Title: Improving the effectiveness of water, sanitation, and hygiene interventions: a simulation approach to generalizing the outcomes of intervention trials

Authors: Andrew F. Brouwer PhD[1], Mondal H. Zahid PhD[1], Marisa C. Eisenberg PhD[1], Benjamin F. Arnold PhD[2,3], Sania Ashraf PhD[4], Jade Benjamin-Chung PhD[5,6], John M. Colford, Jr MD[7], Ayse Ercumen PhD[8], Stephen P. Luby MD[9], Amy J. Pickering PhD[10], Mahbubur Rahman MBBS[4], Joseph N.S. Eisenberg PhD[1,†], Matthew C. Freeman PhD[11,†]

[†]: these authors contributed equally

Affiliations

1. Department of Epidemiology, University of Michigan, Michigan, USA
2. Francis I. Proctor Foundation, University of California, San Francisco, San Francisco, CA, USA
3. Department of Ophthalmology, University of California, San Francisco, San Francisco, CA, USA
4. Environmental Interventions Unit, Infectious Disease Division, The International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh
5. Department of Epidemiology and Population Health, Stanford University, CA, USA
6. Chan Zuckerberg Biohub, San Francisco, CA, USA
7. School of Public Health, University of California Berkeley, Berkeley, CA, USA
8. Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, USA
9. Division of Infectious Diseases and Geographic Medicine, Stanford University, CA, USA
10. Department of Civil and Environmental Engineering, University of California Berkeley, Berkeley, CA, USA
11. Rollins School of Public Health, Emory University, Atlanta, GA, USA

Corresponding author: Andrew F Brouwer; Department of Epidemiology, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109; brouweaf@umich.edu; 734-764-7373

Declaration of conflicts of interest: The authors declare they have nothing to disclose.

Abstract

Background: While water, sanitation, and hygiene (WASH) interventions can reduce diarrhea diarrheal, recent large-scale trials have not found the expected health gains for young children in low-resource settings. Evidence-based guidance is needed to inform programs and future studies.

Objectives: We aimed to determine how sensitive intervention effectiveness is to underlying WASH factors (compliance, efficacy, etc) and to generalize the results of the WASH Benefits Bangladesh randomized controlled trial to other contexts or scenarios using a mechanistic disease transmission model and counterfactual simulations.

Methods: A disease transmission model was developed to account for transmission across multiple environmental pathways, multiple interventions (water (W), sanitation (S), hygiene (H), nutrition (N)) applied individually and in combination, adherence to interventions, and the impact of individuals not enrolled in the study. Leveraging a set of mechanistic parameter combinations fit to the WASH Benefits Bangladesh trial (n=17,187) using a Bayesian sampling approach, we simulated trial outcomes under counterfactual scenarios to estimate how changes in intervention completeness, coverage, fidelity and adherence, and efficacy, as well as baseline WASH conditions and disease burden, impacted intervention effectiveness.

Results: Increasing community coverage was associated with the greatest impact on intervention effectiveness (e.g., median increases in effectiveness of 30.4 and 39.9% points in the WSH and WSHN intervention arms when increasing coverage to 20%). The effect of community coverage on effectiveness depended on how much transmission was along pathways not modified by the interventions. Intervention effectiveness was reduced in counterfactual simulations with lower levels of preexisting WASH conditions or increased baseline disease burden. Individual interventions had complementary but not synergistic effects when combined.

Discussion: Next-generation WASH programs must address community coverage, transmission along pathways not traditionally covered by WASH, and that the effectiveness of individual-level WASH improvements will be blunted the further the community is from achieving herd protection.

Introduction

Enteric diseases, primarily spread through contact with fecally contaminated environments (water, surfaces, food, etc.), are one of the leading causes of morbidity and mortality in young children.[1] An estimated nearly 500,000 children under five years of age die from diarrheal disease globally each year[2,3], and it is hypothesized that repeated sub-clinical infections may lead to growth shortfalls.[4] Much of this burden is in low- and middle-income countries (LMICs)[5]. Enteric pathogens are transmitted by myriad pathways, including fluids, fomites, food, flies, and fauna, as summarized in the classic "F-diagram".[6] Studying and preventing diarrheal disease is complicated because a diverse array of pathogens can cause similar symptoms,[7,8] pathogens can exploit multiple transmission pathways, and asymptomatic infections can contribute to the community pathogen burden.[9]

Diarrheal disease is greatly reduced in communities with robust water, sanitation, and hygiene (WASH) infrastructure, with mutually reinforcing levels of community and individual protection. Household-level WASH improvements can result in considerable reductions in diarrheal disease burden in LMICs, and many WASH interventions—such as improved latrines and handwashing with soap—have demonstrable efficacy to reduce fecal exposure.[10] A recent meta-analysis by Wolf et al. of WASH intervention randomized controlled trials (RCTs) highlighted that WASH interventions reduce diarrhea in children in low-resource settings.[10] However, there is substantial heterogeneity in the effect estimates across the studies, and many of the more recent, large-scale trials reported less-than-expected or null results.[11–17] In particular, the results of the WASH Benefits (WASH-B) Bangladesh and Kenya trials and the SHINE trial, all of which found no impacts of WASH on linear growth but some mixed effects on diarrhea, were particular subjects of substantial discussion in the literature. [18–22] The sub-optimal performance of the interventions in these trials is likely due to a combination of multiple factors, including incomplete blocking of all transmission pathways (low intervenable fraction, also called completeness), inadequate community coverage of the intervention, or a lack of intervention compliance or efficacy.[19,23] Additionally, a community's baseline WASH conditions and disease burden can also impact the real-world intervention effectiveness.[19,23] Assessing which factors are the largest barriers to diarrheal disease reduction will aid policy-makers, practitioners, and researchers in deciding how best to invest in WASH programs and design the next generation of programs and trials.[18–24]

RCTs are considered the gold-standard for estimating causal relationships, and they are rigorous assessments of a particular intervention within a particular context at a particular point in time. But, their findings do not necessarily generalize to other contexts or conditions—e.g., different populations, disease burdens, pathogens, transmission pathways, intervention fidelity and adherence—when there are effect modifiers that vary across field settings and intervention implementations.[20,23] Mechanistic infectious

disease transmission models, unlike meta-analyses, have the potential to generalize findings by directly accounting for these location-specific contexts and conditions, exploring counterfactual questions through simulation of alternate scenarios, and developing location-specific programmatic targets. This approach is used extensively in other contexts to assess public health interventions or counterfactual conditions.[25–27] A mechanistic, counterfactual approach could lead to better-targeted public health WASH interventions, policy recommendations, and field trials.[28,29]

The aims of this work are to use a compartmental transmission model to determine how sensitive WASH intervention effectiveness is to the underlying WASH factors (compliance, efficacy, etc); to evaluate hypotheses about what led to the sub-optimal reductions in diarrhea among intervention households in an RCT (WASH-B Bangladesh); and to provide a framework to support planning of WASH interventions and context-specific WASH programming. We previously developed this model framework accounting for multiple environmental transmission pathways, shared environments, pre-existing WASH conditions, and adherence to multiple interventions and applied it to the empirical trial data.[30] Our approach generates thousands of combinations of coverage, intervention efficacy, and transmission pathway strengths that could reasonably underlie the trial results. In this analysis, in addition to directly exploring the sensitivity of the intervention effectiveness to the underlying WASH factors, we leveraged those parameter combinations to simulate how intervention effectiveness would have been different under alternate scenarios. These counterfactual simulations provide evidence for policy recommendations, programmatic targets, and an evaluation framework for next-generation WASH interventions.[20,21,24]

Methods

Summary of approach. In prior work, we developed a compartmental transmission model framework to explain the outcomes of a RCT.[30] Here, we first explore how effectiveness of a single intervention depends on six WASH factors:

- baseline WASH conditions (i.e., the fraction of the population has WASH infrastructure comparable to that provided by the intervention),
- compliance (i.e., the fraction of participants assigned to an intervention that are actually using it, accounting for both fidelity (whether the intervention was delivered) and adherence (whether participants used the intervention)),
- disease transmission potential (summarize by the basic reproduction number R_0),
- the intervenable fraction of transmission (i.e., how much of the transmission could be prevented in a perfect intervention),
- the intervention efficacy (i.e., how much transmission or shedding the intervention prevents), and

- the community coverage fraction (i.e., what fraction of the population is enrolled in the trial).

Then, we use a multi-intervention version of the model to investigate outcomes in the WASH-B Bangladesh trial specifically. In prior work, we demonstrated how to find mechanistic parameter sets that were consistent with individual-level diarrheal outcomes.[30] Here, we use these mechanistic parameter sets to simulate what intervention effectiveness would have been in each of the WASH-B Bangladesh trial arms under each of six counterfactuals corresponding to the six WASH factors above, accounting for uncertainty in the parameters underlying the real data (original scenario). By simulating what the intervention effectiveness would have been in the trial under alternative circumstances, we evaluated the extent to which each factor may have contributed to the observed outcomes.

Data. The WASH-B Bangladesh trial was a cluster-randomized trial of the efficacy of water, sanitation, hygiene, and nutrition interventions, alone and in combination, on diarrhea prevalence and linear growth.[14] The investigators measured (child-guardian-reported, past-seven-day, all-cause) diarrheal prevalence in children at three time points approximately one year apart. Households in the study area are typically organized into compounds in which a patrilineal family shares a common space and resources, such as a water source and latrine. A total of 5551 compounds were enrolled, contingent on having a pregnant woman in her second trimester during the enrollment period. The study followed one or more target children born after baseline, as well as any other children in the compound who were under age 3 at baseline. These compounds were grouped into 720 clusters. Each cluster was assigned to one of seven arms testing combinations of four interventions: water chlorination (W), a double-pit, pour-flush improved latrine (S), handwashing with soap and water (H), and supplementary nutrition sachets (N). The control arm (C) consisted of 180 compounds, while 90 were assigned to each of the water (W), sanitation (S), handwashing (H), nutrition (N), combined water, sanitation, and handwashing (WSH), and all interventions (WSH-N) arms. Specific details on trial design, interventions, and results are published elsewhere[14,31]. We assessed whether any individual was using an intervention or a substantively equivalent preexisting WASH condition through four indicators defined and assessed by the investigators: detection of free chlorine in water (W), latrine with a functional water seal (S), handwashing station with soap and water (H), at least 50% of nutrition sachets consumed (N). The W and H interventions were intended for the households of the target children, but we were not able to determine whether other children in the compound were in that household or not. For this analysis, we assumed that non-target children were covered by the interventions; any misspecification will attenuate the estimated efficacy of the W and H interventions[14]. We removed individuals with negative reported ages (n=2), missing reported diarrhea (n=2,745), or missing in any of the four use indicators (n=2,660), which left 17,187 individual observations (76% of the original sample) over the three surveys.

Ethics. This secondary analysis of deidentified data was not regulated as human-subjects research.

Model. *Single intervention model*. Our compartmental transmission model, denoted SISE-RCT, is a susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT.[30] The SISE-RCT model accounts for the six key mechanistic factors underlying WASH RCT outcomes described above. In the case of a single intervention, the population is partitioned into individuals with regular exposure (those not enrolled or included in the intervention and those not compliant, with compartments S_- and I_-) and those with exposure or shedding attenuated by the intervention (those compliant with the intervention or an equivalent preexisting condition, with compartments S_+ and I_+). The intervention and control arms are simulated separately, and both the regular and attenuated exposure populations are modeled in both simulations, accounting for the population not enrolled in the study (ω), baseline WASH conditions (ρ_0), and intervention compliance (ρ). The SISE-RCT parameters are given in Table 1, and a model diagram is given in Figure 1. The control arm is simulated with the same parameters as the intervention arms, except that $\rho = \rho_0$. As we previously demonstrated,[30] using a quasi-steady state assumption (i.e., that the environmental compartments are at steady-state), the steady state values for compartments may be found by simulating the following equations to steady state with population fractions $N_+ = \omega\rho + (1-\omega)\rho_0$ and $N_- = \omega(1-\rho) + (1-\omega)(1-\rho_0)$ for the attenuated and regular exposure populations respectively. For brevity, we omit the $\frac{dS}{dt}$ equations, each of which is given by $\frac{dS}{dt} = -\frac{dI}{dt}$ for the corresponding subpopulation (regular or attenuated exposure).

$$\frac{dI_+}{dt} = (\varphi_\beta R_{0,1}E_1 + R_{0,2}E_2)S_+ - I_+,$$

$$\frac{dI_-}{dt} = (R_{0,1}E_1 + R_{0,2}E_2)S_- - I_-,$$

$$E_1 = \varphi_\alpha I_+ + I_-,$$

$$E_2 = I_+ + I_-.$$

The prevalence of disease in the population is denoted $\pi = I_+ + I_-$. The prevalence in the intervention arm π is compared to the prevalence π_C in the control arm, and the intervention effectiveness is defined as $\varepsilon = (\pi_C - \pi)/\pi_C$, namely the fractional reduction in prevalence in the intervention arm relative to the control arm. We simulated this model with the original scenario set of parameters listed in Table 1. We varied each factor one at a time across the range of values given in Table 1, calculating the value needed to achieve disease elimination (defined as disease prevalence <0.1%). We also varied each pair of factors together to investigate potential interactions between factors. Only simulations with $\rho > \rho_0$ were included

to avoid simulation of perverse situations where the intervention reduced use of WASH. This model has been made publicly available as a web app at https://umich-biostatistics.shinyapps.io/sise_rct/.

WASH-B Bangladesh (multi-intervention) model. As discussed in the Data section above, the WASH-B Bangladesh trial included 720 clusters of households each assigned to one of seven arms (control, W, S, H, N, WSH, WSHN). We extended the single-intervention SISE-RCT model to a multi-intervention model by accounting for transmission across three environmental pathways (water, fomites & hands, and all others combined), four interventions applied individually (W, S, H, N) and in combination (WSH, WSHN), and individual-level compliance with interventions or preexisting conditions. In brief, we modeled each of the 720 clusters with susceptible and infectious compartments for each of 2^4=16 combinations of interventions/conditions depending on household adherence, i.e., in every cluster, we modeled the infection prevalence for each combination of having or not having each intervention or equivalent preexisting condition. For example, for a cluster in the WSH arm, we estimated how many people were not using any interventions, how many were using W only, how many were using S only, etc., and what the infection prevalence was among each group given their collective interaction through the shared environments. The 18 model parameters were i) the overall basic reproduction number R_0, which defines the underlying disease pressure measured in the control arm at baseline, ii) two parameters partitioning R_0 into the strengths of the drinking water $R_{0,w}$, fomite & hands $R_{0,f}$, and all other transmission pathways $R_{0,o}$, iii) eight relative reproduction numbers accounting for systematic differences in disease pressure over survey (baseline, midline, and endline) and across arms independently, iv) the community coverage ω, and v) efficacy parameters defining the effect of each intervention (four) or preexisting condition (two) on the transmission pathways (the W intervention (chlorination) reduces transmission via the water pathway $\varphi_{\beta_w,W}$; the S intervention (latrine with water seal) reduces shedding into the shared water environment with different efficacy for preexisting conditions $\bar{\varphi}_{\alpha_w,S}$ and the trial intervention $\varphi_{\alpha_w,S}$; the H (handwashing with soap and water) reduces transmission via the fomite pathway with different efficacy for preexisting conditions $\bar{\varphi}_{\beta_f,H}$ and trial intervention $\varphi_{\beta_f,H}$; and the N intervention (nutrition supplementation) reduces susceptibility to all transmission $\varphi_{\beta,N}$). As above, the 35 original equations (16 each for susceptible and infectious individuals and 3 environmental pathogen concentrations) were simplified using a quasi-steady state assumption on the environmental compartments. Again, for brevity, we omit the $\frac{dS}{dt}$ equations, each of which is given by $\frac{dS}{dt} = -\frac{dI}{dt}$ for the corresponding subpopulation. The basic reproduction number parameters are adjusted by the time and arm-specific and relative basic reproduction numbers corresponding to the cluster being modeled. The

intervention efficacy parameters $\varphi_{\alpha_w,S}$ and $\varphi_{\beta_f,H}$ are replaced by $\bar{\varphi}_{\alpha_w,S}$ and $\bar{\varphi}_{\beta_f,H}$ in clusters without the S and H interventions, respectively, and at baseline.

$$\frac{dI}{dt} = (R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S - I,$$

$$\frac{dI_W}{dt} = (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_W - I_W,$$

$$\frac{dI_S}{dt} = (R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_S - I_S,$$

$$\frac{dI_H}{dt} = (R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_H - I_H,$$

$$\frac{dI_N}{dt} = \varphi_{\beta,N} \cdot (R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_H - I_H,$$

$$\frac{dI_{WS}}{dt} = (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_{WS} - I_{WS},$$

$$\frac{dI_{WH}}{dt} = (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{WH} - I_{WH},$$

$$\frac{dI_{WN}}{dt} = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_{WN} - I_{WN},$$

$$\frac{dI_{SH}}{dt} = (R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{SH} - I_{SH},$$

$$\frac{dI_{SN}}{dt} = \varphi_{\beta,N} \cdot (R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_{SN} - I_{SN},$$

$$\frac{dI_{HN}}{dt} = \varphi_{\beta,N} \cdot (R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{HN} - I_{HN},$$

$$\frac{dI_{WSH}}{dt} = (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{WSH} - I_{WSH},$$

$$\frac{dI_{WSN}}{dt} = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + R_{0,f}E_f + R_{0,o}E_o)S_{WSN} - I_{WSN},$$

$$\frac{dI_{WHN}}{dt} = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{WHN} - I_{WHN},$$

$$\frac{dI_{SHN}}{dt} = \varphi_{\beta,N} \cdot (R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{SHN} - I_{SHN},$$

$$\frac{dI_{WSHN}}{dt} = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w}E_w + \varphi_{\beta_f,H} \cdot R_{0,f}E_f + R_{0,o}E_o)S_{WSH} - I_{WSH},$$

$$E_w = \sum_{S \text{ not in } i} I_i + \varphi_{\alpha,S} \sum_{S \text{ in } i} I_i$$

$$E_f = E_o = \sum_i I_i$$

When solving these differential equations for a given cluster, we use the distribution of interventions and preexisting conditions recorded in the data for those participants and assume that participants not in the study have the same distribution of preexisting conditions as the control arm participants. The initial prevalence is set to 6% for all subgroups, and the model is run to 100 days to reach steady state.

To fit the model to the trial data, we employed a hybrid sampling-importance resampling and estimation framework to obtain 25,000 parameter combinations that represented a good fit to the diarrheal outcomes of each participant using a Bernoulli statistical likelihood.[32] We resampled, with replacement, from our initial 25,000 parameter combinations, based on their goodness of fit; 2,029 unique parameter combinations were included in the final sample, with varying frequency. These parameters sets are similar to but not exactly the same as in Brouwer et al,[30] as they include a small code correction. The fit to the data is given in Figure S1, and the distributions of parameters are given in Figures S2-5.

WASH-B Bangladesh counterfactual analysis. For each of the 25,000 parameter sets *j* identified by fitting the model to WASH-B Bangladesh, we defined the corresponding original scenario matching the WASH-B Bangladesh trial outcomes and the corresponding intervention effectiveness ε_0^j. We considered six types of counterfactual scenarios, detailed in Table 2, addressing baseline WASH conditions, baseline disease burden, the intervenable fraction of transmission, intervention compliance, intervention efficacy, and community coverage. Any parameter sets that eliminated disease in the control arm in a counterfactual simulation were censored from the results as they did not provide information on intervention effectiveness.

The main outcome of a counterfactual simulation was the (absolute) change intervention effectiveness compared to the original scenario, namely $\varepsilon_*^j - \varepsilon_0^j$, where ε_*^j is the intervention effectiveness in the given counterfactual scenario for the *j*th parameter set. We used absolute change rather than percentage change because absolute change, unlike percent change, is bounded (between -100 and 100% points). The intervention is more effective (i.e., a greater reduction in diarrheal prevalence in the intervention arm compared to the control arm) in the counterfactual scenario than the original scenario when the change is positive. To assess whether the intervention factors modified the intervention effectiveness in the counterfactual scenarios, for a subset of counterfactuals, we assessed how the effect of the counterfactual depended on quantiles of the other parameter values.

Results

Dynamics and sensitivity of the single-intervention model

The intervention effectiveness ε at the original scenario parameters in Table 1 was 50%, with a prevalence of 6.4% in the control arm and 3.2% in the intervention arm. Disease elimination would have been achieved in this hypothetical intervention if 1) 36% rather than 25% of the population already had comparable WASH infrastructure; 2) the baseline disease prevalence was 1.1% (R_0=1.17) rather than 6.4% (R_0=1.25); 3) 98% rather than 75% of transmission was along pathways impacted by the intervention; 4) the efficacy of the intervention was 98% rather than 75%; or 5) coverage was increased from 11% to 30%. Disease could not be eliminated by increasing intervention compliance from 75%, even to 100%.

The intervention effectiveness as a function of each pair of the six parameters is given in Figure 2. For most pairs of parameters, there was little evidence of an interaction between the factors (i.e., the contours of the heatmaps are approximately linear and parallel, except at extreme values). The exception to this pattern was coverage. In the inset in Figure 2, we show, as an illustration, the interaction between coverage and compliance on the intervention effectiveness. When coverage is low and compliance is high (point A), it is easier to increase intervention effectiveness by increasing coverage (gray arrow, moving along the x-axis), but when coverage is higher and compliance is low (point B), then it is easier to increase intervention effectiveness by increasing compliance (black arrow, moving along the y-axis). "Easier" here does not reflect cost or feasibility but only the result of a unit change for each individual parameter. Cost-effectiveness is outside of the scope of this work but could be explored in future analysis. Similarly, the coverage needed to achieve disease elimination depended non-linearly on each of the other factors. This observation informed the effect modification analyses included in the WASH-B Bangladesh counterfactual analysis below.

The reader can explore the sensitivity of the model for other values of the WASH factors on the web app available via https://umich-biostatistics.shinyapps.io/sise_rct/.

WASH-B Bangladesh counterfactual analysis

The median baseline disease prevalence in the original scenario was 7.1% (range 5.9–8.2%), decreasing to 5.7% (range 5.2–6.3%) for midline/endline (Figure S1). The median intervention effectivenesses were 8.2% in the W arm, 36.4% in the S arm, 33.1% in the H arm, 30.3% in the WSH arm, 34.0% in N arm, and 34.6% in the WSHN arm (Table 3). The percentage point change varied across arms in each counterfactual scenario. Figure 3 shows the distribution of percentage point change in intervention

effectiveness over the 25,000 parameter sets for each arm and counterfactual scenario, and Table 3 gives the median values.

Eliminate preexisting WASH Conditions. We found that implementing the interventions in a community with no handwashing stations with soap and water or latrines with water seals would have likely resulted in less effective interventions compared to the actual community's higher baseline WASH conditions (e.g., 8.3% points less in the WSH arm; Figure 3a). The W arm is the exception because it had lower effectiveness in the original scenario. The uncertainty in change in intervention effectiveness in each arm was largely driven by uncertainty in what the baseline disease prevalence would have been in the counterfactual scenario (median 8.9%, range 6.4–23.1%).

Double baseline disease prevalence. A higher transmission potential corresponding to doubling the baseline diarrheal disease prevalence (median 14.2% vs median 7.1%) would also have resulted in less-effective interventions compared to the true baseline diarrheal disease prevalence (e.g., 11.9% points less in the WSH arm; Figure 3 b). As above, the W arm is the exception because it had lower effectiveness in the original scenario.

Full compliance. The impact of increasing intervention adherence was negligible-to-modest (e.g., 4.4% points more in the WSH arm; Figure 3c). Note that intervention compliance, as defined by the investigators, was already high.[14,33]

Half of the other pathway transmission can be intervened on. We found that intervention effectiveness could have been greater if more of the total disease transmission was via the water and fomite pathways rather than through pathways that were not intervened on (e.g., 20.3% points more in the WSH arm; Figure 3d). There was potential for a substantial increase in intervention effectiveness, as indicated by the distribution of the individual simulation outcomes, but the median impact was modest, with a less than 25% points increase in effectiveness in the multi-intervention arms and a less than 10% points increase in the single-intervention arms (Table 3). The uncertainty in the potential impact was largely driven by uncertainty in how much of the disease transmission was through other pathways in the original scenario.

Double intervention efficacy. We assessed the impact of increasing efficacy—defined as increasing the reduction of transmission along the relevant pathway(s)—of the four interventions. We found that in each of these increased efficacy scenarios, substantial increases in intervention efficacy could have improved intervention effectiveness in the corresponding arms (Figure 3e–h), with median improvements between 5% and 20% points.

Increase community coverage. The median estimated community coverage in the trial was 6.9%, but this estimate was highly uncertain, ranging from nearly 0% to 20% (Figure S5). For our main coverage counterfactual, we increased the community coverage in each simulation to 20.0%, chosen as a substantial but not unreasonable increase in coverage.[34] This counterfactual scenario was associated with the greatest median increase in intervention effectiveness (among all households now covered by the intervention) of any of the considered counterfactual scenarios (e.g., 30.4% and 39.9% points more in the WSH and WSHN arms; Figure 4a). Following the results of the single-intervention model that highlighted that the effect of coverage depended on the other WASH factors, we plotted the intervention effectiveness distributions for this community coverage counterfactual by quintiles of the values of the other WASH factors. Effect modification is present if the effect of increased coverage depends on the quintile of the WASH factor. Note that when looking at quintiles of one factor, the values of the other factors may not be evenly distributed across the quintiles if values of the factors are correlated. We found that the increase intervention effectiveness with increased community coverage in the W, S, WSH, and WSHN intervention arms depended partly on the strength of transmission via the water pathway (Figure 4b). The increases in intervention effectiveness in these arms could only reach their full potential if the strength of the water pathway were high. A similar, but more modest effect was seen for the H arm and the strength of the fomite pathway (Figure 4c). The greatest overall effect modifier of the impact of increased coverage on intervention effectiveness is the strength of the other pathways (i.e., the intervenable fraction, Figure 4d). When the strength of other pathways was high, increasing coverage had less of an impact. Intervention efficacy also modified the impact of increased coverage but only in the intervention arms with those interventions (Figure 4e–h).

To further understand the joint impact of community coverage and the intervenable fraction (i.e., the strength of the other transmission pathway), we simulated the intervention effectiveness as a function of increased coverage for the highest and lowest quartiles of intervention completeness (Figure 5). The impact of increased coverage on intervention effectiveness depended on the intervenable fraction most strongly for the W, WSH, and WSHN arms, moderately for the S arm, and little for the H and N arms. For example, in the W arm, increasing community coverage to 50% resulted in a median increase of 69% points for samples with the highest intervenable fractions but only 18% points for samples with the lowest intervenable fractions. In contrast, in the H arm, increasing community coverage to 50% resulted in a median increase of 26% points for samples with the highest intervenable fractions compared only 19% points for samples with the lowest intervenable fractions. (Note that the fact that intervenable fraction was relevant for the N arm at all is a result of the correlations between the intervenable fraction and the other parameters in the original parameter sets.)

Discussion

Our model-based analysis used counterfactual simulations to generalize the results of a WASH intervention trial and develop guidance for policymakers and researchers. Our first finding was that increasing community coverage led to the most substantial reduction in disease among people receiving interventions. Second, we found that intervention completeness (i.e., the fraction of disease transmission along pathways that were intervened on) was an important effect modifier of the impact of community coverage on intervention effectiveness, with the impact of increased community coverage enhanced when interventions covered a larger fraction of transmission. Third, our work suggests that interventions are likely to be more effective when disease burden is low, a finding that runs counter to a previous hypothesis about the trial.[14] Finally, we found that multifaceted WASH interventions (WSH) added value over single component interventions (W, S, or H). Each of these findings suggest a path forward for policy and program recommendations for WASH investments and demonstrates how transmission models can be used to design the next generation of WASH interventions and set location-specific programmatic targets.

The importance of ensuring high community WASH coverage to address health outcomes has been highlighted in multiple context, including latrines,[29,35] bed nets,[36] and chemotherapy for helminths,[37] among others. Further work is needed to improve our measures of indirect and direct intervention effects[29,38] to better determine sanitation targets. Our findings support the call for systems-level WASH provisioning and improved universal access, underscoring the fundamental push to achieve the 2030 sustainable development targets.[39] Additionally, because our results suggest that the fraction of transmission an intervention can impact modifies the effect of increasing community coverage, we emphasize the need to better understand the sources not impacted by traditional WASH interventions. For example, contamination of food outside the home or from flies or exposure to feces from animals living near or inside the home may not be reduced by water quality or latrine interventions.[40–42] Capturing and reducing transmission through additional targeted interventions would increase the fraction of transmission intervened on and thereby make increased community coverage even more effective.

Low diarrheal prevalence makes it more difficult to observe a statistically significant reduction in diarrhea.[14] However, from a mechanistic perspective, we found that intervention effectiveness would have been lower had the background disease pressure in the community been higher because individual-level interventions can be overwhelmed by higher disease pressure from the community. This finding is consistent with previous literature that has shown that non-pharmaceutical interventions are more effective for less transmissible pathogens or when the population has a higher degree of population immunity.[43] This is not to say that individual improvements would have no effect but that the effects are

blunted if disease pressure in the rest of the community were not also addressed. These results are supported by the outcomes of WASH-B Kenya trial, which had higher disease prevalence (27% in the control arm) and no significant intervention effects on diarrheal prevelance.[15] Our results highlight limitations of standard statistical methods used to estimate trial sample sizes and point to a potential role for transmission models in study design.[38]

Similarly, it is often suggested that when WASH conditions (e.g., improved latrines) are relatively high, interventions (e.g., latrine water seal) do not provide a substantial improvement.[18,19,22,44] Counter to these hypothesis, we found that intervention effectiveness would have been lower if the baseline WASH conditions were lower: if the baseline WASH conditions were poorer, the community disease burden would be higher, so, as above, it would be more difficult to protect study participants from infection.

Because enteric pathogens can exploit multiple transmission pathways, many studies have tried to determine whether combined WASH interventions (WSH) are more effective than single interventions (W, S, or H).[31,45] Whether or not there is an additional effect of combined interventions depends on whether the interventions are complementary, that is, whether they each block some of the transmission that the other interventions would not have blocked.[46] This complementarity is an assumption in our transmission model framework (as each intervention affects different parts of the disease system), and we found that complementarity is consistent with the observed trial results.[32] Other modeling and empirical studies, support that WASH interventions can complement each other, or even potentially be synergistic.[47,48] In this work, we found that the combined interventions could have a greater effect than the individual interventions, but that the effects were generally sub-additive, meaning that the effectiveness of the combined WSH intervention was less than the sum of its parts (Table 3). Combined interventions offer a substantially better chance of disease elimination, especially at higher coverage levels (Figure 5).

One challenge that WASH RCTs often face is achieving high compliance through both high fidelity (providing the interventions as planned) and high adherence of participants to the use of the intervention. In WASH-B Bangladesh, the intervention compliance, as defined by trial investigators, was high, generally above 90%.[14,33] Accordingly, our full compliance counterfactual was limited in the impact it could detect. In the single-intervention model, however, compliance had the weakest effect on intervention effectiveness and was the only factor unable to eliminate disease on its own (Table 1, Figure 2).

The strength of our approach is underscored by the rich and high-quality data collected by the WASH-B Bangladesh trial (and other RCTs) and in our transmission model framework capturing relative disease prevalence. RCTs provide the gold standard of evidence about intervention effectiveness in a specific

context, and our approach allows us to generalize RCT results to other contexts, providing a tool for powerful policy and programmatic guidance. The SISE-RCT model can be customized for local contexts and interventions and then used to support local decision-making (e.g., to determine whether to invest in community coverage vs intervention efficacy). Future work may also develop recommendations for achieving elimination while minimizing costs. One limitation of our study is the high uncertainty in many of the model parameters, especially the intervenable fraction, which propagates into the counterfactual scenarios. These uncertainties stem from potential trade-offs in the model, e.g., a low intervenable fraction and a low intervention efficacy may have similar effects. Fortunately, our framework has the potential to incorporate additional information about parameters like the intervenable fraction and efficacy through our Bayesian sampling-importance resampling approach, allowing us to tailor projections of intervention effectiveness to specific parameter regions based on additional information (e.g., chlorination efficacy above 75%). One limitation of the data was the inability to distinguish whether non-target children were members of the same household as the target child or not, which introduced misspecification into our classification of W and H exposures, likely attenuating the efficacy estimates for those interventions. Also, we accounted for changes in disease pressure between, but not within, survey periods; future work may more directly address seasonal changes in disease pressure and even pathway strength, as a function of precipitation, seasonal flooding, etc. Another limitation of this study is that our results do not directly address some aspects of the Sustainable Development Goal (SDG) Target 6.2.[39] For example, the sanitation arm did not move households from no or basic sanitation to improved sanitation (as defined by the Joint Monitoring Programme). So, the "sanitation" intervention outcomes we estimated may not directly correspond to the policy-relevant changes required to meet SDG target. Likewise, the "water" intervention focused on water quality improvements (chlorination) but not water quantity. None of these issues are limitations of our modeling framework; rather, they are limitations of our specific application. Applying our methods across other trial datasets could address these limitations by allowing for modeling of other—and perhaps more policy-relevant—WASH exposure parameters.

Our work contributes to the robust discussion[18–24,49] about the future directions of WASH research and programming, and our modeling approach is well-suited to reevaluating current evidence during the "pause for reflection" recommended by a consensus of WASH reserachers.[19] This consensus group said that "the lesson perhaps lies in not seeking to attribute benefits to individual WASH factors but in that the public health dividends are paid when comprehensive services are in place." Our work underscores this conclusion, not only by emphasizing the importance of coverage and completeness of interventions, but also in its rejection of the hypotheses that greater effectiveness might be found in areas with greater disease prevalence or lower preexisting WASH infrastructure. Indeed, our findings suggest that the effect of individual-level WASH improvements will be blunted the further the community is from achieving

herd protection. Accordingly, this analysis provides further evidence supporting community-level interventions seeking to achieve herd protection through high community coverage.

Contributors

JNSE, MCE, MCF, and AFB conceived of the study. JNSE and MCF secured funding for the study. AFB, MCE, and JNSE developed the model. AFB wrote and implemented the software code, completed formal analysis and visualization, and curated the data and code. MHZ validated the software code. AFB wrote the original draft with input from JNSE and MCF. BFA, SA, JBC, JMC, AE, SPL, AJP, and MR contributed equally by aiding in interpretation of the results and providing their expertise in the WASH Benefits trials. All authors reviewed and edited the manuscript. All authors had full access to all study data.

Acknowledgements

This work was funded by the Bill & Melinda Gates Foundation (grant INV-005081) and the National Science Foundation (grant DMS-1853032). The original WASH-B Bangladesh trial was also funded by the Bill & Melinda Gates Foundation (grant OPPGD759). The study sponsors had no role in the study design, the analysis or interpretation of the results, the writing of the report, or in the decision to submit the paper for publication.

Data sharing

The WASH Benefits Bangladesh data is publicly available at https://osf.io/tprw2/. The data and code underlying the results of this paper are available at https://doi.org/10.5281/zenodo.8057236. The SISE-RCT web app with the single-intervention model is available at https://umich-biostatistics.shinyapps.io/sise_rct/. Data underlying each of the figures is given in the Excel spreadsheet supplement.

References

1. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9

2. Troeger C, Blacker B, Khalil IA, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis*. 2018;18(11):1191-1210. doi:10.1016/S1473-3099(18)30310-4

3. Wolf J, Johnston RB, Ambelu A, et al. Burden of disease attributable to unsafe drinking water, sanitation, and hygiene in domestic settings: a global analysis for selected adverse health outcomes. *Lancet*. 2023;401(10393):2060-2071. doi:10.1016/S0140-6736(23)00458-0

4. Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet*. 2009;374(9694):1032-1035. doi:10.1016/S0140-6736(09)60950-8

5. Julian TR. Environmental transmission of diarrheal pathogens in low and middle income countries. *Environ Sci Process Impacts*. 2016;18(8):944-955. doi:10.1039/C6EM00222F

6. Wagner EG, Lanoix JN. Excreta disposal for rural areas and small communities. *Monogr Ser World Health Organ*. 1958;39:1-182. http://www.ncbi.nlm.nih.gov/pubmed/13581743.

7. Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *Lancet*. 2013;382(9888):209-222. doi:10.1016/S0140-6736(13)60844-2

8. Liu J, Platts-Mills JA, Juma J, et al. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. *Lancet*. 2016;388(10051):1291-1301. doi:10.1016/S0140-6736(16)31529-X

9. Watson SI, Rego RTT, Hofer T, Lilford RJ. Evaluations of water, sanitation and hygiene interventions should not use diarrhoea as (primary) outcome. *BMJ Glob Heal*. 2022;7(5):e008521. doi:10.1136/bmjgh-2022-008521

10. Wolf J, Hubbard S, Brauer M, et al. Effectiveness of interventions to improve drinking water, sanitation, and handwashing with soap on risk of diarrhoeal disease in children in low-income and middle-income settings: a systematic review and meta-analysis. *Lancet*. 2022;400(10345):48-59.

doi:10.1016/S0140-6736(22)00937-0

11. Clasen T, Boisson S, Routray P, et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: A cluster-randomised trial. *Lancet Glob Heal*. 2014;2(11):e645-e653. doi:10.1016/S2214-109X(14)70307-9

12. Patil SR, Arnold BF, Salvatore AL, et al. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled trial. *PLoS Med*. 2015;11(8). doi:10.1371/journal.pmed.1001709

13. Pickering AJ, Djebbari H, Lopez C, Coulibaly M, Alzua ML. Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: A cluster-randomised controlled trial. *Lancet Glob Heal*. 2015;3(11):e701-e711. doi:10.1016/S2214-109X(15)00144-8

14. Luby SP, Rahman M, Arnold BF, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised controlled trial. *Lancet Glob Heal*. 2018;6(3):e302-e315. doi:10.1016/S2214-109X(17)30490-4

15. Null C, Stewart CP, Pickering AJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *Lancet Glob Heal*. 2018;6(3):e316-e329. doi:10.1016/S2214-109X(18)30005-6

16. Rogawski McQuade ET, Platts-Mills JA, Gratz J, et al. Impact of Water Quality, Sanitation, Handwashing, and Nutritional Interventions on Enteric Infections in Rural Zimbabwe: The Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial. *J Infect Dis*. 2020;221(8):1379-1386. doi:10.1093/infdis/jiz179

17. Knee J, Sumner T, Adriano Z, et al. Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-after trial. *Elife*. 2021;10. doi:10.7554/eLife.62278

18. Cumming O, Curtis V. Implications of WASH Benefits trials for water and sanitation. *Lancet Glob Heal*. 2018;6(6):e613-e614. doi:10.1016/S2214-109X(18)30192-X

19. Cumming O, Arnold BF, Ban R, et al. 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(1):1-9. doi:10.1186/s12916-019-1410-x

20. Levy K, Eisenberg JNS. Moving towards transformational WASH. *Lancet Glob Heal*. 2019;7(11):e1492. doi:10.1016/S2214-109X(19)30396-1

21. Pickering AJ, Null C, Winch PJ, et al. The WASH Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea. *Lancet Glob Heal*. 2019;7(8):e1139-e1146. doi:10.1016/S2214-109X(19)30268-2

22. Coffey D, Spears D. Implications of WASH Benefits trials for water and sanitation. *Lancet Glob Heal*. 2018;6(6):e615. doi:10.1016/S2214-109X(18)30225-0

23. Haque SS, Freeman MC. The Applications of Implementation Science in Water, Sanitation, and Hygiene (WASH) Research and Practice. *Environ Health Perspect*. 2021;129(6):65002. doi:10.1289/EHP7762

24. Amebelu A, Ban R, Bhagwan J, et al. The Lancet Commission on water, sanitation and hygiene, and health. *Lancet*. 2021;398(10310):1469-1470. doi:10.1016/S0140-6736(21)02005-5

25. Li X, Mukandavire C, Cucunubá ZM, et al. Estimating the health impact of vaccination against ten pathogens in 98 low-income and middle-income countries from 2000 to 2030: a modelling study. *Lancet*. 2021;397(10272):398-408. doi:10.1016/S0140-6736(20)32657-X

26. World Health Organization (WHO). *Impact of the COVID-19 Pandemic on Seven Neglected Tropical Diseases: A Model-Based Analysis*. Geneva; 2021. https://apps.who.int/iris/handle/10665/343993.

27. Lofgren E, Halloran ME, Rivers CM, et al. Mathematical models: A key tool for outbreak response. *Proc Natl Acad Sci*. 2015;112(2):E234-E234. doi:10.1073/pnas.1423846112

28. Coffeng LE, Vaz Nery S, Gray DJ, Bakker R, de Vlas SJ, Clements ACAA. Predicted short and long-term impact of deworming and water, hygiene, and sanitation on transmission of soil-transmitted helminths. Walker M, ed. *PLoS Negl Trop Dis*. 2018;12(12):1-17. doi:10.1371/journal.pntd.0006758

29. Fuller JA, Eisenberg JNS. Herd protection from drinking water, sanitation, and hygiene interventions. *Am J Trop Med Hyg*. 2016;95(5):1201-1210. doi:10.4269/ajtmh.15-0677

30. Brouwer AF, Eisenberg MC, Bakker KM, et al. Leveraging infectious disease models to interpret randomized controlled trials: Controlling enteric pathogen transmission through water, sanitation, and hygiene interventions. Lau EH, ed. *PLOS Comput Biol*. 2022;18(12):e1010748. doi:10.1371/journal.pcbi.1010748

31. Arnold BF, Null C, Luby SP, et al. Cluster-randomised controlled trials of individual and combined water, sanitation, hygiene and nutritional interventions in rural bangladesh and Kenya:

The WASH benefits study design and rationale. *BMJ Open*. 2013;3(8):1-17. doi:10.1136/bmjopen-2013-003476

32. Brouwer AF, Eisenberg MC, Bakker KM, et al. Leveraging infectious disease models to interpret randomized controlled trials: controlling enteric pathogen transmission through water, sanitation, and hygiene interventions. *medRxiv*. January 2022:2022.04.28.22274441. doi:10.1101/2022.04.28.22274441

33. Ercumen A, Pickering AJ, Kwong LH, et al. 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(21):12089-12097. doi:10.1021/acs.est.8b02988

34. Garn J V., Sclar GD, Freeman MC, et al. The impact of sanitation interventions on latrine coverage and latrine use: A systematic review and meta-analysis. *Int J Hyg Environ Health*. 2017;220(2):329-340. doi:10.1016/j.ijheh.2016.10.001

35. Harris M, Alzua ML, 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(12):7219-7227. doi:10.1021/acs.est.7b00178

36. Teklehaimanot A, Sachs JD, Curtis C. Malaria control needs mass distribution of insecticidal bednets. *Lancet*. 2007;369(9580):2143-2146. doi:10.1016/S0140-6736(07)60951-9

37. Anderson R, Truscott J, Hollingsworth TD. The coverage and frequency of mass drug administration required to eliminate persistent transmission of soil-transmitted helminths. *Philos Trans R Soc B*. 2014;369(1645):20130435. doi:10.1098/rstb.2013.0435

38. Halloran ME, Auranen K, Baird S, et al. Simulations for designing and interpreting intervention trials in infectious diseases. *BMC Med*. 2017;15(1):1-8. doi:10.1186/s12916-017-0985-3

39. World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). *Progress on Household Drinking Water, Sanitation and Hygiene 2000-2020: Five Years into the SDGs*. Geneva; 2021.

40. Penakalapati G, Swarthout J, Delahoy MJ, et al. Exposure to Animal Feces and Human Health: A Systematic Review and Proposed Research Priorities. *Environ Sci Technol*. 2017;51(20):11537-11552. doi:10.1021/acs.est.7b02811

41. Khamesipour F, Lankarani KB, Honarvar B, Kwenti TE. A systematic review of human pathogens carried by the housefly (Musca domestica L.). *BMC Public Health*. 2018;18(1):1049. doi:10.1186/s12889-018-5934-3

42. DuPont HL. The Growing Threat of Foodborne Bacterial Enteropathogens of Animal Origin. *Clin Infect Dis*. 2007;45(10):1353-1361. doi:10.1086/522662

43. Baker RE, Saad-Roy CM, Park SW, Farrar J, Metcalf CJE, Grenfell BT. Long-term benefits of nonpharmaceutical interventions for endemic infections are shaped by respiratory pathogen dynamics. *Proc Natl Acad Sci*. 2022;119(49). doi:10.1073/pnas.2208895119

44. Arnold BF, Null C, Luby SP, Colford JM. Implications of WASH Benefits trials for water and sanitation – Authors' reply. *Lancet Glob Heal*. 2018;6(6):e616-e617. doi:10.1016/S2214-109X(18)30229-8

45. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: A systematic review and meta-analysis. *Lancet Infect Dis*. 2005;5(1):42-52. doi:10.1016/S1473-3099(04)01253-8

46. Eisenberg JNS, Scott JC, Porco T. Integrating disease control strategies: Balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. *Am J Public Health*. 2007;97(5):846-852. doi:10.2105/AJPH.2006.086207

47. Eisenberg JNS, Trostle J, Sorensen RJD, Shields KF. Toward a Systems Approach to Enteric Pathogen Transmission: From Individual Independence to Community Interdependence. *Annu Rev Public Health*. 2012;33(1):239-257. doi:10.1146/annurev-publhealth-031811-124530

48. Esrey SA. Water, waste, and well-being: A multicountry study. *Am J Epidemiol*. 1996;143(6):608-623. doi:10.1093/oxfordjournals.aje.a008791

49. D'Mello-Guyett L, Yates T, Bastable A, et al. Setting priorities for humanitarian water, sanitation and hygiene research: A meeting report. *Confl Health*. 2018;12(1):10-14. doi:10.1186/s13031-018-0159-8

Tables

Table 1: Parameters of the SISE-RCT model. The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT.

Parameter	**Definition**	**Original scenario value**	**Sensitivity range**	**Disease elimination value**
ρ_0	Baseline WASH conditions (fraction of individuals in the community with intervention-level WASH infrastructure)	0.25	0–1	0.36
ρ	Compliance (fraction of individuals in intervention arm using intervention)	0.75	0–1	—
$R_0 = R_{0,1} + R_{0,2}$	Transmission potential (basic reproduction number)	1.25	1–1.5	1.17
$R_{0,1}/(R_{0,1} + R_{0,2})$	Intervenable fraction (fraction of transmission that the intervention could theoretically prevent	0.75	0–1	0.98
$1 - \varphi_\alpha$	Intervention efficacy for reducing shedding	0	—	—
$1 - \varphi_\beta$	Intervention efficacy for reducing transmission	0.75	0–1	0.98
ω	Community coverage fraction (fraction of community included in the intervention trial)	0.11	0–1	0.30

Table 2: WASH-B Bangladesh counterfactual scenarios and implementations. WASH = water, sanitation, & hygiene. *indicates parameters in the counterfactual simulation.

Category	Definition	What would have happened if…	Implementation
WASH conditions	Quality of WASH infrastructure at baseline	… no households had preexisting WASH conditions substantively equivalent to the intervention.	Pre-existing conditions were removed from individuals not in the corresponding intervention arms. (Akin to setting $\rho_0^* = 0$ in the single-intervention model).
Compliance	The extent to which individuals assigned to an intervention received it (fidelity) and used it (adherence)	… all households assigned an intervention received and used it.	All individuals in each intervention arm were modelled as using the intervention. (Akin to setting $\rho^* = 1$ in the single-intervention model).
Disease conditions	Disease prevalence at baseline in the absence of preexisting WASH conditions	… the disease pressure was greater	The basic reproduction number is increased such that the baseline prevalence in the absence of preexisting WASH conditions is doubled. (Optimization was used to determine the appropriate value of R_0^* for each parameter set.)
Intervenable fraction	Whether there are transmission pathways that are not affected by the intervention	… more of transmission was along pathways that could be intervened on.	The strength of the other pathway is reduced by 50% and replaced proportionally by the water and fomite pathways. ($R_{0,w}^* = R_{0,w} + \frac{1}{2}R_{0,o} \cdot R_{0,w}/(R_{0,w} + R_{0,f})$, $R_{0,f}^* = R_{0,f} + \frac{1}{2}R_{0,o} \cdot R_{0,f}/(R_{0,w} + R_{0,f})$, $R_{0,o}^* = \frac{1}{2}R_{0,o}$)
Efficacy	The extent to which using the intervention reduced transmission along relevant pathways	… the interventions provided a greater reduction in transmission.	The strength of the reduction in transmission from each intervention (and corresponding preexisting condition) is doubled. ($\varphi^* = \min(2\varphi, 1)$)
Community coverage	The fraction of the at-risk population in a cluster that was provided the intervention	… a different fraction of the population was provided the intervention.	Study coverage is 20%, …, 90%, 100%. ($\omega = 0.2, \ldots, 1.0$)

Table 3. Median intervention effectiveness and median percent change in intervention effectiveness in each intervention arm for each counterfactual scenario compared to the original scenario. Intervention effectiveness (ε) in intervention effectiveness is 1 minus the relative risk of diarrhea in the intervention arm vs the control arm in each scenario, expressed as a percentage. The column Δε gives the median change in intervention effectiveness in percentage points (not the change in median intervention effectiveness); a negative number reflects a decrease in intervention effectiveness.

	W		**S**		**H**		**WSH**		**N**		**WSHN**	
	ε	$\Delta\varepsilon$	ε	$\Delta\varepsilon$	ε	$\Delta\varepsilon$	ε	$\Delta\varepsilon$	ε	$\Delta\varepsilon$	ε	$\Delta\varepsilon$
Original scenario	8.3%	—	36.4%	—	33.1%	—	30.3%	—	34.0%	—	34.6%	—
No WASH baseline conditions	8.2%	-0.2%	22.5%	-14.0%	25.0%	-8.3%	20.7%	-9.3%	27.1%	-7.0%	25.9%	-8.3%
Double baseline disease prevalence	6.8%	-1.5%	15.0%	-21.4%	15.7%	-17.4%	18.5%	-11.9%	17.6%	-16.3%	24.0%	-10.7%
Full compliance	12.0%	+3.6%	36.5%	+0.1%	34.3%	+1.1%	35.0%	+4.4%	34.9%	+0.8%	38.7%	+4.1%
Half of other pathway transmission can be intervened on	14.4%	+6.1%	46.6%	+10.0%	42.8%	+9.6%	50.1%	+20.3%	37.9%	+3.7%	53.1%	+18.4%
Double efficacy of chlorination	15.0%	+6.8%		—		—	36.5%	+5.8%		—	40.4%	+5.5%
Double efficacy of latrine water seal		—	53.9%	+17.2%		—	45.5%	+14.8%		—	48.3%	+13.7%
Double efficacy of handwashing		—		—	53.2%	+20.2%	50.0%	+19.5%		—	52.6%	+17.9%
Double efficacy of nutrition		—		—		—		—	45.1%	+10.9%	44.3%	+9.7%
Increase community coverage to 20%	21.7%	+13.8%	42.4%	+6.0%	40.8%	+7.7%	60.8%	+30.4%	48.0%	+14.4%	74.7%	+39.9%

Figure captions

Figure 1: Single-intervention SISE-RCT model diagram with an attenuated exposure population and a regular exposure population interacting through shared environments. The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. The black lines denote infection and recovery, the blue lines denote shedding from infectious individuals into environmental compartments, the grey lines denote pick-up of pathogens from the environment by susceptible individuals, and the orange lines denote environmental pathogen decay. S_+ and I_+ denote susceptible and infectious fraction of the attenuate exposure population, and S_- and I_- denote susceptible and infectious fraction of the regular exposure population.

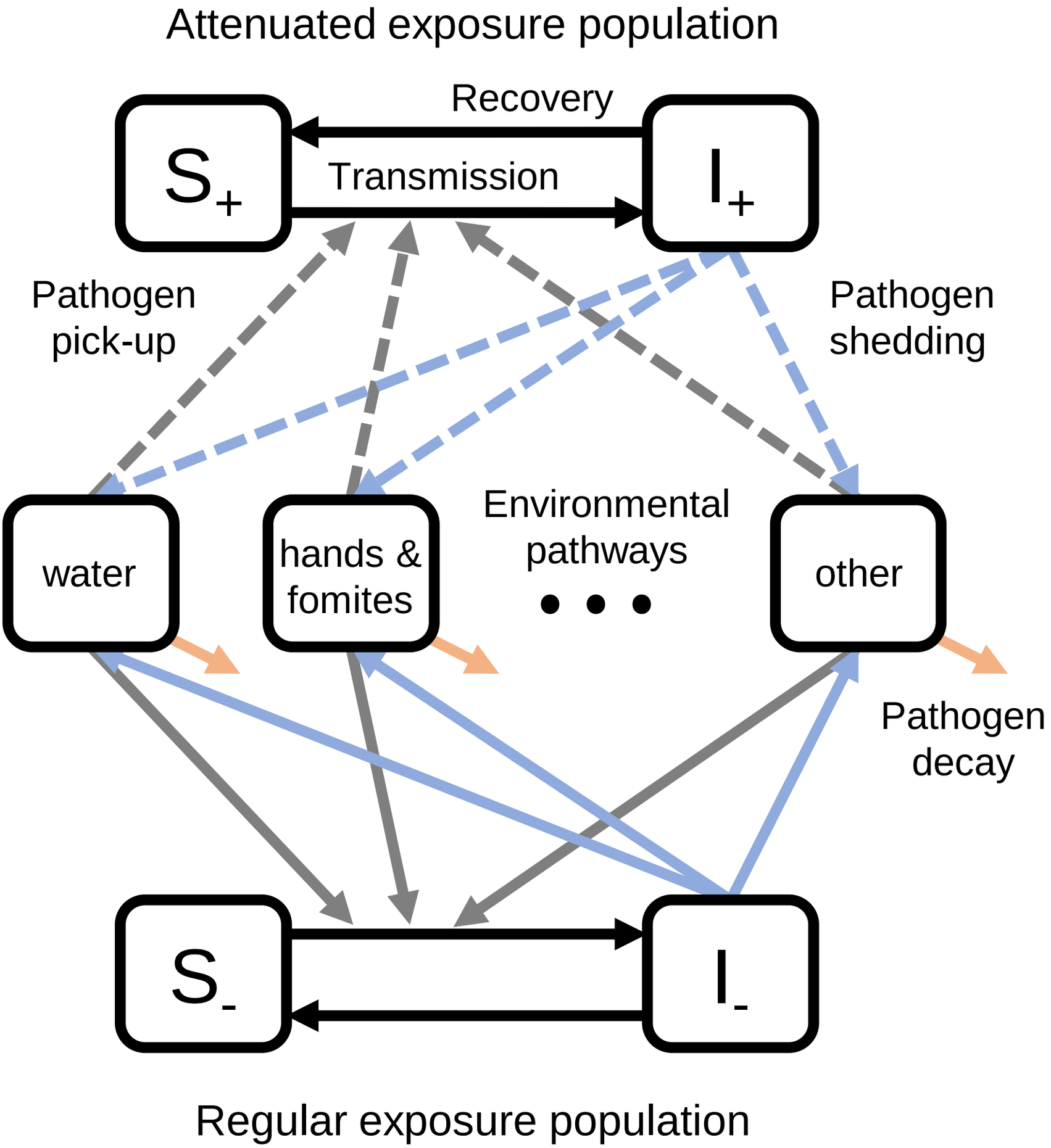
Figure 2: Intervention effectiveness as a function of WASH intervention factors. The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. A single-intervention implementation of the model was simulated at an original scenario (indicated by the white points), and the heatmaps denote how intervention effectiveness depends on each pair of WASH factors. The six WASH factors are baseline WASH conditions (fraction of individuals not enrolled in the intervention arm that are using a pre-existing intervention), compliance (fraction of individuals enrolled in the intervention arm that are using the intervention), disease transmission potential (summarize by the basic reproduction number R_0), intervenable fraction of transmission (how much of the transmission could be prevented in a perfect intervention), intervention efficacy (fraction reduction in transmission or shedding when using the intervention), and the community coverage fraction (fraction of the population enrolled in the trial). The inset enlarges the compliance vs coverage plot and overlays contour lines to show the interaction between the two factors on intervention effectiveness. When coverage is low and compliance is high, it is easier to increase intervention effectiveness by increasing coverage, but when coverage is higher and compliance is low, then it is easier to increase intervention effectiveness by increasing compliance. The underlying data are provided in Excel Table S1. WASH = water, sanitation, & hygiene.

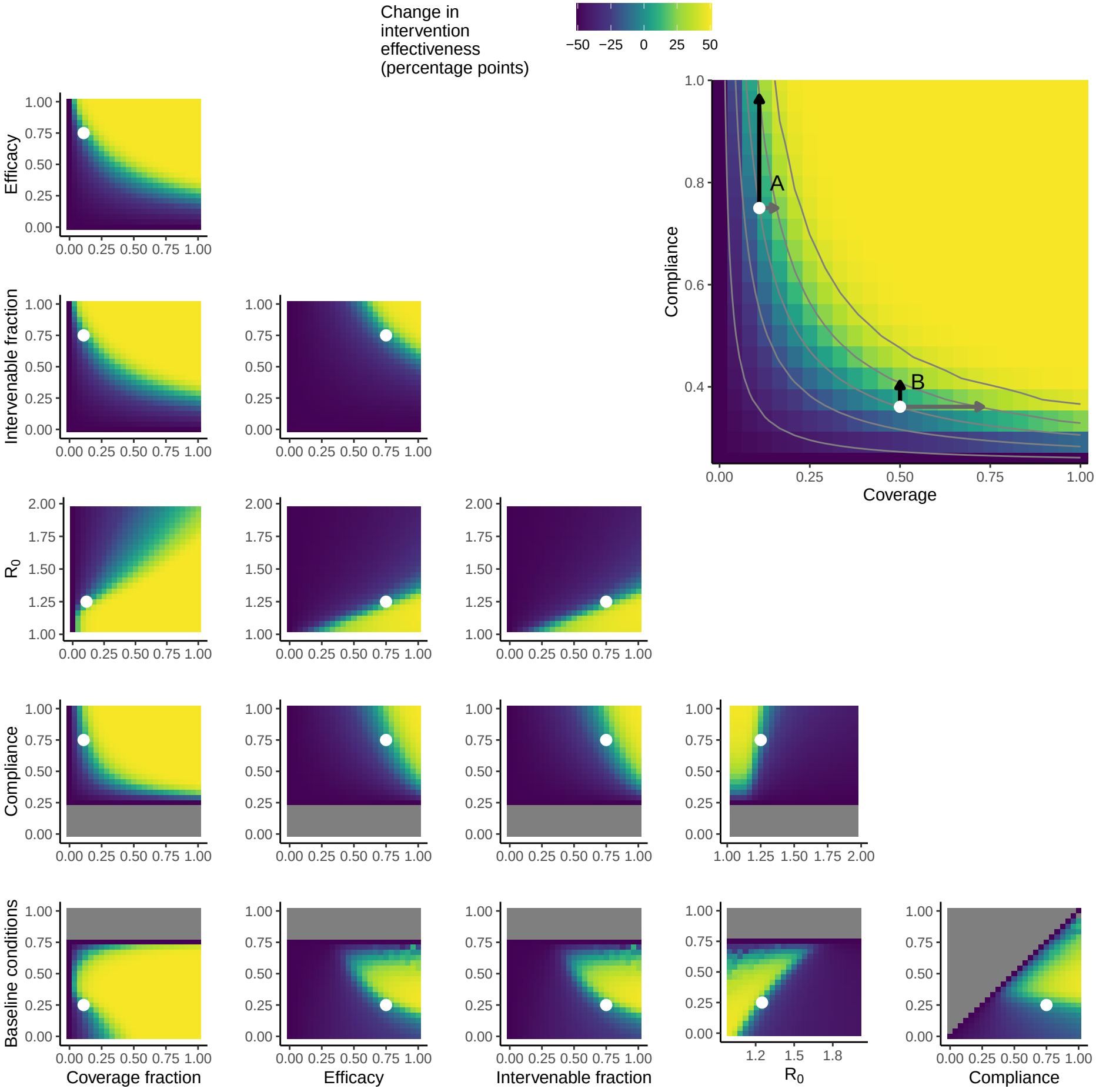
Figure 3: Percentage point change in intervention effectiveness compared to the original scenario in each counterfactual scenario. The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. Here, we applied it to data from the WASH-B Bangladesh trial, selecting 25,000 parameter sets consistent with the trial outcomes. We simulated each parameter set under each counterfactual scenario (Table 2). The violin plots give the distribution of values across the 25,000 simulations, with median points. The underlying data are provided in Excel Table S2A-H. W = water, S = sanitation, H = hygiene, N = nutrition.

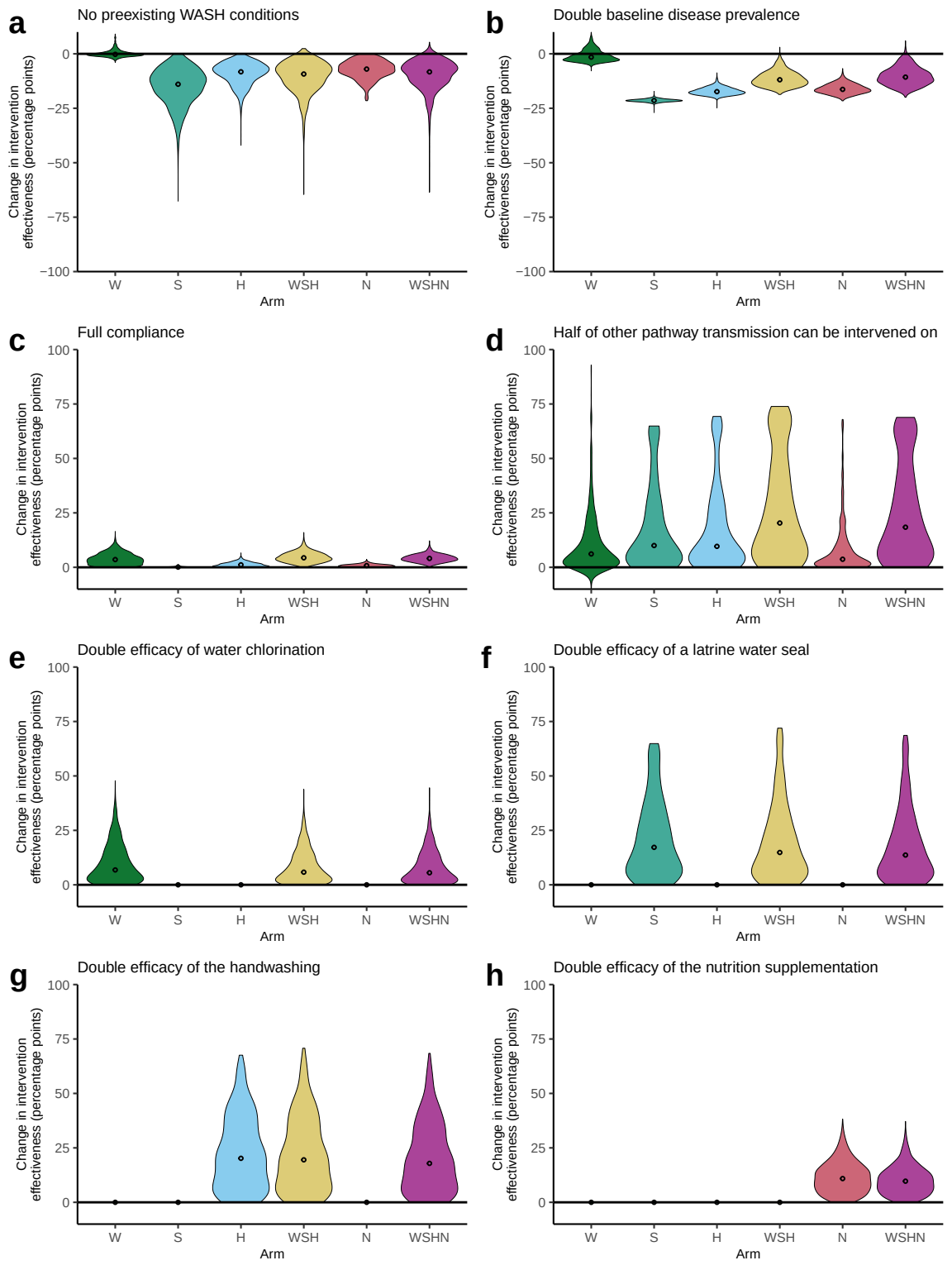
Figure 4: Percentage point change in intervention effectiveness compared to the original scenario in the 20% coverage counterfactual scenario (a) overall and (b) considering other parameters as potential effect modifiers (b-h). The SISE-RCT model is a compartmental susceptible-infectious-susceptible

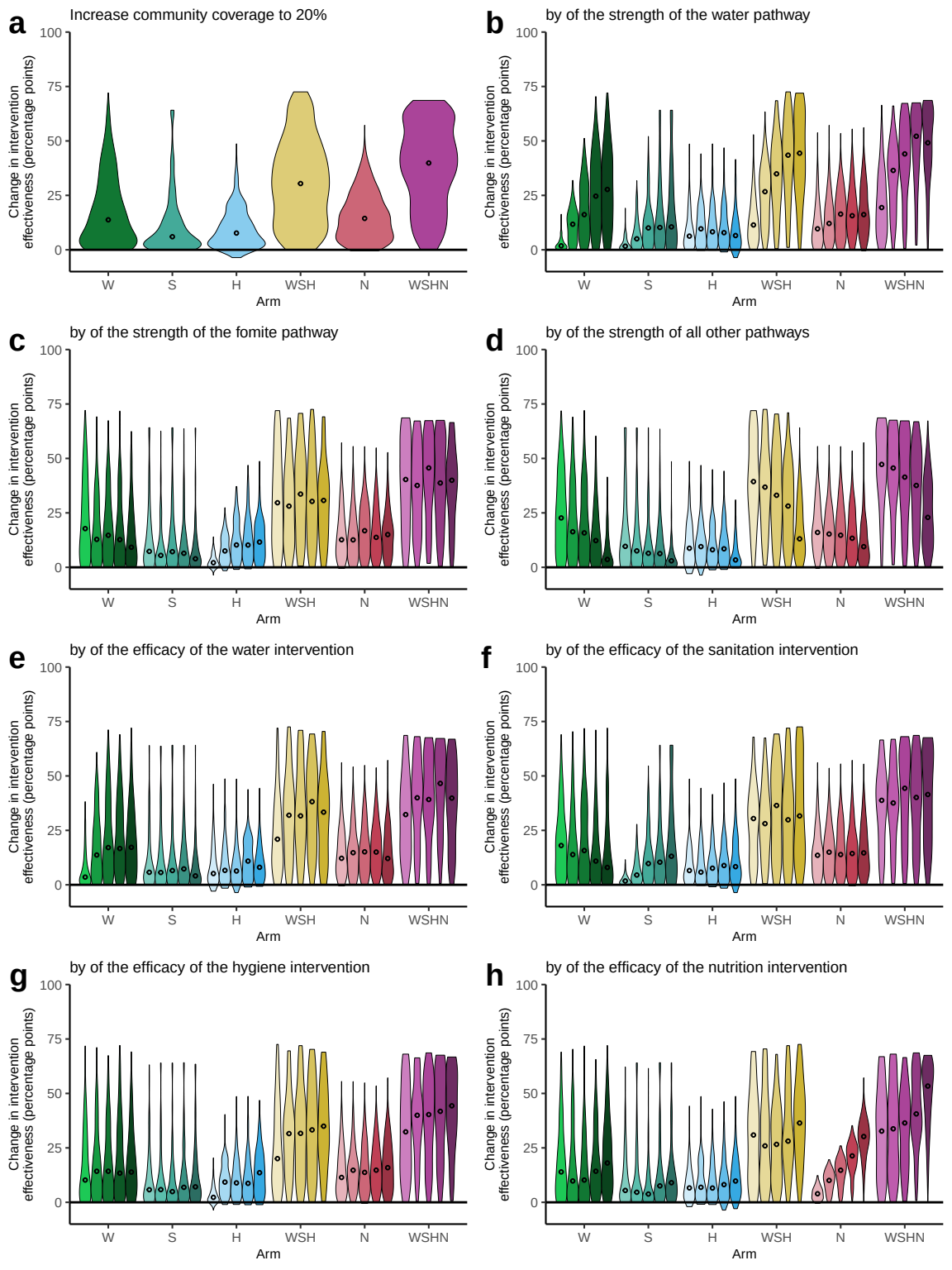
(SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. Here, we applied it to data from the WASH-B Bangladesh trial, selecting 25,000 parameter sets consistent with the trial outcomes. We simulated each parameter set under the 20% coverage counterfactual scenario (Table 2). The violin plots give the distribution of values across the 25,000 simulations, with median points. In plots b-h, the five violin plots give the distributions of the intervention effectiveness across quintiles of the listed potential effect modifier. The underlying data are provided in Excel Table S3A-H. W = water, S = sanitation, H = hygiene, N = nutrition.

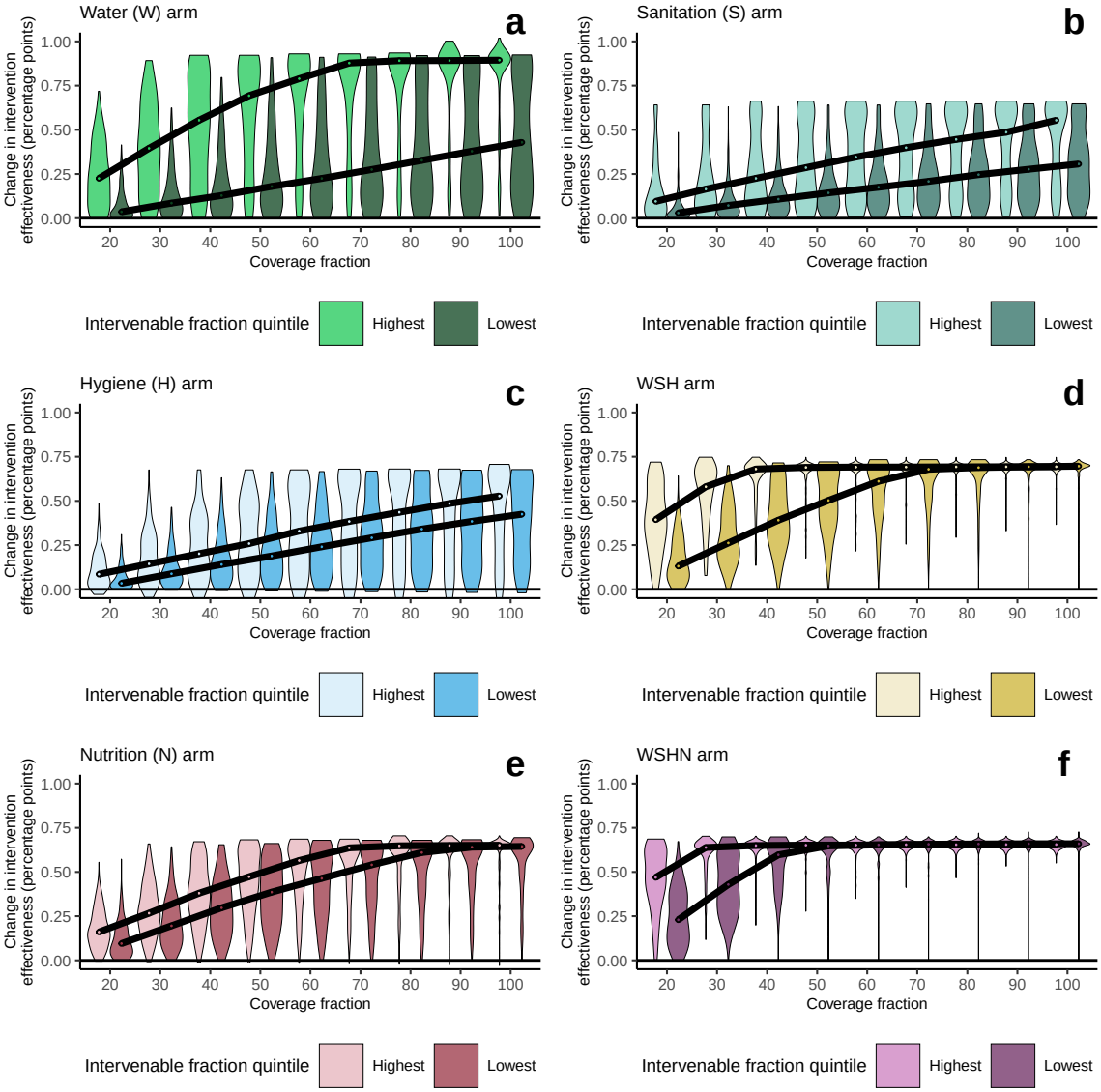
Figure 5: Percentage point change in intervention effectiveness compared to the original scenario in each arm for the lowest and highest quintiles of intervenable fraction. The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. Here, we applied it to data from the WASH-B Bangladesh trial, selecting 25,000 parameter sets consistent with the trial outcomes. We simulated each parameter set for coverage counterfactual scenarios ranging from 20% to 100% (Table 2). The violin plots give the distribution of values across the simulations, with median points and a line connecting the medians, for highest (dark) and lowest (light) quintiles of the intervenable fraction, i.e., the fraction of transmission that the interventions could directly act on. The underlying data are provided in Excel Table S4A-F. W = water, S = sanitation, H = hygiene, N = nutrition.


Attenuated exposure population
Recovery
S_+
Transmission
I_+
Pathogen pick-up
Pathogen shedding
water
hands & fomites
Environmental pathways
other
Pathogen decay
S_-
I_-
Regular exposure population


Change in intervention effectiveness (percentage points)
−50 −25 0 25 50
Efficacy
Intervenable fraction
R_0
Compliance
Baseline conditions
Coverage fraction
Efficacy
Intervenable fraction
R_0
Compliance
Coverage
A
B


a
No preexisting WASH conditions
b
Double baseline disease prevalence
c
Full compliance
d
Half of other pathway transmission can be intervened on
e
Double efficacy of water chlorination
f
Double efficacy of a latrine water seal
g
Double efficacy of the handwashing
h
Double efficacy of the nutrition supplementation
Change in intervention effectiveness (percentage points)
Arm
W
S
H
WSH
N
WSHN


a
Increase community coverage to 20%
b
by of the strength of the water pathway
c
by of the strength of the fomite pathway
d
by of the strength of all other pathways
e
by of the efficacy of the water intervention
f
by of the efficacy of the sanitation intervention
g
by of the efficacy of the hygiene intervention
h
by of the efficacy of the nutrition intervention
Change in intervention effectiveness (percentage points)
100
75
50
25
0
W
S
H
WSH
N
WSHN
Arm


Water (W) arm
a
Sanitation (S) arm
b
Hygiene (H) arm
c
WSH arm
d
Nutrition (N) arm
e
WSHN arm
f
Change in intervention effectiveness (percentage points)
1.00
0.75
0.50
0.25
0.00
20
30
40
50
60
70
80
90
100
Coverage fraction
Intervenable fraction quintile
Highest
Lowest