012**, *46*, 5989−6001.
(52) Haugland, R. A.; Varma, M.; Sivaganesan, M.; Kelty, C.; Peed, L.; Shanks, O. C. Evaluation of Genetic Markers from the 16S RRNA Gene V2 Region for Use in Quantitative Detection of Selected Bacteroidales Species and Human Fecal Waste by QPCR. *Syst. Appl. Microbiol.* **2010**, *33*, 348−357.
(53) Sivaganesan, M.; Haugland, R. A.; Chern, E. C.; Shanks, O. C. Improved Strategies and Optimization of Calibration Models for Real-Time PCR Absolute Quantification. *Water Res.* **2010**, *44*, 4726−4735.
(54) Wing, C.; Simon, K.; Bello-Gomez, R. A. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annu. Rev. Public Health* **2018**, *39*, 453−469.
(55) Abadie, A. Semiparametric Difference-in-Differences Estimators. *Rev. Econ. Stud.* **2005**, *72*, 1−19.
(56) Gelman, A.; Greenland, S. Are Confidence Intervals Better Termed "Uncertainty Intervals"? *Br. Med. J.* **2019**, *97*, No. l5381.
(57) Gelman, A.; Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*; Cambridge University Press, 2007.
(58) Bürkner, P.-C. Brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Softw.* **2017**, *80*, No. 1.
(59) McElreath, R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. *CRC Texts in Statistical Science*, 2nd ed.; Taylor & Francis, CRC Press: Boca Raton, 2020.
(60) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
(61) Schreiner, M.; Nsthandoca, H.; Lory, D.; Dd, D.; Yy, M. M. A Simple Poverty Scorecard for Mozambique, 2013, Microfinance.com.
(62) Rogan, W. J.; Gladen, B. Estimating Prevalence from the Results of a Screening Test. *Am. J. Epidemiol.* **1978**, *107*, 71−76.
(63) Gelman, A.; Carpenter, B. Bayesian Analysis of Tests with Unknown Specificity and Sensitivity. *J. R. Stat. Soc. C* **2020**, *69*, No. 1269.
(64) Sima, L. C.; Ng, R.; Elimelech, M. Modeling Risk Categories to Predict the Longitudinal Prevalence of Childhood Diarrhea in Indonesia. *Am. J. Trop. Med. Hyg.* **2013**, *89*, 884−891.
(65) Divelbiss, D. W.; Boccelli, D. L.; Succop, P. A.; Oerther, D. B. Environmental Health and Household Demographics Impacting Biosand Filter Maintenance and Diarrhea in Guatemala: An Application of Structural Equation Modeling. *Environ. Sci. Technol.* **2013**, *47*, 1638−1645.
(66) Ghitza, Y.; Gelman, A. Deep Interactions with MRP: Election Turnout and Voting Patterns Among Small Electoral Subgroups. *Am. J. Polit. Sci.* **2013**, *57*, 762−776.
(67) Carpenter, B.; Gelman, A.; Hoffman, M. D.; Lee, D.; Goodrich, B.; Betancourt, M.; Brubaker, M.; Guo, J.; Li, P.; Riddell, A. Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **2017**, *76*, No. 1.
(68) Stan Development Team. *RStan: The R Interface to Stan*; 2020.
(69) Gelman, A.; Pardoe, I. 2. Average Predictive Comparisons for Models with Nonlinearity, Interactions, and Variance Components. *Sociol. Methodol.* **2007**, *37*, 23−51.
(70) Dengo-Baloi, L.; Berendes, D.; Holcomb, D.; Knee, J.; Langa, J.; Magaia, A.; Nalá, R.; Brown, J. *Increased Antibiotic Resistance in the Community in Maputo, Mozambique: Preliminary Results of Microbial Analysis of a Pathogen Rich Environment*, VI Jornadas Nacionais De Saúde (National Health Journeys), Maputo, Mozambique, September 2018.
(71) Araneda-Cabrera, R. J.; Bermúdez, M.; Puertas, J. Assessment of the Performance of Drought Indices for Explaining Crop Yield Variability at the National Scale: Methodological Framework and Application to Mozambique. *Agric. Water Manage.* **2021**, *246*, No. 106692.
(72) Ercumen, A.; Pickering, A. J.; Kwong, L. H.; Mertens, A.; Arnold, B. F.; Benjamin-Chung, J.; Hubbard, A. E.; Alam, M.; Sen, D.; Islam, S.; Rahman, Md. Z.; Kullmann, C.; Chase, C.; Ahmed, R.; Parvez, S. M.; Unicomb, L.; Rahman, M.; Ram, P. K.; Clasen, T.; Luby, S. P.; Colford, J. M. Do Sanitation Improvements Reduce Fecal Contamination of Water, Hands, Food, Soil, and Flies? Evidence from a Cluster-Randomized Controlled Trial in Rural Bangladesh. *Environ. Sci. Technol.* **2018**, *52*, 12089−12097.
(73) Contreras, J. D.; Islam, M.; Mertens, A.; Pickering, A. J.; Kwong, L. H.; Arnold, B. F.; Benjamin-Chung, J.; Hubbard, A. E.; Alam, M.; Sen, D.; Islam, S.; Rahman, M.; Unicomb, L.; Luby, S. P.; Colford, J. M.; Ercumen, A. Longitudinal Effects of a Sanitation Intervention on Environmental Fecal Contamination in a Cluster-Randomized Controlled Trial in Rural Bangladesh. *Environ. Sci. Technol.* **2021**, *55*, No. 8169.
(74) Nshimyimana, J. P.; Cruz, M. C.; Thompson, R. J.; Wuertz, S. Bacteroidales Markers for Microbial Source Tracking in Southeast Asia. *Water Res.* **2017**, *118*, 239−248.
(75) Stewart, J. R.; Boehm, A. B.; Dubinsky, E. A.; Fong, T.-T.; Goodwin, K. D.; Griffith, J. F.; Noble, R. T.; Shanks, O. C.; Vijayavel, K.; Weisberg, S. B. Recommendations Following a Multi-Laboratory Comparison of Microbial Source Tracking Methods. *Water Res.* **2013**, *47*, 6829−6838.

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(76) Zhu, K.; Suttner, B.; Pickering, A.; Konstantinidis, K. T.; Brown, J. A Novel Droplet Digital PCR Human MtDNA Assay for Fecal Source Tracking. *Water Res.* **2020**, No. 116085.
(77) Pehrsson, E. C.; Tsukayama, P.; Patel, S.; Mejía-Bautista, M.; Sosa-Soto, G.; Navarrete, K. M.; Calderon, M.; Cabrera, L.; Hoyos-Arango, W.; Bertoli, M. T.; Berg, D. E.; Gilman, R. H.; Dantas, G. Interconnected Microbiomes and Resistomes in Low-Income Human Habitats. *Nature* **2016**, *533*, 212−216.
(78) Bauza, V.; Madadi, V. O.; Ocharo, R. M.; Nguyen, T. H.; Guest, J. S. 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. *Environ. Sci. Technol.* **2019**, *53*, 8271−8281.
(79) Hu, Y. O. O.; Ndegwa, N.; Alneberg, J.; Johansson, S.; Logue, J. B.; Huss, M.; Käller, M.; Lundeberg, J.; Fagerberg, J.; Andersson, A. F. Stationary and Portable Sequencing-Based Approaches for Tracing Wastewater Contamination in Urban Stormwater Systems. *Sci. Rep.* **2018**, *8*, No. 11907.
(80) Acharya, K.; Khanal, S.; Pantha, K.; Amatya, N.; Davenport, R. J.; Werner, D. A Comparative Assessment of Conventional and Molecular Methods, Including MinION Nanopore Sequencing, for Surveying Water Quality. *Sci. Rep.* **2019**, *9*, No. 15726.
(81) Capone, D.; Berendes, D.; Cumming, O.; Holcomb, D.; Knee, J.; Konstantinidis, K. T.; Levy, K.; Nalá, R.; Risk, B. B.; Stewart, J.; Brown, J. Impact of an Urban Sanitation Intervention on Enteric Pathogen Detection in Soils. *Environ. Sci. Technol.* **2021**, *55*, 9989−10000.
(82) Li, X.; Harwood, V. J.; Nayak, B.; Staley, C.; Sadowsky, M. J.; Weidhaas, J. A Novel Microbial Source Tracking Microarray for Pathogen Detection and Fecal Source Identification in Environmental Systems. *Environ. Sci. Technol.* **2015**, *49*, 7319−7329.
(83) Ercumen, A.; Arnold, B. F.; Naser, A. M.; Unicomb, L.; Colford, J. M.; Luby, S. P. Potential Sources of Bias in the Use of *Escherichia coli* to Measure Waterborne Diarrhoea Risk in Low-Income Settings. *Trop. Med. Int. Health* **2017**, *22*, 2−11.
(84) Delahoy, M. J.; Wodnik, B.; McAliley, L.; Penakalapati, G.; Swarthout, J.; Freeman, M. C.; Levy, K. Pathogens Transmitted in Animal Feces in Low- and Middle-Income Countries. *Int. J. Hyg. Environ. Health* **2018**, *221*, No. 661.
(85) Penakalapati, G.; Swarthout, J.; Delahoy, M. J.; McAliley, L.; Wodnik, B.; Levy, K.; Freeman, M. C. Exposure to Animal Feces and Human Health: A Systematic Review and Proposed Research Priorities. *Environ. Sci. Technol.* **2017**, *51*, 11537−11552.
(86) Berendes, D. M.; Yang, P. J.; Lai, A.; Hu, D.; Brown, J. Estimation of Global Recoverable Human and Animal Faecal Biomass. *Nat. Sustainability* **2018**, *1*, 679−685.
(87) Berendes, D. M.; Leon, J. S.; Kirby, A. E.; Clennon, J. A.; Raj, S. J.; Yakubu, H.; Robb, K. A.; Kartikeyan, A.; Hemavathy, P.; Gunasekaran, A.; Roy, S.; Ghale, B. C.; Kumar, J. S.; Mohan, V. R.; Kang, G.; Moe, C. L. Associations between Open Drain Flooding and Pediatric Enteric Infections in the MAL-ED Cohort in a Low-Income, Urban Neighborhood in Vellore, India. *BMC Public Health* **2019**, *19*, No. 926.
(88) Berendes, D.; Kirby, A.; Clennon, J. A.; Raj, S.; Yakubu, H.; Leon, J.; Robb, K.; Kartikeyan, A.; Hemavathy, P.; Gunasekaran, A.; Ghale, B.; Kumar, J. S.; Mohan, V. R.; Kang, G.; Moe, C. The Influence of Household- and Community-Level Sanitation and Fecal Sludge Management on Urban Fecal Contamination in Households and Drains and Enteric Infection in Children. *Am. J. Trop. Med. Hyg.* **2017**, *96*, No. 1404.
(89) Harris, M.; Alzua, M. L.; Osbert, N.; Pickering, A. Community-Level Sanitation Coverage More Strongly Associated with Child Growth and Household Drinking Water Quality than Access to a Private Toilet in Rural Mali. *Environ. Sci. Technol.* **2017**, *51*, 7219−7227.
(90) Huda, T. Md. N.; Schmidt, W.-P.; Pickering, A. J.; Unicomb, L.; Mahmud, Z. H.; Luby, S. P.; Biran, A. Effect of Neighborhood Sanitation Coverage on Fecal Contamination of the Household Environment in Rural Bangladesh. *Am. J. Trop. Med. Hyg.* **2019**, *100*, No. 717.
(91) Bauza, V.; Majorin, F.; Routray, P.; Sclar, G. D.; Caruso, B. A.; Clasen, T. Child Feces Management Practices and Fecal Contamination: A Cross-Sectional Study in Rural Odisha, India. *Sci. Total Environ.* **2019**, No. 136169.
(92) Waller, A.; Lakhanpaul, M.; Godfrey, S.; Parikh, P. Multiple and Complex Links between BabyWASH and Stunting: An Evidence Synthesis. *J. Water, Sanit. Hyg. Dev.* **2020**, *10*, No. 786.
(93) Cumming, O.; Arnold, B. F.; Ban, R.; Clasen, T.; Esteves Mills, J.; Freeman, M. C.; Gordon, B.; Guiteras, R.; Howard, G.; Hunter, P. R.; Johnston, R. B.; Pickering, A. J.; Prendergast, A. J.; Prüss-Ustün, A.; Rosenboom, J. W.; Spears, D.; Sundberg, S.; Wolf, J.; Null, C.; Luby, S. P.; Humphrey, J. H.; Colford, J. M. The Implications of Three Major New Trials for the Effect of Water, Sanitation and Hygiene on Childhood Diarrhea and Stunting: A Consensus Statement. *BMC Med.* **2019**, *17*, No. e302.
(94) Pickering, A. J.; Null, C.; Winch, P. J.; Mangwadu, G.; Arnold, B. F.; Prendergast, A. J.; Njenga, S. M.; Rahman, M.; Ntozini, R.; Benjamin-Chung, J.; Stewart, C. P.; Huda, T. M. N.; Moulton, L. H.; Colford, J. M.; Luby, S. P.; Humphrey, J. H. The WASH Benefits and SHINE Trials: Interpretation of WASH Intervention Effects on Linear Growth and Diarrhoea. *Lancet Glob. Health* **2019**, *7*, e1139−e1146.
(95) Husseini, M.; Darboe, M. K.; Moore, S. E.; Nabwera, H. M.; Prentice, A. M. Thresholds of Socio-Economic and Environmental Conditions Necessary to Escape from Childhood Malnutrition: A Natural Experiment in Rural Gambia. *BMC Med.* **2018**, *16*, No. e473.
(96) Tidwell, J.; Chipungu, J.; Ross, I.; Antwi-Agyei, P.; Alam, M.-U.; Tumwebaze, I. K.; Norman, G.; Cumming, O.; Simiyu, S. Where Shared Sanitation Is the Only Immediate Option: A Research Agenda for Shared Sanitation in Densely Populated Low-Income Urban Settings. *Am. J. Trop. Med. Hyg.* **2021**, *104*, No. 429.
(97) Sclar, G. D.; Penakalapati, G.; Caruso, B. A.; Rehfuess, E. A.; Garn, J. V.; Alexander, K. T.; Freeman, M. C.; Boisson, S.; Medlicott, K.; Clasen, T. Exploring the Relationship between Sanitation and Mental and Social Well-Being: A Systematic Review and Qualitative Synthesis. *Soc. Sci. Med.* **2018**, *217*, No. 121.
(98) Shiras, T.; Cumming, O.; Brown, J.; Muneme, B.; Nala, R.; Dreibelbis, R. Shared Latrines in Maputo, Mozambique: Exploring Emotional Well-Being and Psychosocial Stress. *BMC Int. Health Hum. Rights* **2018**, *18*, No. 30.

https://doi.org/10.1021/acs.est.1c01538
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Impacts of an urban sanitation intervention on fecal indicators and the prevalence of human fecal contamination in Mozambique: Supporting Information

David A. Holcomb, Jackie Knee, Drew Capone, Trent Sumner, Zaida Adriano, Rassul Nalá, Oliver Cumming, Joe Brown, Jill R. Stewart

This supporting information contains 47 pages, 1 figure, and 10 tables.

Table of Contents

S1. Site selection criteria

Intervention sites were selected by the nongovernmental organization (NGO) Water and Sanitation for the Urban Poor (WSUP) according to feasibility and demand criteria (Table S1). The MapSan researchers were not involved in the design of the intervention or selection of the intervention sites, but did recruit similar compounds to serve as control sites according to a reduced set of the same selection criteria applied to intervention sites.

Table S1. Baseline site selection criteria for intervention and control compounds

	required for	
criterion	**intervention compounds**	**control compounds**
located in the 11 pre-defined implementation neighborhoods	yes	no[a]
residents share sanitation in poor condition	yes	yes
at least 12 residents	yes	yes
residents willing to contribute to latrine construction costs	yes	yes
sufficient space available for construction of the new facility	yes	no
accessible for transportation of construction materials and tank-emptying activities	yes	no
legal access to piped water supply	yes	yes
groundwater level deep enough to accommodate a septic tank	yes	no
at least one child younger than 48 months old in residence	no	yes

[a] also recruited from 6 similar, adjacent neighborhoods; see Knee et al. (2021)[1]

All children under 48 months old living in study compounds at baseline enrollment were invited to participate in the MapSan health impact trial following written, informed consent from a parent or guardian. At 12-month follow-up, any children living in study compounds who were not previously enrolled but would have been under 48 months of age at the time the compound was enrolled at baseline were also invited to participate, including children born after baseline. Knee et al. (2021) provide additional details on eligibility and enrollment into the MapSan child health study component.[1]

We collected environmental samples from a subset of compounds already scheduled for baseline enrollment during May – August 2015. The specific intervention compounds at which we collected baseline environmental samples were selected opportunistically, largely prioritizing

compounds with visits scheduled earlier in the morning to ensure sufficient time for sample transport and laboratory processing. Control compounds were similarly selected for environmental sampling among those already scheduled for baseline enrollment, although we prioritized control compounds with similar numbers of residents as the intervention compounds that had been selected for environmental sampling in the preceding two weeks.

The compounds selected for environmental sampling at baseline were revisited in June – September 2016, 12 months (± 2 weeks) following the opening of the new latrine for intervention compounds and 12 months (± 2 weeks) after baseline enrollment for control compounds. Four compounds at which environmental samples were collected at baseline were unavailable at follow up due to travel or relocation of eligible children for the health impact study. The provision of intervention latrines was substantially delayed for two additional compounds following baseline enrollment, rendering them outside the 12-month (± 2 weeks) follow-up window for the duration of the June – September 2016 environmental sampling campaign. However, environmental samples were collected from 13 additional compounds during the 12-month follow-up phase that had not been sampled at baseline but had been enrolled in the larger MapSan trial at baseline. These additional compounds were selected opportunistically in the same manner as the initial baseline sample set, but intervention compounds (n = 11) were prioritized over control compounds (n = 2), which had been overrepresented at baseline.

S2. Samples collected

We visited additional households and compounds in the follow-up study phase, collecting more samples than at baseline (Table S2).[2] Fewer source water samples were collected at follow-up because the municipal water supply was available less often. One control compound (with

two enrolled households) independently upgraded the latrine after baseline and was excluded from the follow-up sample.

Table S2. Number of observations of compounds, households, and each sample type

	before		after	
observation unit	**control**	**intervention**	**control**	**intervention**
compounds	33	25	30	34
households	86	66	86	90
latrine entrance soil	33	23	30	30
household entrance soil	49	36	47	45
compound source water	23	21	19	22
household stored water	50	41	52	55
food preparation surfaces	50	40	53	51

We requested that respondents provide both household stored water and household food preparation surfaces in the manner or condition in which they would typically be used. For stored water, we requested that the respondent provide it to us as if they were giving a child water to drink. Similarly, we requested that the household respondent identify or provide a surface they would typically use to prepare food, in the condition in which they would use it. If multiple surfaces were available, the respondent decided which to present as representative of their ordinary food preparation practices. Additional descriptions of specific sampling procedures and the observed baseline characteristics of each sample type are provided in Holcomb et al. (2020).[2]

S3. qPCR assay details

All qPCR assays were performed using TaqMan Environmental Master Mix 2.0 and were subjected to an initial 10 minute incubation at 95°C, after which the cycling conditions specified by the assay developers were followed for each assay (Table S3).

Table S3. qPCR assay details

assay[reference]	cycles	parameters	oligonucleotide (nM)	sequence (5'-3')
EC23S857[3]	40	15 s: 95 °C 60 s: 60 °C	F (1000)	GGTAGAGCACTGTTTtGGCA
			R (1000)	TGTCTCCCGTGATAACtTTCTC
			P (80)	6-FAM-TCATCCCGACTTACCAACCCG-BHQ1
HF183/ BacR287[4]	40	15 s: 95 °C 60 s: 60 °C	HF183 (1000)	ATCATGAGTTCACATGTCCG
			BacR287 (1000)	CTTCCTCTCAGAACCCCTATCC
			BacP234MGB (80)	6-FAM-CTAATGGAACGCATCCC-BHQplus
Mnif [5]	50	10 s: 95 °C 30 s: 57 °C	Mnif-202F (800)	GAAAGCGGAGGTCCTGAA
			Mnif-353R (800)	ACTGAAAAACCTCCGCAAAC
			Mnif-236P (240)	6-FAM-CCGGACGTGGTGTAAC AGTAGCTA-BHQ1
Sketa22[6,7]	40	15 s: 95 °C 60 s: 60 °C	SketaF2 (1000)	GGTTTCCGCAGCTGGG
			SketaR22 (1000)	CCGAGCCGTCCTGGTC
			SketaP2 (80)	6-FAM-AGTCGCAGGCGGCCACCGT-BHQ1

S4. Calibration curves

Calibration curves for quantifying molecular fecal indicator gene copies were fit to ten-fold dilution series of positive controls (PC) that had been spiked with reference material for all three assays and extracted alongside each batch of samples.[2] Three gBlock linear DNA fragments (Integrated DNA Technologies, Skokie, Il, USA) containing composite reference sequences for the three fecal source tracking assays considered in this paper as well as additional assays considered in the associated validation study were used as standard reference material for the positive controls (Table S4).[2,8,9] Extracting the reference material accounted for loss of target DNA during extraction but induced additional variability between dilution series constructed from each PC. We therefore allowed both the slopes and intercepts of the calibration curves to vary by qPCR instrument run and extraction batch to account for the additional variation. The resulting curves were relatively linear (R^2 > 95%) when averaged across all instrument runs and extraction batches but were somewhat inefficient, particularly for HF183 (Table S5).

Table S4. Synthetic DNA reference material spiked into positive controls

assays covered	sequence [5'-3']	GenBank (base positions)	length [bases]
BacHum-UCD[10] BacUni-UCD[10] HF183/BacR287[11]	CCAGGATGGGATCATGAGTTCACATGTCCGCATGAT TAAAGGTATTTTCCGGTAGACGATGGGGATGCGTTC CATTAGATAGTAGGCGGGGTAACGGCCCACCTAGTC AACGATGGATAGGGGTTCTGAGAGGAAGGTCCCCCA CATTGGAACTGAGACACGGTCCAAACTCCTACGGGA GGCAGCAGTGAGGAATATTGGTCAATGGGCGATGGC CTGAACCAGCCAAGTAGCGTGAAGGATGACTGCCCT ATGGGTTGTAAACTTCTTTTATAAAGGAATAAAGTC GGGTATGCATACCCGTTTGCATGTACTTTATGAATAA GGATCGGCTAACTCCGTGCCAGCAGCCGCGGTAATA CGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTT AAAGGGAGCGTAGATGGATGTTTAAGTCAGTTGTGA AAGTTTGCGGCTCAACCGTAAAATTGCAGTTGATAC TGGATGTCTTGAGTGCAGTTGAGGCAGGCGGAATTC GTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGA ACTCCGATTGCGAAGGCAGC	AB242142 (170-730)	560
GFD[12] LA35[13]	TGGGTCTAATACCGGATACGACCATCTGCCGCATGG CGGGTGGTGGAAAGTTTTTCGATTGGGGATGGGCTC GCGGCCTATCAGTTTGTTGGTGGGGTAATGGCCTAC CAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGA CCGGCCACACTGGGACTGAGACACGGCCCAGACTCC TACGGGAGGCAGCAGTGGGGAATATTGCACAATGG GGGAAACCCTGATGCAGCGACGCAGCGTGCGGGAT GACGGCCTTCGGGTTGTAAACCGCTTTCAGCAGGGA AGAAGCCTTCGGGTGACGGTACCTGCAGAAGAAGTA CCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT AGGGTACGAGCGTTGTCCGGAATTATTGGGCGTAAA GAGCTCGTAGGTGGTTGGTCACGTCTGCTGTGGAAA CGCAACGCTTAACGTTGCGCGGGCAGTGGGTACGGG CTGACTAGAGTGCAGTAGGGGAGTCTGGAATTCCTG GTGTAGCGGTGAAATGCGCAGATATCAGGAGGAAC ACCGGTGGCGAAGGCGGGACTCTGGGCTGTGACTGA CACTGGGGAGCGAAAGCATTGCTAACAGTTcGGCTG AGCACTCTAGGGAGACTGCCTTCGCAAGGAGGAGGA AGGTGAGGACGACGTCAAGTCATCATGGCCCTTACG CCTAGGGCTACACACGTGCTACAATGGGATGTACAA AGAGACGCAATACCGCGA	FJ462358 (156-746) JN084061 (29-171)	732
EC23S857[3] HAdV[14] Mnif[5]	TAACTATGGTCATCGTTCGTCAGCAGTAACAGTAATT GCTACACCTGCTGAAACCACTGTCCCTTTTTCTTGGG CAACTCTTGTTTATGTGTTGAAAGCGGAGGTCCTGA ACCGGGTGTTGGCTGTGCCGGACGTGGTGTAACAGT AGCTATGAAAAGACTTGAAAACTTAGGTGTTTTTGA TAAGGATTTGGATGTAGTCATTTATGGTGTACTTGGA GATGTTGTTTGCGGAGGTTTTTCAGTGCCTTTACGTT CTCGGGCCAGGACGCCTCGGAGTACCTGAGCCCCGG GCTGGTGCAGTTTGCCCGCGCCACCGAGACGTACTT CAGCCTGAATAACAAGTTTAGAAACCCCACGGTGGC GCCTACGCATCTCCGGGGGTAGAGCACTGTTTCGGC AAGGGGGTCATCCCGACTTACCAACCCGATGCAAAC TGCGAATACCGGAGAATGTTATCACGGGAGACACAC GGCGGGT	AE015928 (4515891- 4515973) AB019138 (192-363) AC_000008 (18885-19000) DQ682619 (847-954)	475

Table S5. Mean (95% CI) estimates of calibration curve parameters

assay	intercept	slope	efficiency (%)	R^2
EC23S	47.9 (47.3, 48.6)	-3.50 (-3.64, -3.37)	93.1 (88.1, 98.1)	0.98 (0.97, 0.98)
HF183	47.5 (46.4, 48.6)	-3.85 (-4.07, -3.67)	81.8 (76.1, 87.4)	0.98 (0.97, 0.98)
Mnif	48.8 (47.7, 50.3)	-3.47 (-3.79, -3.23)	94.5 (83.6, 104.0)	0.95 (0.93, 0.95)

S5. Detection limits

The limit of detection (LOD) for each assay was obtained in terms of the quantification cycle (Cq), the number of amplification cycles above which the target would be considered absent from the reaction, using receiver operating characteristic (ROC) analysis.[2,15] We performed ROC analysis on the local validation study data presented in Holcomb et al. (2020), considering Cq cutoff values from 10 to the maximum number of cycles indicated by the assay developers, in full-cycle increments. We calculated diagnostic sensitivity and specificity for each cutoff value, considering any reactions with a Cq value below the cutoff as positive. The highest Cq value that maximized the Youden index, $J = sensitivity + specificity - 1$, for each assay was considered the assay LOD (Table S6).[15,16]

Process limits of detection (PLOD) were estimated for each sample from the assay LOD Cq values, the extraction batch- and instrument run-specific calibration curve estimates, and the amount of sample processed, i.e. volume of water filtered or the dry mass or surface area represented by the eluate filtered for soil samples and surface swabs, respectively. For *E. coli* enumerated by culture (cEC), we assumed an assay LOD of 1 cfu per plate and calculated the PLOD for each sample on the basis of the amount of sample represented by the least-diluted plate read for that sample.[2] The PLOD averages indicate that relatively high gene copy concentrations were required for reliable detection by any of the three qPCR-based assays, on the order of >10,000 gc/dry gram of soil, nearly 1000 gc/100 mL water, and >3000 gc/100 cm^2 of food preparation surface (Table S6).

Table S6. qPCR assay limits of detection and mean sample-specific process limits of detection by sample matrix

		process limit of detection					
		soil		water		surface	
	assay	log_{10} gc[c]/dry g		log_{10} gc/100 mL		log_{10} gc/100 cm^2	
assay	LOD[a] Cq[b]	n	mean (SD[d])	n	mean (SD)	n	mean (SD)
EC23S	39	298	4.51 (0.11)	292	3.27 (0.12)	199	3.96 (0.30)
HF183	39	299	4.24 (0.33)	292	2.85 (0.29)	199	3.45 (0.36)
Mnif	41	298	4.18 (0.17)	291	3.03 (0.11)	196	3.53 (0.35)

[a] limit of detection
[b] quantification cycle
[c] gene copies
[d] standard deviation

S6. Laboratory quality control

Sterile PBS was filtered as a laboratory blank for approximately every 10 samples filtered, all of which were culture-negative for cEC (n = 151). At least three no template control (NTC) reactions were included on every qPCR run using 5 µL nuclease-free water in place of sample template. Each qPCR run typically included samples from three extraction batches, and a negative extraction control (NEC) was included from each extraction batch represented on the plate. HF183 was absent in all NTC (n = 46) and NEC (n = 46) reactions. Mnif was likewise not detected in any NTC (n = 42) or NEC (n = 46) reaction. However, EC23S was detected in 2% of NTC (1/45) and 11% of NEC (5/46) reactions. EC23S concentrations were low in the contaminated negative controls, with a mean Cq of 38.1 and a minimum Cq of 37.1—only slightly above the detection limit of 39 cycles. Such low levels of contamination have been reported previously and are thought to be due to residual *E. coli* DNA in the Environmental Master Mix from the production process.[2,17] One latrine soil sample was considered inhibited based on a greater than 3 Cq deviation from the mean Sketa22 Cq of all NECs and positive controls on the same plate. This sample was diluted 1:5 in all further analyses. For each sample type, 10% of samples were randomly selected to be analyzed by each qPCR assay in technical duplicate reactions. The detection status of EC23S agreed in 96% (74/77) of replicate pairs.

HF183 results matched for 86% (66/77) of paired reactions and Mnif agreed in 95% (71/75). Samples analyzed in duplicate were considered positive for a given target if at least one of the replicate reactions was above the limit of detection.

S7. Conditional probability

The conditional probability of human contamination, C, given the detection of two human-associated markers, $M_1$and M_2 is given by

$$P(C^+|M_1^+ \cap M_2^+) = \frac{P(M_1^+|C^+) \times P(M_2^+|C^+) \times P(C^+)}{P(M_1^+|C^+) \times P(M_2^+|C^+) \times P(C) + P(M_1^+|C^-) \times P(M_2^+|C^-) \times P(C^-)} = \frac{Se_1 \times Se_2 \times P(C^+)}{Se_1 \times Se_2 \times P(C^+) + (1 - Sp_1) \times (1 - Sp_2) \times \left(1 - P(C^+)\right)} \quad (1)$$

where $Se_1 = P(M_1^+|C^+)$, $Se_2 = P(M_2^+|C^+)$, $Sp_1 = P(M_1^-|C^-)$, $Sp_2 = P(M_2^-|C^-)$, and A^+ indicates the presence, and A^- the absence, of variable A.[18] We calculated the conditional probability of contamination for HF183 and Mnif separately and for each combination of the two indicators (M_1^+, M_2^+; M_1^+, M_2^-; M_1^-, M_2^+; M_1^-, M_2^-) by sample type.

S8. Difference-in-differences estimates

The effect of the intervention was estimated using difference-in-differences approaches. Crude DID estimates of *E. coli* $\log_{10}$-concentration and human FST marker prevalence were calculated using 2000 bootstrap samples (Table S7). The DID estimate was calculated as

$$DID = \left(\mathrm{E}\left[Y_{T=1,P=1}\right] - \mathrm{E}\left[Y_{T=1,P=0}\right]\right) - \left(\mathrm{E}\left[Y_{T=0,P=1}\right] - \mathrm{E}\left[Y_{T=0,P=0}\right]\right) \quad (2)$$

where $Y_{T,P}$ are the observed indicator values for treatment group T in study phase P. The value of T is 0 for control compound observations and 1 for intervention compounds; likewise, P takes the value 0 for pre-treatment (baseline) observations and 1 post-treatment (follow-up). Observations below the PLOD or too numerous to count were imputed by sample type from a truncated normal distribution with mean and standard deviation obtained through maximum

likelihood estimation, assuming the $\log_{10}$ concentration was normally distributed and subject to left- and right-censoring.[2,19] Model-based DID estimates were obtained using the product-term representation of the DID estimator in regression models, which permits the inclusion of additional covariates. We produced both crude model-based estimates, which only included terms for the DID estimator, as well as adjusted estimates that included additional terms for meteorological, compound, household, sample characteristics (Table S8).

Table S7. Bootstrap difference-in-differences estimates

		before				after				
		control		intervention		control		intervention		DID[a]
indicator	sample type	N	estimate	N	estimate	N	estimate	N	estimate	estimate
					***E. coli* $\log_{10}$ concentration, mean (95% CI)**					
cEC	latrine soil	33	4.0 (3.6, 4.3)	23	4.0 (3.5, 4.4)	30	3.3 (2.9, 3.7)	30	3.0 (2.6, 3.5)	-0.3 (-1.1, 0.5)
	household soil	49	4.1 (3.8, 4.3)	36	4.2 (3.9, 4.5)	47	3.7 (3.5, 4.0)	45	3.4 (3.1, 3.8)	-0.4 (-1.0, 0.1)
	stored water	50	1.8 (1.4, 2.2)	41	1.6 (1.1, 2.0)	52	1.1 (0.7, 1.4)	55	1.0 (0.7, 1.3)	0.2 (-0.6, 0.9)
	food surface	50	3.3 (2.8, 3.8)	40	2.9 (2.3, 3.6)	53	2.0 (1.5, 2.5)	51	1.9 (1.4, 2.4)	0.3 (-0.8, 1.3)
EC23S	latrine soil	33	6.2 (5.8, 6.6)	23	6.9 (6.4, 7.4)	30	6.7 (6.4, 7.1)	30	6.5 (6.0, 7.0)	-0.9 (-1.8, 0.0)
	household soil	49	6.8 (6.5, 7.0)	36	6.7 (6.4, 7.0)	47	6.7 (6.4, 7.0)	45	6.6 (6.4, 6.9)	0.0 (-0.5, 0.5)
	stored water	50	4.1 (3.9, 4.3)	41	4.4 (4.2, 4.7)	52	4.2 (3.9, 4.4)	55	4.0 (3.8, 4.2)	-0.4 (-0.9, 0.0)
	food surface	50	4.4 (4.2, 4.7)	40	5.0 (4.8, 5.3)	53	4.7 (4.4, 5.0)	51	4.5 (4.2, 4.8)	-0.8 (-1.4, -0.3)
					human marker prevalence, % (95% CI)					
HF183	latrine soil	33	33 (17, 50)	23	43 (23, 64)	30	57 (39, 75)	30	43 (26, 61)	-23 (-60, 14)
	household soil	49	17 (7, 28)	36	36 (20, 52)	47	49 (35, 64)	45	38 (24, 52)	-30 (-57, -1)
	stored water	50	12 (4, 22)	41	22 (10, 35)	52	10 (2, 20)	55	19 (9, 30)	-1 (-21, 19)
	food surface	50	2 (0, 7)	40	0 (0, 0)	53	9 (2, 18)	51	2 (0, 7)	-5 (-16, 4)
Mnif	latrine soil	33	51 (35, 69)	23	65 (45, 84)	30	50 (32, 68)	30	36 (19, 54)	-27 (-63, 8)
	household soil	49	43 (30, 57)	36	23 (9, 37)	47	25 (13, 39)	45	24 (12, 38)	20 (-7, 46)
	stored water	50	0 (0, 0)	41	2 (0, 8)	52	0 (0, 0)	55	0 (0, 0)	-2 (-8, 0)
	food surface	50	0 (0, 0)	40	3 (0, 8)	53	2 (0, 7)	51	0 (0, 0)	-4 (-12, 0)

[a] difference-in-differences: (intervention, after – intervention, before) – (control, after – control, before)

Table S8. Model-based difference-in-differences estimates

	latrine entrance soil				household entrance soil				household stored water				food preparation surfaces			
	crude[a]		adjusted[b]		crude		adjusted		crude		adjusted		crude		adjusted	
target	N	DID[c]	N	DID	N	DID	N	DID	N	DID	N	DID	N	DID	N	DID
									***E. coli* $\log_{10}$ concentration change (95% CI)**							
cEC	111	-0.37 (-1.17, 0.44)	95	-0.42 (-1.28, 0.40)	175	-0.34 (-0.92, 0.22)	146	0.05 (-0.62, 0.72)	194	0.15 (-0.61, 0.92)	170	-0.42 (-1.28, 0.44)	192	0.18 (-0.9, 1.26)	169	-0.11 (-1.4, 1.17)
EC23S	116	-0.84 (-1.64, -0.02)	98	-1.22 (-2.11, -0.30)	176	0.06 (-0.46, 0.57)	147	0.36 (-0.26, 1.01)	193	-0.46 (-0.89, -0.04)	170	-0.41 (-0.92, 0.12)	193	-0.82 (-1.46, -0.19)	171	-0.56 (-1.32, 0.19)
									human target prevalence odds ratio (95% CI)							
HF183	116	0.94 (0.40, 1.87)	98	0.92 (0.38, 1.87)	176	0.83 (0.39, 1.54)	147	0.90 (0.38, 1.80)	193	1.12 (0.48, 2.27)	170	1.05 (0.44, 2.15)				
Mnif	116	0.73 (0.31, 1.44)	98	0.71 (0.29, 1.47)	175	1.24 (0.53, 2.46)	146	1.00 (0.39, 2.06)								

[a] not adjusted for covariates; include terms for study phase, treatment arm, and their product, and compound-varying intercepts

[b] adjusted for animal presence, population density, household wealth, temperature, precipitation, and sample-specific variables: sun exposure and surface wetness for soils, presence of lid and width of container mouth for stored water, and food surface material.

[c] difference-in-differences estimated as the regression coefficient on the product of study phase and treatment arm indicators

S9. Validation studies

Local differences in diet, geography, and population history affecting the gut microbiome composition of a given population are expected to play the key role in determining local fecal source tracking performance.[20,21] However, multi-laboratory comparisons have also demonstrated that assay performance can vary meaningfully between labs analyzing the same set of challenge samples.[22] Furthermore, assay design and other intrinsic characteristics may also impact potential performance—that is, some assays may have more robust designs that increase the likelihood of performing well in a variety of settings and populations. Source tracking validation studies typically report crude sensitivity and specificity values without quantifying the uncertainty in these estimates, which could be substantial considering the limited number of samples analyzed in many studies, particularly in resource-limited settings (Table S9).

We incorporate a meta-analysis of FST validation studies for our selected human markers, which provides an opportunity to partially pool information across a variety of locations to potentially refine the sensitivity and specificity estimates for our specific study setting.[23] The hierarchical structure of the meta-analytic model means that the estimates for any given location are driven by the data from that location. Sensitivity and specificity estimates at a location with a large amount of data would be almost entirely determined by the local data and largely unaffected by the other studies, while estimates from locations with sparse data, which would be uncertain on their own, incorporate more information pooled from other studies.[24] The extent of information pooling is also determined by the consistency of marker performance across studies. Highly variable marker performance between study locations suggests that performance is mostly driven by differences in local characteristics, limiting the amount of information that can be gained by considering performance in other locations. Accordingly, meta-analytic

performance estimates in a particular location will be determined largely by the local samples and minimally influenced by data from other locations. On the other hand, similar marker performance across locations suggests that intrinsic assay characteristics play a larger role in its performance, enabling information from other locations to refine the estimates at specific locations with limited local data.[25] We adjust for diagnostic accuracy using the meta-analytic sensitivity and specificity estimates for our study location, such that the estimates are driven by our local data but are influenced by the broader trends in assay performance across all the studies considered. The degree of influence by outside data depends on the degree of uncertainty in the local data (a function of sample size) and the similarity of performance estimates across all the studies considered, which is reflected in the between-study standard deviations, σ^{Se} and σ^{Sp}.

Validation studies included in the performance meta-analysis were identified from Google Scholar records of articles citing the original assay publications and from the references of each published validation study identified.[5,7,11] We only included HF183 studies that assessed the HF183/BFDRev assay or its modification used in this study, HF183/BacR287.[11] We identified HF183 validation studies conducted on five continents, including two in Africa, seven in Asia, two in Australia, four in North America, and one in South America (Table S9). All identified Mnif studies used samples from the USA except for our study in Mozambique. The rural/urban status of the settings from which samples were collected could not be determined for all studies and were more often specified for studies conducted in low- and middle-income countries. We used estimates that considered samples detected below the limit of quantification (DNQ) as negatives when separate estimates were reported by DNQ definition.[26]

Table S9. Published validation studies of human fecal indicators HF183 and Mnif

			human samples		non-human samples				
study	**location**	**setting**	**N**	**positive**	**N**	**positive**	**sensitivity**	**specificity**	**ref**
HF183									
1	Mozambique	urban	14	9	27	9	64%	67%	[2]
2	Bangladesh	urban	5	3	20	12	60%	40%	[21]
3	Bangladesh	rural	5	4	20	10	80%	50%	[27]
4	India	rural	35	10[a]	60	12[a]	29%	80%	[28]
5	Kenya	rural	17	11	25	0	65%	100%	[29]
6	Thailand		28	25	100	19	89%	81%	[30]
7	Costa Rica		8	8	47	3	100%	94%	[31]
8	Singapore	urban	56	42	85	9	75%	89%	[32]
9	Nepal		10	10	44	30	100%	32%	[33]
10	USA	rural	4	4	109	0	100%	100%	[34]
11	Japan		20	20	15	0	100%	100%	[35]
12	Australia		32	24	359	12	72%	96%	[36]
13	USA, Australia				184	6		97%	[37]
14	USA		60	57[a]	130	10 [a]	95%	92%	[26]
15	Peru	peri-urban	30	23	71	24	77%	66%	[38]
Mnif									
1	Mozambique	urban	14	10	27	8	71%	70%	[2]
2	USA		60	36[a]	130	31[a]	60%	76%	[26]
3	Indiana		59	40	120	9	68%	93%	[18]
4	Mississippi		62	51	243	101	82%	58%	[18]

[a] samples below the limit of quantification considered negative

S10. Diagnostic accuracy

We considered HF183 validation data from 14 studies of diagnostic sensitivity and 15 studies of specificity (Table S7). For Mnif, we incorporated data from four validation studies. The number of samples ranged from 5–62 for human feces and 15–359 for non-human. Reported crude sensitivity ranged from 29-100% for HF183 and 60-82% for Mnif. Crude specificities were reported from 32–100% for HF183 and 58-93% for Mnif.

By incorporating indicator validation data into models of human fecal contamination, we obtained estimates of indicator sensitivity and specificity that were partially informed by observations in environmental samples with unknown fecal contamination status. For this reason, accuracy estimates from the same model differed by sample type (main text, Table 2). We first

considered only validation data from our study area (Model 2), which produced slightly lower point estimates for HF183 sensitivity (60%) than obtained from crude calculations (64%). The sensitivity estimates were similar for all three sample types, with the greatest uncertainty in stored water (95% CI: 38-81%). The HF183 specificity estimate was similar to the crude value (67%) for latrine soil [66% (53-80%)] but the estimates were higher for household soil [72% (61-83%)] and stored water [85% (77-91%)], in which HF183 was detected less frequently. Sensitivity and specificity patterns for Mnif followed similar patterns.

Meta-analytic sensitivity and specificity estimates were slightly higher than the single-study estimates when using a single indicator (Model 3) and when combining both indicators (Model 4). However, the 95% CIs remained similar for all three models, suggesting the inclusion of additional information minimally impacted estimates of diagnostic accuracy. Model 5, which simultaneously considered both indicators in all sample types, provided a single set of sensitivity and specificity estimates for each indictor. The estimates were comparable to values observed for individual sample types but featured narrower 95% CIs, indicating reduced uncertainty in indicator performance with the inclusion of data from multiple sample types. While the estimated sensitivity for Mnif [71% (59-83%)] remained similar to the crude value (71%), the estimated specificity of both indicators and HF183 sensitivity were somewhat higher than expected from crude calculations.

S11. Accuracy-adjusted effects

We constructed a series of models to estimate sanitation intervention effects on the prevalence of human fecal contamination. In the first model, human-associated fecal indicators were used as direct proxies for human fecal contamination, equivalent to assuming 100% sensitivity and specificity. The estimated intervention effect, given by the DID prevalence odds

ratio, was less than one for both indicators in latrine soil and household soil and above one for HF183 in stored water (Figure S1). However, the estimates were imprecise and encompassed both large increases and large reductions within the 95% CIs, providing little evidence for an effect of the intervention on either indicator in any sample type. Incorporating local validation data to account for indicator accuracy (Model 2) further widened the 95% CIs, indicating additional uncertainty about the intervention effect on the prevalence of human fecal contamination. Effect estimates incorporating multiple validation datasets (Model 3) were largely similar to those using local validation data alone. When using HF183 and Mnif to jointly estimate human fecal contamination (Model 4), the DID estimates were similar to estimates for individual indicators. Similarly, combining observations from all three sample types produced an estimated intervention effect close to the null with the 95% CI encompassing both reductions and increases in the odds of human fecal contamination at the compound level [POR: 0.93 (95% CI: 0.42-2.1)]. Across all model formulations, there was little evidence for an effect of the sanitation intervention on the prevalence of human fecal contamination.

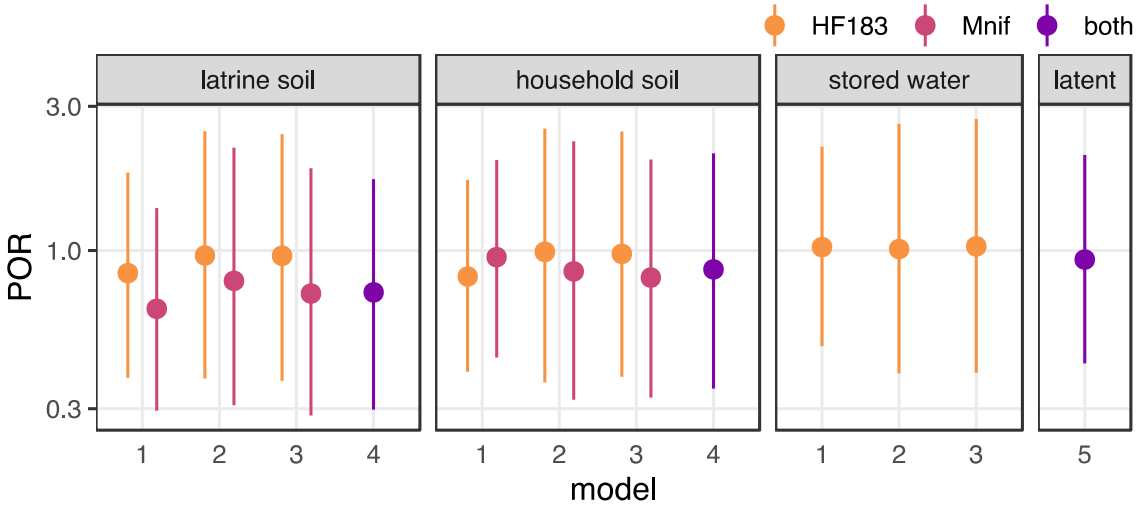


Figure S1. Difference-in-difference prevalence odds ratio (POR) estimates of the sanitation intervention effect on human fecal contamination under five different models. Model 1 made no correction for indicator accuracy, Model 2 used local validation data to account for indicator sensitivity and specificity, and Models 3, 4 and 5 used meta-analytic estimates of local indicator accuracy. Model 3 was fit separately by indicator and sample type, Model 4 was fit to both indicators by sample type, and Model 5 used both indicators in all sample types to estimate the latent compound prevalence of human fecal contamination. All models were adjusted for population density, presence of animals, wealth score, temperature, antecedent precipitation, and sun exposure and surface wetness for soil samples and storage container mouth width and cover status for water samples.

S12. Human fecal contamination prevalence estimates

Posterior predictions from each of the five models were used to estimate stratum-specific prevalence. We fit unadjusted models that included only intercepts and the DID terms (Table S10) as well as adjusted models with meteorological, compound, household, and sample characteristics included as covariates (Table 2, main text).

Table S10. Estimated sensitivity, specificity, and prevalence of human fecal contamination from unadjusted models

					prevalence estimate (95% CI)				
					control		intervention		
	target	sensitivity (95% CI)	specificity (95% CI)	N	before	after	before	after	prevalence DID (95% CI)
latrine soil									
bootstrap	HF183	1	1	116	0.33 (0.17, 0.50)	0.57 (0.39, 0.75)	0.43 (0.23, 0.64)	0.43 (0.26, 0.61)	-0.23 (-0.60, 0.14)
	Mnif	1	1	116	0.51 (0.35, 0.69)	0.50 (0.32, 0.68)	0.65 (0.45, 0.84)	0.36 (0.19, 0.54)	-0.27 (-0.63, 0.08)
model 1	HF183	1	1	116	0.41 (0.28, 0.54)	0.49 (0.35, 0.64)	0.40 (0.26, 0.56)	0.45 (0.30, 0.61)	-0.04 (-0.22, 0.14)
	Mnif	1	1	116	0.54 (0.41, 0.67)	0.48 (0.35, 0.62)	0.57 (0.41, 0.71)	0.42 (0.27, 0.57)	-0.09 (-0.27, 0.09)
model 2	HF183	0.59 (0.41, 0.80)	0.65 (0.52, 0.79)	116	0.41 (0.05, 0.90)	0.44 (0.05, 0.92)	0.41 (0.04, 0.91)	0.43 (0.04, 0.93)	-0.01 (-0.20, 0.19)
	Mnif	0.68 (0.50, 0.86)	0.67 (0.52, 0.82)	116	0.53 (0.11, 0.92)	0.48 (0.09, 0.91)	0.53 (0.10, 0.92)	0.44 (0.07, 0.93)	-0.05 (-0.25, 0.14)
model 3	HF183	0.65 (0.45, 0.85)	0.68 (0.55, 0.83)	116	0.39 (0.06, 0.87)	0.43 (0.06, 0.89)	0.38 (0.05, 0.88)	0.41 (0.04, 0.91)	-0.01 (-0.19, 0.18)
	Mnif	0.70 (0.55, 0.83)	0.70 (0.55, 0.84)	116	0.54 (0.17, 0.87)	0.49 (0.13, 0.88)	0.55 (0.14, 0.88)	0.44 (0.09, 0.88)	-0.05 (-0.25, 0.14)
model 4	HF183	0.64 (0.45, 0.84)	0.69 (0.55, 0.83)	116	0.47 (0.16, 0.79)	0.46 (0.13, 0.79)	0.47 (0.13, 0.81)	0.40 (0.10, 0.77)	-0.06 (-0.25, 0.13)
	Mnif	0.71 (0.56, 0.84)	0.69 (0.54, 0.83)						
model 5	HF183	0.72 (0.55, 0.87)	0.83 (0.76, 0.90)	107	0.48 (0.25, 0.73)	0.49 (0.27, 0.72)	0.43 (0.19, 0.70)	0.47 (0.25, 0.71)	0.04 (-0.14, 0.23)
	Mnif	0.71 (0.58, 0.84)	0.79 (0.70, 0.88)						
household soil									
bootstrap	HF183	1	1	176	0.17 (0.07, 0.28)	0.49 (0.35, 0.64)	0.36 (0.20, 0.52)	0.38 (0.24, 0.52)	-0.30 (-0.57, -0.01)
	Mnif	1	1	175	0.43 (0.30, 0.57)	0.25 (0.13, 0.39)	0.23 (0.09, 0.37)	0.24 (0.12, 0.38)	0.20 (-0.07, 0.46)
model 1	HF183	1	1	176	0.27 (0.18, 0.37)	0.41 (0.30, 0.53)	0.31 (0.20, 0.44)	0.40 (0.28, 0.53)	-0.05 (-0.21, 0.11)
	Mnif	1	1	175	0.37 (0.26, 0.48)	0.29 (0.19, 0.40)	0.29 (0.18, 0.41)	0.24 (0.14, 0.35)	0.03 (-0.11, 0.17)
model 2	HF183	0.58 (0.36, 0.80)	0.71 (0.61, 0.83)	176	0.25 (0.03, 0.78)	0.30 (0.02, 0.81)	0.26 (0.02, 0.78)	0.31 (0.02, 0.84)	0.00 (-0.18, 0.19)
	Mnif	0.65 (0.41, 0.86)	0.75 (0.66, 0.85)	175	0.22 (0.03, 0.55)	0.18 (0.02, 0.48)	0.18 (0.02, 0.49)	0.13 (0.01, 0.43)	0.00 (-0.13, 0.13)
model 3	HF183	0.65 (0.40, 0.85)	0.74 (0.63, 0.85)	176	0.23 (0.03, 0.64)	0.30 (0.03, 0.71)	0.23 (0.03, 0.66)	0.30 (0.02, 0.76)	0.00 (-0.18, 0.18)
	Mnif	0.69 (0.52, 0.84)	0.77 (0.68, 0.86)	175	0.22 (0.03, 0.50)	0.17 (0.02, 0.41)	0.17 (0.02, 0.41)	0.12 (0.01, 0.34)	0.00 (-0.12, 0.13)
model 4	HF183	0.67 (0.44, 0.86)	0.71 (0.62, 0.80)	175	0.15 (0.02, 0.35)	0.16 (0.02, 0.39)	0.13 (0.02, 0.31)	0.13 (0.01, 0.34)	-0.01 (-0.12, 0.11)
	Mnif	0.69 (0.52, 0.83)	0.76 (0.67, 0.85)						
model 5	HF183	0.72 (0.55, 0.87)	0.83 (0.76, 0.90)	156	0.24 (0.09, 0.43)	0.24 (0.10, 0.42)	0.20 (0.06, 0.39)	0.23 (0.08, 0.42)	0.03 (-0.11, 0.17)
	Mnif	0.71 (0.58, 0.84)	0.79 (0.70, 0.88)						
stored water									
bootstrap	HF183	1	1	193	0.12 (0.04, 0.22)	0.10 (0.02, 0.20)	0.22 (0.10, 0.35)	0.19 (0.09, 0.30)	-0.01 (-0.21, 0.19)
model 1	HF183	1	1	193	0.14 (0.08, 0.23)	0.12 (0.06, 0.20)	0.19 (0.11, 0.29)	0.18 (0.10, 0.28)	0.01 (-0.09, 0.11)
model 2	HF183	0.58 (0.34, 0.80)	0.84 (0.78, 0.89)	193	0.08 (0.01, 0.23)	0.07 (0.01, 0.22)	0.08 (0.01, 0.27)	0.07 (0.01, 0.26)	0.00 (-0.08, 0.09)
model 3	HF183	0.65 (0.40, 0.85)	0.85 (0.79, 0.90)	193	0.07 (0.01, 0.20)	0.06 (0.01, 0.19)	0.08 (0.01, 0.25)	0.07 (0.01, 0.24)	0.00 (-0.08, 0.08)
model 5	HF183	0.72 (0.55, 0.87)	0.83 (0.76, 0.90)	176	0.08 (0.01, 0.20)	0.09 (0.01, 0.21)	0.07 (0.01, 0.18)	0.08 (0.01, 0.20)	0.01 (-0.05, 0.08)
latent compound									
m5	HF183	0.72 (0.55, 0.87)	0.83 (0.76, 0.90)	113	0.26 (0.08, 0.55)	0.26 (0.08, 0.55)	0.22 (0.06, 0.50)	0.25 (0.07, 0.53)	0.03 (-0.11, 0.18)
	Mnif	0.71 (0.58, 0.84)	0.79 (0.70, 0.88)						

S13. Prior distributions

For all models, we sought to select regularizing ("weakly informative") priors where feasible. Regularizing priors impose soft constraints on parameter values, discouraging the model fitting algorithm from exploring extreme values that can reasonably be assumed implausible *a priori*.[25] In addition to the practical benefit of often aiding model convergence, particularly for complex, high-dimensional models, regularization can help improve the precision of estimates for parameters for which the underlying data are somewhat noisy. This also has the effect of shrinking estimates towards a common value—often the null, in the case of regression coefficients, which mildly increases the strength of the evidence necessary to demonstrate a probable effect but provides the advantage of reducing false positives that can arise from typical sampling variation.[25] While setting regularizing priors is subjective in the strictest sense, one generally possesses sufficient information to determine a broadly plausible range of parameter values that arguably are more readily accepted than the assumption implicit in flat ("non-informative") priors that the range of potential parameter values is essentially infinite.

For the basic DID models used to assess the intervention effect on individual fecal indicators in separate sample types (Table S8), we included a compound-varying intercept that used the **brms** package default positive-constrained student-t prior with 3 degrees of freedom (df) and scale determined from the link-transformed data.[39] Censored linear regression was used to estimate the intervention impact on the $\log_{10}$ concentrations of the two *E. coli* assays, cEC and EC23S, using regularizing normal priors with standard deviation (SD) = 5 on the population-level intercept and SD = 2 on the predictor coefficients, including the DID terms.[25,39,40] We estimated the effect of the intervention on human-associated indicator prevalence using logistic regression and the prevalence odds ratio (POR) as the measure of effect. Under the logistic link,

priors are defined on the continuous log-odds scale but are best understood on the probability scale, which is more intuitive but constrained between 0 and 1 prevalence.[41] As such, the population-level intercept and predictor coefficients were given regularizing normal priors with SD = 1.5 and SD = 0.5, respectively, which on the probability scale corresponded to a ~95% chance the population-level prevalence was between 0.05 and 0.95 and an effect of up to ± 0.23 for each predictor at a population-level prevalence of 0.5.[40] This represents a substantial effect size on the probability scale—a reduction in absolute risk from 50% to 27%, for example—which would be unexpected in an environmental context.[23] Furthermore, as a soft constraint, estimates could still exceed this effect size given sufficient and sufficiently strong data in support of a larger effect.

The diagnostic accuracy-corrected models of human fecal contamination prevalence shared the same basic structure as the human-associated indicator prevalence model described above. Accordingly, we set the same priors for the population-level intercept, the DID terms, and the covariates in the adjustment set:

$$y_i \sim Bernoulli(p_i)$$

$$p_i = \pi_i$$

$$logit(\pi_i) = \beta^0 + \beta^P P_i + \beta^T T_i + \beta^{DID} P_i \times T_i + \boldsymbol{X}_i \boldsymbol{\gamma} \qquad \textbf{Model 1}$$

$$\beta^0 \sim normal(0,1.5)$$

$$\beta^P, \beta^T, \beta^{DID}, \boldsymbol{\gamma} \sim normal(0,0.5)$$

Model 2 introduced parameters for sensitivity (Se) and specificity (Sp), which required priors as well. Due the modest sample sizes of the local validation analysis, particularly of human samples (n=14), the use of non-informative priors risks unrealistically broad sensitivity and specificity estimates.[23,42] As we possessed prior information on the typical ranges for Se and

Sp in resource-limited settings (Table S9), we used informative $beta(3,2)$ priors as soft constraints, corresponding to a 95% chance they fall between 0.19 and 0.93 on the probability scale, with mean 0.6.

$$y_i \sim Bernoulli(p_i)$$
$$p_i = Se \times \pi_i + (1 - Sp)(1 - \pi_i)$$
$$logit(\pi_i) = \beta^0 + \beta^P P_i + \beta^T T_i + \beta^{DID} P_i \times T_i + \boldsymbol{X}_i \boldsymbol{\gamma}$$
$$y^{Se} \sim binomial(n^{Se}, Se) \qquad \textbf{Model 2}$$
$$y^{Sp} \sim binomial(n^{Sp}, Sp)$$
$$Se, Sp \sim beta(3,2)$$
$$\beta^0 \sim normal(0,1.5);\ \beta^P, \beta^T, \beta^{DID}, \boldsymbol{\gamma} \sim normal(0,0.5)$$

Because our validation sample set was small and performance estimates vary widely between studies, we fit a third model (Model 3) featuring a meta-analysis of indicator sensitivity and specificity. We assumed the log-odds of the sensitivity in the kth study, $Se_{[k]}$, were normally distributed with mean μ^{Se} and SD σ^{Se}, with an equivalent structure for specificity. We assigned μ^{Se} and μ^{Sp} $normal(0.5, 1)$ priors, which provided approximately equivalent coverage on the probability scale as the previous beta priors in Model 2, with weakly-informative $normal^{+}(0, 0.5)$ [half-normal] priors on σ^{Se} and σ^{Sp}, as discussed by Gelman and Carpenter.[23]

$$y_i \sim Bernoulli(p_i)$$

$$p_i = Se_{[1]} \times \pi_i + \left(1 - Sp_{[1]}\right)(1 - \pi_i)$$

$$logit(\pi_i) = \beta^0 + \beta^P P_i + \beta^T T_i + \beta^{DID} P_i \times T_i + \boldsymbol{X}_i \boldsymbol{\gamma}$$

$$y_{[k]}^{Se} \sim binomial\left(n_{[k]}^{Se}, Se_{[k]}\right);\ y_{[k]}^{Sp} \sim binomial\left(n_{[k]}^{Sp}, Sp_{[k]}\right)$$

$$logit\left(Se_{[k]}\right) \sim normal(\mu^{Se}, \sigma^{Se});\ logit\left(Sp_{[k]}\right) \sim normal(\mu^{Sp}, \sigma^{Sp})$$

Model 3

$$\mu^{Se}, \mu^{Sp} \sim normal(0.5, 1)$$

$$\sigma^{Se}, \sigma^{Sp} \sim normal^{+}(0, 0.5)$$

$$\beta^0 \sim normal(0{,}1.5);\ \beta^P, \beta^T, \beta^{DID}, \boldsymbol{\gamma} \sim normal(0{,}0.5)$$

Model 4 followed the same structure as Model 3, but incorporated two different fecal indicators, represented as $y_i^{[assay]}$, with a corresponding duplication of the sensitivity and specificity model components. All priors remain the same as Model 3.

$$y_i^{[assay]} \sim Bernoulli\left(p_i^{[assay]}\right)$$

$$p_i^{[assay]} = Se_{[1]}^{[assay]} \times \pi_i + \left(1 - Sp_{[1]}^{[assay]}\right)(1 - \pi_i)$$

$$logit(\pi_i) = \beta^0 + \beta^P P_i + \beta^T T_i + \beta^{DID} P_i \times T_i + \boldsymbol{X}_i \boldsymbol{\gamma}$$

$$y_{[k]}^{[assay],Se} \sim binomial\left(n_{[k]}^{[assay],Se,}, Se_{[k]}^{[assay]}\right);\ y_{[k]}^{[assay],Sp} \sim binomial\left(n_{[k]}^{[assay],Sp,}, Sp_{[k]}^{[assay]}\right)$$

$$logit\left(Se_{[k]}^{[assay]}\right) \sim normal\left(\mu^{[assay],Se}, \sigma^{[assay],Se}\right)$$

Model 4

$$logit\left(Sp_{[k]}^{[assay]}\right) \sim normal\left(\mu^{[assay],Sp}, \sigma^{[assay],Sp}\right)$$

$$\mu^{[assay],Se}, \mu^{[assay],Sp} \sim normal(0.5, 1)$$

$$\sigma^{[assay],Se}, \sigma^{[assay],Sp} \sim normal^{+}(0, 0.5)$$

$$\beta^0 \sim normal(0{,}1.5);\ \beta^P, \beta^T, \beta^{DID}, \boldsymbol{\gamma} \sim normal(0{,}0.5)$$

Finally, Model 5 included multiple sample types with type-specific prevalence variables, $\pi_i^{[type]}$, derived from sample-type specific intercepts $\alpha^{[type]}$ and a shared compound-varying

intercept $\alpha_{[j]}^{comp}$ corresponding to the $j$the compound that replaced the previously fixed, population-level intercept β^0. As before, all β and γ parameters were given normal priors with SD = 0.5. We assumed $\alpha_{[j]}^{comp} \sim normal(\mu^{comp}, \sigma^{comp})$ and $\alpha^{[type]} \sim normal(0, \sigma^{type})$, with half-normal, SD = 0.5 priors on σ^{comp} and σ^{type} and the same $normal(0, 1.5)$ prior on μ^{comp} used for the fixed intercept β^0 in previous models.

Model 5

$$y_i^{[assay,type]} \sim Bernoulli\left(p_i^{[assay,type]}\right)$$

$$p_i^{[assay,type]} = Se_{[1]}^{[assay,type]} \times \pi_i^{[type]} + \left(1 - Sp_{[1]}^{[assay,type]}\right)\left(1 - \pi_i^{[type]}\right)$$

$$logit(\pi_i^{[type]}) = \alpha^{[type]} + \boldsymbol{X}_i^{[type]}\boldsymbol{\gamma}^{[type]} + logit\left(\pi_{[j]}^{comp}\right)$$

$$logit\left(\pi_{[j]}^{comp}\right) = \alpha_{[j]}^{comp} + \beta^P P_{[j]} + \beta^T T_{[j]} + \beta^{DID} P_{[j]} \times T_{[j]} + \boldsymbol{X}_{[j]}^{comp}\boldsymbol{\gamma}^{comp}$$

$$y_{[k]}^{[assay],Se} \sim binomial\left(n_{[k]}^{[assay],Se,}, Se_{[k]}^{[assay]}\right);\ y_{[k]}^{[assay],Sp} \sim binomial\left(n_{[k]}^{[assay],Sp,}, Sp_{[k]}^{[assay]}\right)$$

$$logit\left(Se_{[k]}^{[assay]}\right) \sim normal\left(\mu^{[assay],Se}, \sigma^{[assay],Se}\right)$$

$$logit\left(Sp_{[k]}^{[assay]}\right) \sim normal\left(\mu^{[assay],Sp}, \sigma^{[assay],Sp}\right)$$

$$\alpha_{[j]}^{comp} \sim normal(\mu^{comp}, \sigma^{comp});\ \alpha^{[type]} \sim normal(0, \sigma^{type})$$

$$\mu^{[assay],Se}, \mu^{[assay],Sp} \sim normal(0.5, 1);\ \mu^{comp} \sim normal(0,1.5)$$

$$\sigma^{[type]}, \sigma^{[assay],Se}, \sigma^{[assay],Sp} \sim normal^{+}(0, 0.5)$$

$$\beta^P, \beta^T, \beta^{DID}, \boldsymbol{\gamma}^{[type]}, \boldsymbol{\gamma}^{comp} \sim normal(0,0.5)$$

S14. Stan code

S14.1. Model 1

```
// Model 1:
// single sample type, single target
// no correction for sensitivity/specificity
// no compound-varying intercepts

data{
// sample data
  int<lower = 0> N_samp;  // number of sample observations
  int<lower = 0, upper = 1> y_samp[N_samp];  // sample observations
  vector<lower = 0, upper = 1>[N_samp] phase;  // survey phase
  vector<lower = 0, upper = 1>[N_samp] treat;  // treatment arm
  vector<lower = 0, upper = 1>[N_samp] did;  // phase*treat interaction
  int<lower = 0> K;  // number of predictors
  matrix[N_samp, K] X;  // predictor values

// prior predictive check?
  int<lower = 0, upper = 1> prior_only;  // toggle on to sample priors only
}

parameters{
// linear model parameters
  real b0;  // intercept
  real bP;  // phase
  real bT;  // treatment
  real bD;  // DID
  vector[K] g;  // predictor coefficients
}

model{
// linear model for probability of human contamination
  vector[N_samp] p_samp = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// Likelihood
  if(prior_only == 0){
    y_samp ~ binomial(1, p_samp);
  }

// linear model priors
  b0 ~ normal(0, 1.5);
  bD ~ normal(0, 0.5);
  bT ~ normal(0, 0.5);
  bP ~ normal(0, 0.5);
  g ~ normal(0, 0.5);
}

generated quantities{
// posterior predictions (or prior predictions, if prior_only == 1)
// define predicted variables
  int<lower = 0, upper = 1> y_pred[N_samp];  // predicted sample observations
  int<lower = 0> n_pos;  // number of predicted positives
  real<lower = 0, upper = 1> p_samp_avg;  // mean target prevalence in samples
```

```
// calculate human contamination probability
  vector[N_samp] p_samp_sim = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// predict sample observations
  y_pred = binomial_rng(1, p_samp_sim);
  n_pos = sum(y_pred);

// summarise prevalence calculations
  p_samp_avg = mean(p_samp_sim);
}
```

S14.2. Model 2

```
// Model 2:
// single sample type, single target
// diagnostic performance data from this study only
// no compound-varying intercepts

data{
// sample data
  int<lower = 0> N_samp;  // number of sample observations
  int<lower = 0, upper = 1> y_samp[N_samp];  // sample observations
  vector<lower = 0, upper = 1>[N_samp] phase;  // survey phase
  vector<lower = 0, upper = 1>[N_samp] treat;  // treatment arm
  vector<lower = 0, upper = 1>[N_samp] did;  // phase*treat interaction
  int<lower = 0> K;  // number of predictors
  matrix[N_samp, K] X;  // predictor values

// sens/spec data
  int<lower = 0> y_spec;  // number of true negatives observed
  int<lower = 0> n_spec;  // number of non-target samples
  int<lower = 0> y_sens;  // number of true positives observed
  int<lower = 0> n_sens;  // number of target samples

// prior predictive check?
  int<lower = 0, upper = 1> prior_only;  // toggle on to sample priors only
}

parameters{
// linear model parameters
  real b0;  // intercept
  real bP;  // phase
  real bT;  // treatment
  real bD;  // DID
  vector[K] g;  // predictor coefficients

// diagnostic performance parameters
  real<lower=0, upper=1> spec;  // specificity
  real<lower=0, upper=1> sens;  // sensitivity
}

model{
// linear model for probability of human contamination
  vector[N_samp] p = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// adjust for sens/spec
  vector[N_samp] p_samp = sens * p + (1 - spec) * (1 - p);

// Likelihoods
  if(prior_only == 0){
  // samples
    y_samp ~ binomial(1, p_samp);

  // validation studies
    y_spec ~ binomial(n_spec, spec);
    y_sens ~ binomial(n_sens, sens);
  }
```

```
// linear model priors
  b0 ~ normal(0, 1.5);
  bD ~ normal(0, 0.5);
  bT ~ normal(0, 0.5);
  bP ~ normal(0, 0.5);
  g ~ normal(0, 0.5);

// validation priors
  sens ~ beta(3, 2);
  spec ~ beta(3, 2);
}

generated quantities{
// posterior predictions (or prior predictions, if prior_only == 1)
// define predicted variables
  int<lower = 0, upper = 1> y_pred[N_samp];  // predicted sample observations
  int<lower = 0> n_pos;  // number of predicted positives
  real<lower = 0, upper = 1> p_avg;  //  mean human contamination prevalence
  real<lower = 0, upper = 1> p_samp_avg;  // mean target prevalence in samples

// calculate human contamination probability
  vector[N_samp] p_sim = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// adjust for sens/spec
  vector[N_samp] p_samp_sim = sens * p_sim + (1 - spec) * (1 - p_sim);

// predict sample observations
  y_pred = binomial_rng(1, p_samp_sim);
  n_pos = sum(y_pred);

// summarise prevalence calculations
  p_avg = mean(p_sim);
  p_samp_avg = mean(p_samp_sim);
}
```

S14.3. Model 3

```
// Model 3:
// single sample type, single target
// diagnostic performance meta-analysis
// no compound-varying intercepts

data{
// sample data
  int<lower = 0> N_samp;  // number of sample observations
  int<lower = 0, upper = 1> y_samp[N_samp];  // sample observations
  vector<lower = 0, upper = 1>[N_samp] phase;  // survey phase
  vector<lower = 0, upper = 1>[N_samp] treat;  // treatment arm
  vector<lower = 0, upper = 1>[N_samp] did;  // phase*treat interaction
  int<lower = 0> K;  // number of predictors
  matrix[N_samp, K] X;  // predictor values

// sens/spec data
  int<lower = 0> J_spec;
  int<lower = 0> y_spec[J_spec];
  int<lower = 0> n_spec[J_spec];
  int<lower = 0> J_sens;
  int<lower = 0> y_sens[J_sens];
  int<lower = 0> n_sens[J_sens];

// prior predictive check?
  int<lower = 0, upper = 1> prior_only;  // toggle on to sample priors only
}

parameters{
// linear model parameters
  real b0;  // intercept
  real bP;  // phase
  real bT;  // treatment
  real bD;  // DID
  vector[K] g;  // predictor coefficients

// sens/spec meta-analysis parameters
  real mu_logit_spec;  // mean spec on the logit scale
  real mu_logit_sens;
  real<lower = 0> sigma_logit_spec;  // spec SD on logit scale
  real<lower = 0> sigma_logit_sens;
  // non-centered parameterization of logit-transformed sens/spec
  vector<offset = mu_logit_spec, multiplier = sigma_logit_spec>[J_spec] logit_spec;
  vector<offset = mu_logit_sens, multiplier = sigma_logit_sens>[J_sens] logit_sens;
}

transformed parameters{
// recover sens/spec on probability scale
  vector[J_spec] spec = inv_logit(logit_spec);
  vector[J_sens] sens = inv_logit(logit_sens);
}

model{
// linear model for probability of human contamination
  vector[N_samp] p = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);
```

```
// adjust for sens/spec
  vector[N_samp] p_samp = sens[1] * p + (1 - spec[1]) * (1 - p);

// Likelihoods
  if(prior_only == 0){
  // samples
    y_samp ~ binomial(1, p_samp);

  // validation studies
    y_spec ~ binomial(n_spec, spec);
    y_sens ~ binomial(n_sens, sens);
  }

// linear model priors
  b0 ~ normal(0, 1.5);
  bD ~ normal(0, 0.5);
  bT ~ normal(0, 0.5);
  bP ~ normal(0, 0.5);
  g ~ normal(0, 0.5);

// validation priors
  logit_spec ~ normal(mu_logit_spec, sigma_logit_spec);
  logit_sens ~ normal(mu_logit_sens, sigma_logit_sens);
  sigma_logit_spec ~ normal(0, .5);
  sigma_logit_sens ~ normal(0, .5);
  mu_logit_spec ~ normal(.5, 1);
  mu_logit_sens ~ normal(.5, 1);
}

generated quantities{
// posterior predictions (or prior predictions, if prior_only == 1)
// define predicted variables
  int<lower = 0, upper = 1> y_pred[N_samp];  // predicted sample observations
  int<lower = 0> n_pos;  // number of predicted positives
  real<lower = 0, upper = 1> p_avg;  //  mean human contamination prevalence
  real<lower = 0, upper = 1> p_samp_avg;  // mean target prevalence in samples

// calculate human contamination probability
  vector[N_samp] p_sim = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// adjust for sens/spec
  vector[N_samp] p_samp_sim = sens[1] * p_sim + (1 - spec[1]) * (1 - p_sim);

// predict sample observations
  y_pred = binomial_rng(1, p_samp_sim);
  n_pos = sum(y_pred);

// summarise prevalence calculations
  p_avg = mean(p_sim);
  p_samp_avg = mean(p_samp_sim);
}
```

S14.4. Model 4

```
// Model 4:
// single sample type, two targets
// diagnostic performance meta-analysis
// no compound-varying intercepts

data{
// sample data
  int<lower = 0> N_samp;  // number of sample observations
  int<lower = 0, upper = 1> y_hf[N_samp];  // HF183 observations
  int<lower = 0, upper = 1> y_mn[N_samp];  // Mnif observations
  vector<lower = 0, upper = 1>[N_samp] phase;  // survey phase
  vector<lower = 0, upper = 1>[N_samp] treat;  // treatment arm
  vector<lower = 0, upper = 1>[N_samp] did;  // phase*treat interaction
  int<lower = 0> K;  // number of predictors
  matrix[N_samp, K] X;  // predictor values

// sens/spec data
  int<lower = 0> J_spec_hf;
  int<lower = 0> y_spec_hf[J_spec_hf];
  int<lower = 0> n_spec_hf[J_spec_hf];
  int<lower = 0> J_sens_hf;
  int<lower = 0> y_sens_hf[J_sens_hf];
  int<lower = 0> n_sens_hf[J_sens_hf];
  int<lower = 0> J_spec_mn;
  int<lower = 0> y_spec_mn[J_spec_mn];
  int<lower = 0> n_spec_mn[J_spec_mn];
  int<lower = 0> J_sens_mn;
  int<lower = 0> y_sens_mn[J_sens_mn];
  int<lower = 0> n_sens_mn[J_sens_mn];

// prior predictive check?
  int<lower = 0, upper = 1> prior_only;  // toggle on to sample priors only
}

parameters{
// linear model parameters
  real b0;  // intercept
  real bP;  // phase
  real bT;  // treatment
  real bD;  // DID
  vector[K] g;  // predictor coefficients

// sens/spec meta-analysis parameters
  real mu_logit_spec_hf;  // mean spec for HF183 on the logit scale
  real mu_logit_sens_hf;
  real mu_logit_spec_mn;
  real mu_logit_sens_mn;
  real<lower = 0> sigma_logit_spec_hf;  // spec SD on logit scale
  real<lower = 0> sigma_logit_sens_hf;
  real<lower = 0> sigma_logit_spec_mn;
  real<lower = 0> sigma_logit_sens_mn;
  // non-centered parameterization of logit-transformed sens/spec for each target
  vector<offset = mu_logit_spec_hf, multiplier = sigma_logit_spec_hf>[J_spec_hf]
logit_spec_hf;
```

```
  vector<offset = mu_logit_sens_hf, multiplier = sigma_logit_sens_hf>[J_sens_hf]
logit_sens_hf;
  vector<offset = mu_logit_spec_mn, multiplier = sigma_logit_spec_mn>[J_spec_mn]
logit_spec_mn;
  vector<offset = mu_logit_sens_mn, multiplier = sigma_logit_sens_mn>[J_sens_mn]
logit_sens_mn;
}

transformed parameters{
// recover sens/spec on probability scale
  vector[J_spec_hf] spec_hf = inv_logit(logit_spec_hf);
  vector[J_sens_hf] sens_hf = inv_logit(logit_sens_hf);
  vector[J_spec_mn] spec_mn = inv_logit(logit_spec_mn);
  vector[J_sens_mn] sens_mn = inv_logit(logit_sens_mn);
}

model{
// linear model for probability of human contamination
  vector[N_samp] p = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// adjust for sens/spec
  // by convention the first sens/spec element represents this current study
  vector[N_samp] p_hf = sens_hf[1] * p + (1 - spec_hf[1]) * (1 - p);
  vector[N_samp] p_mn = sens_mn[1] * p + (1 - spec_mn[1]) * (1 - p);

// Likelihoods
  if(prior_only == 0){
  // samples
    y_hf ~ binomial(1, p_hf);
    y_mn ~ binomial(1, p_mn);

  // validation studies
    y_spec_hf ~ binomial(n_spec_hf, spec_hf);
    y_sens_hf ~ binomial(n_sens_hf, sens_hf);
    y_spec_mn ~ binomial(n_spec_mn, spec_mn);
    y_sens_mn ~ binomial(n_sens_mn, sens_mn);
  }

// linear model priors
  b0 ~ normal(0, 1.5);
  bD ~ normal(0, 0.5);
  bT ~ normal(0, 0.5);
  bP ~ normal(0, 0.5);
  g ~ normal(0, 0.5);

// validation priors
  logit_spec_hf ~ normal(mu_logit_spec_hf, sigma_logit_spec_hf);
  logit_sens_hf ~ normal(mu_logit_sens_hf, sigma_logit_sens_hf);
  sigma_logit_spec_hf ~ normal(0, .5);
  sigma_logit_sens_hf ~ normal(0, .5);
  mu_logit_spec_hf ~ normal(.5, 1);
  mu_logit_sens_hf ~ normal(.5, 1);
  logit_spec_mn ~ normal(mu_logit_spec_mn, sigma_logit_spec_mn);
  logit_sens_mn ~ normal(mu_logit_sens_mn, sigma_logit_sens_mn);
  sigma_logit_spec_mn ~ normal(0, .5);
  sigma_logit_sens_mn ~ normal(0, .5);
  mu_logit_spec_mn ~ normal(.5, 1);
```

```
  mu_logit_sens_mn ~ normal(.5, 1);
}

generated quantities{
// posterior predictions (or prior predictions, if prior_only == 1)
// define predicted variables
  int<lower = 0, upper = 1> y_pred_hf[N_samp];  // predicted HF183 observations
  int<lower = 0> n_pos_hf;  // number of predicted HF183 positives
  real<lower = 0, upper = 1> p_samp_avg_hf;  // mean HF183 prevalence in samples
  int<lower = 0, upper = 1> y_pred_mn[N_samp];  // predicted Mnif observations
  int<lower = 0> n_pos_mn;  // number of predicted Mnif positives
  real<lower = 0, upper = 1> p_samp_avg_mn;  // mean Mnif prevalence in samples
  real<lower = 0, upper = 1> p_avg;  //  mean human contamination prevalence

// calculate human contamination probability
  vector[N_samp] p_sim = inv_logit(b0 + bP*phase + bT*treat + bD*did + X*g);

// adjust for sens/spec
  vector[N_samp] p_hf_sim = sens_hf[1] * p_sim + (1 - spec_hf[1]) * (1 - p_sim);
  vector[N_samp] p_mn_sim = sens_mn[1] * p_sim + (1 - spec_mn[1]) * (1 - p_sim);

// predict sample observations
  y_pred_hf = binomial_rng(1, p_hf_sim);
  n_pos_hf = sum(y_pred_hf);
  y_pred_mn = binomial_rng(1, p_mn_sim);
  n_pos_mn = sum(y_pred_mn);

// summarise prevalence calculations
  p_avg = mean(p_sim);
  p_samp_avg_hf = mean(p_hf_sim);
  p_samp_avg_mn = mean(p_mn_sim);
}
```

S14.5. Model 5

```
// Model 5:
// three sample types, two targets
// diagnostic performance meta-analysis
// compound-varying intercept, type-varying intercept

data{
  int<lower = 0> N_type;  // number of sample types considered

// compound data
  int<lower = 0> J_comp;  // number of unique compounds
  int<lower = 0> N_comp;  // number of compound observations
  int<lower = 1, upper = J_comp> comp[N_comp]; // compound index
  int<lower = 0> K_comp;  // number of compound-level predictors
  matrix[N_comp, K_comp] X_comp;  // compound-level predictors

// hw sample data
  int<lower = 0> N_hw;  // number of compound observations
  int<lower = 0, upper = 1> y_hw_hf[N_hw];  // hw HF183 observations
  // int<lower = 0, upper = 1> y_hw_mn[N_hw];  // no hw Mnif observations
  int<lower = 1, upper = J_comp> comp_hw[N_hw]; // hw compound index
  int<lower = 0> K_hw;  // number of hw sample-level predictors
  matrix[N_hw, K_hw] X_hw;  // hw predictors

// ds sample data
  int<lower = 0> N_ds;  // number of compound observations
  int<lower = 0, upper = 1> y_ds_hf[N_ds];  // ds HF183 observations
  int<lower = 0, upper = 1> y_ds_mn[N_ds];  // ds Mnif observations
  int<lower = 1, upper = J_comp> comp_ds[N_ds]; // ds compound index
  int<lower = 0> K_ds;  // number of ds sample-level predictors
  matrix[N_ds, K_ds] X_ds;  // ds predictors

// ls sample data
  int<lower = 0> N_ls;  // number of compound observations
  int<lower = 0, upper = 1> y_ls_hf[N_ls];  // ls HF183 observations
  int<lower = 0, upper = 1> y_ls_mn[N_ls];  // ls Mnif observations
  int<lower = 1, upper = J_comp> comp_ls[N_ls]; // ls compound index
  int<lower = 0> K_ls;  // number of ls sample-level predictors
  matrix[N_ls, K_ls] X_ls;  // ls predictors

// sens/spec data
  int<lower = 0> J_spec_hf;
  int<lower = 0> y_spec_hf[J_spec_hf];
  int<lower = 0> n_spec_hf[J_spec_hf];
  int<lower = 0> J_sens_hf;
  int<lower = 0> y_sens_hf[J_sens_hf];
  int<lower = 0> n_sens_hf[J_sens_hf];
  int<lower = 0> J_spec_mn;
  int<lower = 0> y_spec_mn[J_spec_mn];
  int<lower = 0> n_spec_mn[J_spec_mn];
  int<lower = 0> J_sens_mn;
  int<lower = 0> y_sens_mn[J_sens_mn];
  int<lower = 0> n_sens_mn[J_sens_mn];

// prior predictive check?
```

```
  int<lower = 0, upper = 1> prior_only;
}

parameters{
// compound prevalence linear model parameters
  vector[K_comp] b_comp; // compound-level coefficients
  real mu_comp;  // mean of compound-varying intercept
  real<lower = 0> sigma_comp;  // SD of compound-varying intercept
  vector<offset = mu_comp, multiplier = sigma_comp>[J_comp] a_comp;

// sample-level parameters
  vector[K_hw] g_hw;  // hw predictor coefficients
  vector[K_ds] g_ds;  // ds predictor coefficients
  vector[K_ls] g_ls;  // ls predictor coefficients
  real<lower = 0> sigma_type;  // SD of sample differences
  vector<multiplier = sigma_type>[N_type] a_type;  // sample type-varying intercept

// sens/spec meta-analysis parameters
  real mu_logit_spec_hf;
  real mu_logit_sens_hf;
  real mu_logit_spec_mn;
  real mu_logit_sens_mn;
  real<lower = 0> sigma_logit_spec_hf;
  real<lower = 0> sigma_logit_sens_hf;
  real<lower = 0> sigma_logit_spec_mn;
  real<lower = 0> sigma_logit_sens_mn;
  vector<offset = mu_logit_spec_hf, multiplier = sigma_logit_spec_hf>[J_spec_hf]
logit_spec_hf;
  vector<offset = mu_logit_sens_hf, multiplier = sigma_logit_sens_hf>[J_sens_hf]
logit_sens_hf;
  vector<offset = mu_logit_spec_mn, multiplier = sigma_logit_spec_mn>[J_spec_mn]
logit_spec_mn;
  vector<offset = mu_logit_sens_mn, multiplier = sigma_logit_sens_mn>[J_sens_mn]
logit_sens_mn;
}

transformed parameters{
  vector[J_spec_hf] spec_hf = inv_logit(logit_spec_hf);
  vector[J_sens_hf] sens_hf = inv_logit(logit_sens_hf);
  vector[J_spec_mn] spec_mn = inv_logit(logit_spec_mn);
  vector[J_sens_mn] sens_mn = inv_logit(logit_sens_mn);
}

model{
// linear model for compound contamination
  vector[N_comp] logit_p_comp = a_comp[comp] + X_comp * b_comp;

// linear models for sample-type specific prevalence
  vector[N_hw] p_hw = inv_logit(logit_p_comp[comp_hw] + a_type[1] + X_hw * g_hw);
  vector[N_ds] p_ds = inv_logit(logit_p_comp[comp_ds] + a_type[2] + X_ds * g_ds);
  vector[N_ls] p_ls = inv_logit(logit_p_comp[comp_ls] + a_type[3] + X_ls * g_ls);

// adjust for sens/spec
  vector[N_hw] p_hw_hf = sens_hf[1] * p_hw + (1 - spec_hf[1]) * (1 - p_hw);
  vector[N_hw] p_hw_mn = sens_mn[1] * p_hw + (1 - spec_mn[1]) * (1 - p_hw);
  vector[N_ds] p_ds_hf = sens_hf[1] * p_ds + (1 - spec_hf[1]) * (1 - p_ds);
  vector[N_ds] p_ds_mn = sens_mn[1] * p_ds + (1 - spec_mn[1]) * (1 - p_ds);
```

```
  vector[N_ls] p_ls_hf = sens_hf[1] * p_ls + (1 - spec_hf[1]) * (1 - p_ls);
  vector[N_ls] p_ls_mn = sens_mn[1] * p_ls + (1 - spec_mn[1]) * (1 - p_ls);

// Likelihoods
  if(prior_only == 0){
  // samples
    y_hw_hf ~ binomial(1, p_hw_hf);
    // y_hw_mn ~ binomial(1, p_hw_mn); no Mnif for HW samples
    y_ds_hf ~ binomial(1, p_ds_hf);
    y_ds_mn ~ binomial(1, p_ds_mn);
    y_ls_hf ~ binomial(1, p_ls_hf);
    y_ls_mn ~ binomial(1, p_ls_mn);

  // validation studies
    y_spec_hf ~ binomial(n_spec_hf, spec_hf);
    y_sens_hf ~ binomial(n_sens_hf, sens_hf);
    y_spec_mn ~ binomial(n_spec_mn, spec_mn);
    y_sens_mn ~ binomial(n_sens_mn, sens_mn);
  }

// sample priors
  // compound-level
  a_comp ~ normal(mu_comp, sigma_comp);
  mu_comp ~ normal(0, 1.5);
  sigma_comp ~ normal(0, 0.5);
  b_comp ~ normal(0, 0.5);

  // sample-level
  a_type ~ normal(0, sigma_type);
  sigma_type ~ normal(0, 0.5);
  g_hw ~ normal(0, 0.5);
  g_ds ~ normal(0, 0.5);
  g_ls ~ normal(0, 0.5);

// validation priors
  logit_spec_hf ~ normal(mu_logit_spec_hf, sigma_logit_spec_hf);
  logit_sens_hf ~ normal(mu_logit_sens_hf, sigma_logit_sens_hf);
  sigma_logit_spec_hf ~ normal(0, .5);
  sigma_logit_sens_hf ~ normal(0, .5);
  mu_logit_spec_hf ~ normal(.5, 1);
  mu_logit_sens_hf ~ normal(.5, 1);
  logit_spec_mn ~ normal(mu_logit_spec_mn, sigma_logit_spec_mn);
  logit_sens_mn ~ normal(mu_logit_sens_mn, sigma_logit_sens_mn);
  sigma_logit_spec_mn ~ normal(0, .5);
  sigma_logit_sens_mn ~ normal(0, .5);
  mu_logit_spec_mn ~ normal(.5, 1);
  mu_logit_sens_mn ~ normal(.5, 1);
}

generated quantities{
// simulated sample outcome containers
  int y_hw_hf_sim[N_hw];
  int y_hw_mn_sim[N_hw];
  int y_ds_hf_sim[N_ds];
  int y_ds_mn_sim[N_ds];
  int y_ls_hf_sim[N_ls];
  int y_ls_mn_sim[N_ls];
```

```
  int n_pos_hw_hf;
  int n_pos_hw_mn;
  int n_pos_ds_hf;
  int n_pos_ds_mn;
  int n_pos_ls_hf;
  int n_pos_ls_mn;
  real<lower = 0, upper = 1> p_avg_comp;
  real<lower = 0, upper = 1> p_avg_hw;
  real<lower = 0, upper = 1> p_avg_ds;
  real<lower = 0, upper = 1> p_avg_ls;
  real<lower = 0, upper = 1> p_samp_avg_hw_hf;
  real<lower = 0, upper = 1> p_samp_avg_hw_mn;
  real<lower = 0, upper = 1> p_samp_avg_ds_hf;
  real<lower = 0, upper = 1> p_samp_avg_ds_mn;
  real<lower = 0, upper = 1> p_samp_avg_ls_hf;
  real<lower = 0, upper = 1> p_samp_avg_ls_mn;

// simulate type-specific probabilities
  vector[N_comp] logit_p_comp_sim = a_comp[comp] + X_comp * b_comp;
  vector[N_comp] p_comp_sim = inv_logit(logit_p_comp_sim);
  vector[N_hw] p_hw_sim = inv_logit(logit_p_comp_sim[comp_hw] + a_type[1] + X_hw *
g_hw);
  vector[N_ds] p_ds_sim = inv_logit(logit_p_comp_sim[comp_ds] + a_type[2] + X_ds *
g_ds);
  vector[N_ls] p_ls_sim = inv_logit(logit_p_comp_sim[comp_ls] + a_type[3] + X_ls *
g_ls);
// adjust simulations for sens/spec
  vector[N_hw] p_hw_hf_sim = sens_hf[1] * p_hw_sim + (1 - spec_hf[1]) * (1 -
p_hw_sim);
  vector[N_hw] p_hw_mn_sim = sens_mn[1] * p_hw_sim + (1 - spec_mn[1]) * (1 -
p_hw_sim);
  vector[N_ds] p_ds_hf_sim = sens_hf[1] * p_ds_sim + (1 - spec_hf[1]) * (1 -
p_ds_sim);
  vector[N_ds] p_ds_mn_sim = sens_mn[1] * p_ds_sim + (1 - spec_mn[1]) * (1 -
p_ds_sim);
  vector[N_ls] p_ls_hf_sim = sens_hf[1] * p_ls_sim + (1 - spec_hf[1]) * (1 -
p_ls_sim);
  vector[N_ls] p_ls_mn_sim = sens_mn[1] * p_ls_sim + (1 - spec_mn[1]) * (1 -
p_ls_sim);

// simulate sample observations
  y_hw_hf_sim = binomial_rng(1, p_hw_hf_sim);
  y_hw_mn_sim = binomial_rng(1, p_hw_mn_sim);
  y_ds_hf_sim = binomial_rng(1, p_ds_hf_sim);
  y_ds_mn_sim = binomial_rng(1, p_ds_mn_sim);
  y_ls_hf_sim = binomial_rng(1, p_ls_hf_sim);
  y_ls_mn_sim = binomial_rng(1, p_ls_mn_sim);

// summarize simulated samples
  n_pos_hw_hf = sum(y_hw_hf_sim);
  n_pos_hw_mn = sum(y_hw_mn_sim);
  n_pos_ds_hf = sum(y_ds_hf_sim);
  n_pos_ds_mn = sum(y_ds_mn_sim);
  n_pos_ls_hf = sum(y_ls_hf_sim);
  n_pos_ls_mn = sum(y_ls_mn_sim);

// mean prevalence predictions
```

```
  p_avg_comp = mean(p_comp_sim);
  p_avg_hw = mean(p_hw_sim);
  p_avg_ds = mean(p_ds_sim);
  p_avg_ls = mean(p_ls_sim);
  p_samp_avg_hw_hf = mean(p_hw_hf_sim);
  p_samp_avg_hw_mn = mean(p_hw_mn_sim);
  p_samp_avg_ds_hf = mean(p_ds_hf_sim);
  p_samp_avg_ds_mn = mean(p_ds_mn_sim);
  p_samp_avg_ls_hf = mean(p_ls_hf_sim);
  p_samp_avg_ls_mn = mean(p_ls_mn_sim);
}
```

S15. Supplementary References

(1) Knee, J.; Sumner, T.; Adriano, Z.; Anderson, C.; Bush, F.; Capone, D.; Casmo, V.; Holcomb, D. A.; Kolsky, P.; MacDougall, A.; Molotkova, E.; Braga, J. M.; Russo, C.; Schmidt, W. P.; Stewart, J.; Zambrana, W.; Zuin, V.; Nalá, R.; Cumming, O.; Brown, J. 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. https://doi.org/10.7554/eLife.62278.

(2) Holcomb, D. A.; Knee, J.; Sumner, T.; Adriano, Z.; de Bruijn, E.; Nalá, R.; Cumming, O.; Brown, J.; Stewart, J. R. Human Fecal Contamination of Water, Soil, and Surfaces in Households Sharing Poor-Quality Sanitation Facilities in Maputo, Mozambique. *International Journal of Hygiene and Environmental Health* **2020**, *226*, 113496. https://doi.org/10.1016/j.ijheh.2020.113496.

(3) Chern, E. C.; Siefring, S.; Paar, J.; Doolittle, M.; Haugland, R. A. Comparison of Quantitative PCR Assays for Escherichia Coli Targeting Ribosomal RNA and Single Copy Genes. *Letters in Applied Microbiology* **2011**, *52* (3), 298–306. https://doi.org/10.1111/j.1472-765X.2010.03001.x.

(4) Green, H. C.; Haugland, R. A.; Varma, M.; Millen, H. T.; Borchardt, M. A.; Field, K. G.; Walters, W. A.; Knight, R.; Sivaganesan, M.; Kelty, C. A.; Shanks, O. C. Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Applied and Environmental Microbiology* **2014**, *80* (10), 3086–3094. https://doi.org/10.1128/AEM.04137-13.

(5) Johnston, C.; Ufnar, J. a; Griffith, J. F.; Gooch, J. a; Stewart, J. R. A Real-Time QPCR Assay for the Detection of the NifH Gene of Methanobrevibacter Smithii, a Potential

Indicator of Sewage Pollution. *Journal of applied microbiology* **2010**, *109* (6), 1946–1956. https://doi.org/10.1111/j.1365-2672.2010.04824.x.

(6) Haugland, R. a; Siefring, S. C.; Wymer, L. J.; Brenner, K. P.; Dufour, A. P. Comparison of Enterococcus Measurements in Freshwater at Two Recreational Beaches by Quantitative Polymerase Chain Reaction and Membrane Filter Culture Analysis. *Water Research* **2005**, *39* (4), 559–568. https://doi.org/10.1016/j.watres.2004.11.011.

(7) Haugland, R. a; Varma, M.; Sivaganesan, M.; Kelty, C.; Peed, L.; Shanks, O. C. Evaluation of Genetic Markers from the 16S RRNA Gene V2 Region for Use in Quantitative Detection of Selected Bacteroidales Species and Human Fecal Waste by QPCR. *Systematic and Applied Microbiology* **2010**, *33* (6), 348–357. https://doi.org/10.1016/j.syapm.2010.06.001.

(8) Kodani, M.; Winchell, J. M. Engineered Combined-Positive-Control Template for Real-Time Reverse Transcription-PCR in Multiple-Pathogen-Detection Assays. *Journal of Clinical Microbiology* **2012**, *50* (3), 1057–1060. https://doi.org/10.1128/JCM.05987-11.

(9) Liu, J.; Gratz, J.; Amour, C.; Kibiki, G.; Becker, S.; Janaki, L.; Verweij, J. J.; Taniuchi, M.; Sobuz, S. U.; Haque, R.; Haverstick, D. M.; Houpt, E. R. A Laboratory-Developed TaqMan Array Card for Simultaneous Detection of 19 Enteropathogens. *Journal of Clinical Microbiology* **2013**, *51* (2), 472–480. https://doi.org/10.1128/JCM.02658-12.

(10) Kildare, B. J.; Leutenegger, C. M.; McSwain, B. S.; Bambic, D. G.; Rajal, V. B.; Wuertz, S. 16S RRNA-Based Assays for Quantitative Detection of Universal, Human-, Cow-, and Dog-Specific Fecal Bacteroidales: A Bayesian Approach. *Water Research* **2007**, *41* (16), 3701–3715. https://doi.org/10.1016/j.watres.2007.06.037.

(11) Green, H. C.; Haugland, R. A.; Varma, M.; Millen, H. T.; Borchardt, M. A.; Field, K. G.; Walters, W. A.; Knight, R.; Sivaganesan, M.; Kelty, C. A.; Shanks, O. C. Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Applied and Environmental Microbiology* **2014**, *80* (10), 3086–3094. https://doi.org/10.1128/AEM.04137-13.

(12) Green, H. C.; Dick, L. K.; Gilpin, B.; Samadpour, M.; Field, K. G. Genetic Markers for Rapid PCR-Based Identification of Gull, Canada Goose, Duck, and Chicken Fecal Contamination in Water. *Applied and environmental microbiology* **2012**, *78* (2), 503–510. https://doi.org/10.1128/AEM.05734-11.

(13) Weidhaas, J. L.; Macbeth, T. W.; Olsen, R. L.; Sadowsky, M. J.; Norat, D.; Harwood, V. J. Identification of a Brevibacterium Marker Gene Specific to Poultry Litter and Development of a Quantitative PCR Assay. *J Appl Microbiol* **2010**, *109* (1), 334–347. https://doi.org/10.1111/j.1365-2672.2010.04666.x.

(14) Jothikumar, N.; Cromeans, T. L.; Vincent, R.; Lu, X.; Sobsey, M. D.; Erdman, D. D.; Hill, V. R. Quantitative Real-Time PCR Assays for Detection of Human Adenoviruses and Identification of Serotypes 40 and 41 Quantitative Real-Time PCR Assays for Detection of Human Adenoviruses and Identification of Serotypes 40 and 41. *Appl Environ Microbiol.* **2005**, *6* (71), 3131–3136. https://doi.org/10.1128/AEM.71.6.3131.

(15) Nutz, S.; Döll, K.; Karlovsky, P. Determination of the LOQ in Real-Time PCR by Receiver Operating Characteristic Curve Analysis: Application to QPCR Assays for Fusarium Verticillioides and F. Proliferatum. *Analytical and Bioanalytical Chemistry* **2011**, *401* (2), 717–726. https://doi.org/10.1007/s00216-011-5089-x.

(16) Fluss, R.; Faraggi, D.; Reiser, B. Estimation of the Youden Index and Its Associated Cutoff Point. *Biometrical Journal* **2005**, *47* (4), 458–472. https://doi.org/10.1002/bimj.200410135.

(17) Shrestha, A.; Dorevitch, S. Evaluation of Rapid QPCR Method for Quantification of E. Coli at Non-Point Source Impacted Lake Michigan Beaches. *Water Research* **2019**, *156*, 395–403. https://doi.org/10.1016/j.watres.2019.03.034.

(18) Johnston, C.; Byappanahalli, M. N.; Gibson, J. M.; Ufnar, J. a.; Whitman, R. L.; Stewart, J. R. Probabilistic Analysis Showing That a Combination of Bacteroides and Methanobrevibacter Source Tracking Markers Is Effective for Identifying Waters Contaminated by Human Fecal Pollution. *Environmental Science & Technology* **2013**, *47* (23), 13621–13628. https://doi.org/10.1021/es403753k.

(19) Messier, K. P.; Akita, Y.; Serre, M. L. Integrating Address Geocoding, Land Use Regression, and Spatiotemporal Geostatistical Estimation for Groundwater Tetrachloroethylene. *Environmental Science & Technology* **2012**, *46* (5), 2772–2780. https://doi.org/10.1021/es203152a.

(20) Stewart, J. R.; Boehm, A. B.; Dubinsky, E. A.; Fong, T.-T.; Goodwin, K. D.; Griffith, J. F.; Noble, R. T.; Shanks, O. C.; Vijayavel, K.; Weisberg, S. B. Recommendations Following a Multi-Laboratory Comparison of Microbial Source Tracking Methods. *Water Research* **2013**, *47* (18), 6829–6838. https://doi.org/10.1016/j.watres.2013.04.063.

(21) Harris, A. R.; Pickering, A. J.; Harris, M.; Doza, S.; Islam, Md. S.; Unicomb, L.; Luby, S.; Davis, J.; Boehm, A. B. Ruminants Contribute Fecal Contamination to the Urban Household Environment in Dhaka, Bangladesh. *Environmental Science & Technology* **2016**, *50* (9), 4642–4649. https://doi.org/10.1021/acs.est.5b06282.

(22) Ebentier, D. L.; Hanley, K. T.; Cao, Y.; Badgley, B. D.; Boehm, A. B.; Ervin, J. S.; Goodwin, K. D.; Gourmelon, M.; Griffith, J. F.; Holden, P. A.; Kelty, C. A.; Lozach, S.; McGee, C.; Peed, L. A.; Raith, M.; Ryu, H.; Sadowsky, M. J.; Scott, E. A.; Domingo, J. S.; Schriewer, A.; Sinigalliano, C. D.; Shanks, O. C.; Van De Werfhorst, L. C.; Wang, D.; Wuertz, S.; Jay, J. A. Evaluation of the Repeatability and Reproducibility of a Suite of QPCR-Based Microbial Source Tracking Methods. *Water research* **2013**, *47* (18), 6839–6848. https://doi.org/10.1016/j.watres.2013.01.060.

(23) Gelman, A.; Carpenter, B. Bayesian Analysis of Tests with Unknown Specificity and Sensitivity. *J. R. Stat. Soc. C* **2020**, rssc.12435. https://doi.org/10.1111/rssc.12435.

(24) Gelman, A.; Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*; Cambridge University Press, 2007. https://doi.org/10.2277/0521867061.

(25) McElreath, R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, 2nd ed.; CRC texts in statistical science; Taylor and Francis, CRC Press: Boca Raton, 2020.

(26) Layton, B. a; Cao, Y.; Ebentier, D. L.; Hanley, K.; Ballesté, E.; Brandão, J.; Byappanahalli, M.; Converse, R.; Farnleitner, A. H.; Gentry-Shields, J.; Gidley, M. L.; Gourmelon, M.; Lee, C. S.; Lee, J.; Lozach, S.; Madi, T.; Meijer, W. G.; Noble, R.; Peed, L.; Reischer, G. H.; Rodrigues, R.; Rose, J. B.; Schriewer, A.; Sinigalliano, C.; Srinivasan, S.; Stewart, J.; Van De Werfhorst, L. C.; Wang, D.; Whitman, R.; Wuertz, S.; Jay, J.; Holden, P. a; Boehm, A. B.; Shanks, O.; Griffith, J. F. Performance of Human Fecal Anaerobe-Associated PCR-Based Assays in a Multi-Laboratory Method Evaluation Study. *Water Research* **2013**, *47* (18), 6897–6908. https://doi.org/10.1016/j.watres.2013.05.060.

(27) Boehm, A. B.; Wang, D.; Ercumen, A.; Shea, M.; Harris, A. R.; Shanks, O. C.; Kelty, C.; Ahmed, A.; Mahmud, Z. H.; Arnold, B. F.; Chase, C.; Kullmann, C.; Colford, J. M.; Luby, S. P.; Pickering, A. J. Occurrence of Host-Associated Fecal Markers on Child Hands, Household Soil, and Drinking Water in Rural Bangladeshi Households. *Environmental Science & Technology Letters* **2016**, acs.estlett.6b00382. https://doi.org/10.1021/acs.estlett.6b00382.

(28) Odagiri, M.; Schriewer, A.; Hanley, K.; Wuertz, S.; Misra, P. R.; Panigrahi, P.; Jenkins, M. W. Validation of Bacteroidales Quantitative PCR Assays Targeting Human and Animal Fecal Contamination in the Public and Domestic Domains in India. *Science of The Total Environment* **2015**, *502*, 462–470. https://doi.org/10.1016/j.scitotenv.2014.09.040.

(29) Jenkins, M. W.; Tiwari, S.; Lorente, M.; Gichaba, C. M.; Wuertz, S. Identifying Human and Livestock Sources of Fecal Contamination in Kenya with Host-Specific Bacteroidales Assays. *Water research* **2009**, *43* (19), 4956–4966. https://doi.org/10.1016/j.watres.2009.07.028.

(30) Somnark, P.; Chyerochana, N.; Mongkolsuk, S.; Sirikanchana, K. Performance Evaluation of Bacteroidales Genetic Markers for Human and Animal Microbial Source Tracking in Tropical Agricultural Watersheds. *Environmental Pollution* **2018**, *236*, 100–110. https://doi.org/10.1016/j.envpol.2018.01.052.

(31) Symonds, E. M.; Young, S.; Verbyla, M. E.; McQuaig-Ulrich, S. M.; Ross, E.; Jiménez, J. A.; Harwood, V. J.; Breitbart, M. Microbial Source Tracking in Shellfish Harvesting Waters in the Gulf of Nicoya, Costa Rica. *Water Research* **2017**, *111*, 177–184. https://doi.org/10.1016/j.watres.2017.01.004.

(32) Nshimyimana, J. P.; Cruz, M. C.; Thompson, R. J.; Wuertz, S. Bacteroidales Markers for Microbial Source Tracking in Southeast Asia. *Water Research* **2017**, *118*, 239–248. https://doi.org/10.1016/j.watres.2017.04.027.

(33) Malla, B.; Ghaju Shrestha, R.; Tandukar, S.; Bhandari, D.; Inoue, D.; Sei, K.; Tanaka, Y.; Sherchand, J. B.; Haramoto, E. Validation of Host-specific Bacteroidales Quantitative PCR Assays and Their Application to Microbial Source Tracking of Drinking Water Sources in the Kathmandu Valley, Nepal. *Journal of Applied Microbiology* **2018**, *125* (2), 609–619. https://doi.org/10.1111/jam.13884.

(34) Li, X.; Sivaganesan, M.; Kelty, C. A.; Zimmer-Faust, A.; Clinton, P.; Reichman, J. R.; Johnson, Y.; Matthews, W.; Bailey, S.; Shanks, O. C. Large-Scale Implementation of Standardized Quantitative Real-Time PCR Fecal Source Identification Procedures in the Tillamook Bay Watershed. *PLOS ONE* **2019**, *14* (6), e0216827. https://doi.org/10.1371/journal.pone.0216827.

(35) Haramoto, E.; Osada, R. Assessment and Application of Host-Specific Bacteroidales Genetic Markers for Microbial Source Tracking of River Water in Japan. *PLOS ONE* **2018**, *13* (11), e0207727. https://doi.org/10.1371/journal.pone.0207727.

(36) Ahmed, W.; Gyawali, P.; Feng, S.; McLellan, S. L. Host Specificity and Sensitivity of Established and Novel Sewage-Associated Marker Genes in Human and Nonhuman Fecal Samples. *Applied and Environmental Microbiology* **2019**, *85* (14). https://doi.org/10.1128/AEM.00641-19.

(37) Feng, S.; McLellan, S. L. Highly Specific Sewage-Derived Bacteroides Quantitative PCR Assays Target Sewage-Polluted Waters. *Applied and Environmental Microbiology* **2019**, *85* (6), 1–15. https://doi.org/10.1128/AEM.02696-18.

(38) Schiaffino, F.; Pisanic, N.; Colston, J. M.; Rengifo, D.; Paredes Olortegui, M.; Shapiama, V.; Peñataro Yori, P.; Heaney, C. D.; Davis, M. F.; Kosek, M. N. Validation of Microbial Source Tracking Markers for the Attribution of Fecal Contamination in Indoor-Household Environments of the Peruvian Amazon. *Science of The Total Environment* **2020**, *743*, 140531. https://doi.org/10.1016/j.scitotenv.2020.140531.

(39) Bürkner, P.-C. Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **2017**, *80* (1). https://doi.org/10.18637/jss.v080.i01.

(40) Stan Development Team. Prior Choice Recommendations https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations (accessed 2020 -10 -14).

(41) Gelman, A.; Jakulin, A.; Pittau, M. G.; Su, Y.-S. A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models. *The Annals of Applied Statistics* **2008**, *2* (4), 1360–1383. https://doi.org/10.1214/08-AOAS191.

(42) Guo, J.; Riebler, A.; Rue, H. Bayesian Bivariate Meta-Analysis of Diagnostic Test Studies with Interpretable Priors. *Statist. Med.* **2017**, *36* (19), 3039–3058. https://doi.org/10.1002/sim.7313.