

1 **Title:** Understanding the effectiveness of water, sanitation, and hygiene interventions: a counterfactual
2 simulation approach to generalizing the outcomes of intervention trials

3 **Authors:** Andrew F. Brouwer PhD¹, Mondal H. Zahid PhD¹, Marisa C. Eisenberg PhD¹, Benjamin F.
4 Arnold PhD^{2,3}, Sania Ashraf PhD⁴, Jade Benjamin-Chung PhD^{5,6}, John M. Colford, Jr MD⁷, Ayse
5 Ercumen PhD⁸, Stephen P. Luby MD⁹, Amy J. Pickering PhD¹⁰, Mahbubur Rahman MBBS⁴, Alicia N.M.
6 Kraay PhD¹¹, Joseph N.S. Eisenberg PhD^{1,†}, Matthew C. Freeman PhD^{12,†}

7 †: these authors contributed equally

8 **Affiliations**

- 9 1. Department of Epidemiology, University of Michigan, Michigan, USA
- 10 2. Francis I. Proctor Foundation, University of California, San Francisco, San Francisco, CA, USA
- 11 3. Department of Ophthalmology, University of California, San Francisco, San Francisco, CA, USA
- 12 4. Environmental Interventions Unit, Infectious Disease Division, The International Centre for
13 Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh
- 14 5. Department of Epidemiology and Population Health, Stanford University, CA, USA
- 15 6. Chan Zuckerberg Biohub, San Francisco, CA, USA
- 16 7. School of Public Health, University of California Berkeley, Berkeley, CA, USA
- 17 8. Department of Forestry and Environmental Resources, North Carolina State University, Raleigh,
18 NC, USA
- 19 9. Division of Infectious Diseases and Geographic Medicine, Stanford University, CA, USA
- 20 10. Department of Civil and Environmental Engineering, University of California Berkeley,
21 Berkeley, CA, USA
- 22 11. Institute for Disease Modeling, a program within the Global Health Division of the Bill &
23 Melinda Gates Foundation, Seattle, Washington, USA
- 24 12. Rollins School of Public Health, Emory University, Atlanta, GA, USA

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

27 **Declaration of conflicts of interest:** ANMK's contributions were directly funded by the Bill & Melinda
28 Gates Foundation and not as part of the foundation grant to the authors. All other authors declare no
29 conflicts of interest.

30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

Abstract

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

Objectives: We aimed to estimate how sensitive the intervention effectiveness found in the WASH Benefits Bangladesh randomized controlled trial was to underlying WASH contextual and intervention factors (e.g., baseline disease prevalence, compliance, community coverage, efficacy) and to generalize the results of the trial other contexts or scenarios.

Methods: We developed a disease transmission model 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, compliance, and efficacy, as well as preexisting WASH conditions and baseline disease burden, impacted intervention effectiveness.

Results: Increasing community coverage had the greatest impact on intervention effectiveness (e.g., median increases in effectiveness of 34.0 and 45.5% 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 by lower levels of preexisting WASH conditions or increased baseline disease burden. Individual interventions had complementary but not synergistic effects when combined.

Discussion: To realize the expected health gains, future WASH interventions must address community coverage and transmission along pathways not traditionally covered by WASH. The effectiveness of individual-level WASH improvements will be blunted the further the community is from the high community coverage needed to achieve herd protection.

57 **Introduction**

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

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

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

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

96 The aims of this work are to evaluate hypotheses about what led to the sub-optimal reductions in diarrhea
97 among intervention households in an RCT (WASH-B Bangladesh) using a compartmental transmission
98 model and to provide a framework to support planning of WASH interventions and context-specific
99 WASH programming. We previously developed this model framework accounting for multiple
100 environmental transmission pathways, shared environments, pre-existing WASH conditions, and
101 adherence to multiple interventions and applied it to the empirical trial data.³⁰ Our approach generates
102 thousands of combinations of coverage, intervention efficacy, and transmission pathway strengths that
103 could reasonably underlie the trial results. In this analysis, we leveraged those parameter combinations to
104 simulate how intervention effectiveness would have been different under alternate scenarios. These
105 counterfactual simulations provide evidence for policy recommendations, programmatic targets, and an
106 evaluation framework for next-generation WASH interventions.^{20,21,24}

107 **Methods**

108 **Summary of approach.** In prior work, we developed a compartmental transmission model framework to
109 explain the outcomes of a RCT.³⁰ Then, we explored how effectiveness of a single intervention depends
110 on six WASH factors:³¹

- 111 • *Preexisting WASH conditions:* We account for the fact that a fraction of the population already
112 has WASH infrastructure comparable to that provided by the intervention trial.
- 113 • *Disease transmission potential:* We account for the baseline disease prevalence through the basic
114 reproduction number R_0 , a summary measure of the disease transmission potential. Note that,
115 given the values of the other WASH factors, there is a one-to-one correspondence between the
116 disease prevalence and the basic reproduction number. In this analysis, we vary the baseline
117 disease prevalence rather than R_0 , as it what intervention trials measure.
- 118 • *Intervention compliance:* We account for both intervention fidelity (whether the intervention was
119 delivered) and adherence (whether participants used the intervention), defining compliance as the
120 fraction of participants assigned to an intervention who are using it.

- 121 • *Intervenable fraction of transmission*: Any individual intervention targets some, but often not all,
122 of the transmission pathways that pathogens exploit. We account for how much transmission is
123 along pathways that the intervention attenuates (even if imperfectly) and how much even a
124 perfect intervention would not affect.
- 125 • *Intervention efficacy*: Because interventions do not result in perfect reduction of transmission or
126 shedding, we account for the actual reduction.
- 127 • *Community coverage fraction*: Intervention trials typically only provide the intervention to a
128 subset of the population. We account for the fraction of the population is enrolled in the trial.

129
130 In this analysis, we use a multi-intervention version of the model to investigate outcomes in the WASH-B
131 Bangladesh randomized controlled trial specifically. In prior work, we demonstrated how to find
132 mechanistic parameter sets that were consistent with individual-level diarrheal outcomes.³⁰ Here, we use
133 these mechanistic parameter sets to simulate what intervention effectiveness would have been in each of
134 the WASH-B Bangladesh trial arms under each of six counterfactuals corresponding to the six WASH
135 factors above, accounting for uncertainty in the parameters underlying the real data (original scenario). By
136 simulating what the intervention effectiveness would have been in the trial under alternative
137 circumstances, we evaluated the extent to which each factor may have contributed to the observed
138 outcomes.

139 **Data.** The WASH-B Bangladesh trial was a cluster-randomized trial of the efficacy of water, sanitation,
140 hygiene, and nutrition interventions, alone and in combination, on diarrhea prevalence and linear
141 growth.¹⁴ The investigators measured (child-guardian-reported, past-seven-day, all-cause) diarrheal
142 prevalence in children at three time points approximately one year apart. Households in the study area are
143 typically organized into compounds in which a patrilineal family shares a common space and resources,
144 such as a water source and latrine. A total of 5551 compounds were enrolled, contingent on having a
145 pregnant woman in her second trimester during the enrollment period. The study followed one or more
146 target children born after baseline, as well as any other children in the compound who were under age 3 at
147 baseline. These compounds were grouped into 720 clusters. Each cluster was assigned to one of seven
148 arms testing combinations of four interventions: water chlorination (W), a double-pit, pour-flush
149 improved latrine (S), handwashing with soap and water (H), and supplementary nutrition sachets (N). The
150 control arm (C) consisted of 180 compounds, while 90 were assigned to each of the water (W), sanitation
151 (S), handwashing (H), nutrition (N), combined water, sanitation, and handwashing (WSH), and all
152 interventions (WSH-N) arms. Specific details on trial design, interventions, and results are published
153 elsewhere^{14,32}. We assessed whether any individual was using an intervention or a substantively

154 equivalent preexisting WASH condition through four indicators defined and assessed by the investigators:
155 detection of free chlorine in water (W), latrine with a functional water seal (S), handwashing station with
156 soap and water (H), at least 50% of nutrition sachets consumed (N). The W and H interventions were
157 intended for the households of the target children, but we were not able to determine whether other
158 children in the compound were in that household or not. For this analysis, we assumed that non-target
159 children were covered by the interventions; any misspecification will attenuate the estimated efficacy of
160 the W and H interventions.¹⁴ We removed individuals with negative reported ages (n=2), missing reported
161 diarrhea (n=2,745), or missing in any of the four use indicators (n=2,660), which left 17,187 individual
162 observations (76% of the original sample) over the three surveys.

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

164 **Model.** Our compartmental transmission model, denoted SISE-RCT, is a susceptible-infectious-
165 susceptible (SIS) model with transmission through environmental (E) compartments and simulated to
166 steady state to approximate an RCT.³⁰ The SISE-RCT model accounts for the six key mechanistic factors
167 underlying WASH RCT outcomes described above. We previously developed the single-intervention
168 model,³¹ and we include a description of it in the supplemental material for convenience.

169 As discussed in the Data section above, the WASH-B Bangladesh trial included 720 clusters of
170 households each assigned to one of seven arms (control, W, S, H, N, WSH, WSHN). We extended the
171 single-intervention SISE-RCT model to a multi-intervention model by accounting for transmission across
172 three environmental pathways (water, fomites & hands, and all others combined), four interventions
173 applied individually (W, S, H, N) and in combination (WSH, WSHN), and individual-level compliance
174 with interventions or preexisting conditions. In brief, we modeled each of the 720 clusters with
175 susceptible and infectious compartments for each of $2^4=16$ combinations of interventions/conditions
176 depending on household adherence, i.e., in every cluster, we modeled the infection prevalence for each
177 combination of having or not having each intervention or equivalent preexisting WASH condition. For
178 example, for a cluster in the WSH arm, we estimated how many people were not using any interventions,
179 how many were using W only, how many were using S only, etc., and what the infection prevalence was
180 among each group given their collective interaction through the shared environments.

181 Specifically, for a given cluster, we denote the fraction of the population that is susceptible to infection
182 and is using intervention(s) or preexisting WASH condition(s) i in $\{0, W, S, H, N, WS, \dots, WSHN\}$ as S_i ,
183 where 0 indicates the use of no intervention or preexisting WASH condition. Analogously, we denote the
184 fraction of the population that is infected analogously by I_i . The intervention and control arms are
185 simulated separately. The populations with regular and attenuated exposure are modeled in every cluster

186 simulation, accounting for the fraction of the population 1) enrolled in the study (the community
187 coverage, ω), 2) with preexisting WASH conditions (ρ_0), and 3) complying with the intervention (ρ).
188 Note that ρ_0 and ρ are vectors of length 16 that each sum to 1; i.e., everyone is categorized into one of the
189 16 exposure groups. For each cluster, the overall fraction of the population in each exposure group is
190 given by the vector

$$N = (1 - \omega)\rho_0 + \omega\rho. \#(1)$$

191 That is, the fraction of the population enrolled in the study (ω) follows the intervention compliance
192 distribution of exposure groups (ρ), and the fraction of the population not enrolled ($1 - \omega$) follows the
193 preexisting WASH condition distribution of exposure groups (ρ_0). In the control arm, $\rho = \rho_0$.

194 The environment is partitioned into the three environmental pathways: water (E_w), fomites & hands (E_f),
195 and all other pathways (E_o). We assume that chlorination reduces transmission along the water pathway,
196 sanitation reduces shedding into the water pathway, handwashing reduces transmission along the fomite
197 pathway, and nutrition reduces susceptibility (transmission) for all three pathways. For each of the three
198 pathways j , an environmental compartment E_j is characterized by the shedding into the environment (α_j),
199 the decay of pathogens in the environment (ξ_j), and the transmission of pathogens from the environment
200 to susceptible individuals (β_j). The relative magnitude of shedding into E_j and relative transmission from
201 E_j for the attenuated compared to the exposed populations are given by ϕ_{α_j} and ϕ_{β_j} , respectively. We
202 also accounted for the possibility that the preexisting conditions were less efficacious than the RCT
203 intervention. Once infected, individuals clear the infection at rate γ .

204 The model diagram of the multi-intervention SISE-RCT is given in Figure 1, although only two of the
205 sixteen different exposure populations are shown. The full equations are given below (Eqs 2). A
206 transmission term $\beta_j E_j$ denotes transmission from the environmental pathway j . The transmission term
207 $\beta_j E_j$ is attenuated by ϕ_{β_j} only for people in an attenuated exposure group (S_i) using an intervention or
208 preexisting condition that reduces transmission from pathway j , and contamination of that environmental
209 pathway is attenuated by ϕ_{α_j} only for infectious people in an exposure group (I_i) using an intervention or
210 preexisting condition that reduces shedding into pathway j . In the following equations, the subscripts w, f ,
211 and o represent the water, fomite and other pathways. The subscripts θ, W, S, H, N , and any combinations
212 represent the exposure groups as defined above. Parameters ω, ρ , and ρ_0 do not show up in these
213 equations but are accounted for in the constraints, as discussed below. For brevity, we omit the $\frac{dS_i}{dt}$
214 equations, each of which is given by $\frac{dS_i}{dt} = -\frac{dI_i}{dt}$ for the corresponding subpopulation.

$$\frac{dI_0}{dt} = (\beta_w E_w + \beta_f E_f + \beta_o E_o) S_0 - \gamma I_0 \#(2)$$

$$\frac{dI_W}{dt} = (\varphi_{\beta_w, W} \cdot \beta_w E_w + \beta_f E_f + \beta_o E_o) S_W - \gamma I_W$$

$$\frac{dI_S}{dt} = (\beta_w E_w + \beta_f E_f + \beta_o E_o) S_S - \gamma I_S$$

$$\frac{dI_H}{dt} = (\beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_H - \gamma I_H$$

$$\frac{dI_N}{dt} = \varphi_{\beta, N} \cdot (\beta_w E_w + \beta_f E_f + \beta_o E_o) S_H - \gamma I_H$$

$$\frac{dI_{WS}}{dt} = (\varphi_{\beta_w, W} \cdot \beta_w E_w + \beta_f E_f + \beta_o E_o) S_{WS} - \gamma I_{WS}$$

$$\frac{dI_{WH}}{dt} = (\varphi_{\beta_w, W} \cdot \beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_{WH} - \gamma I_{WH}$$

$$\frac{dI_{WN}}{dt} = \varphi_{\beta, N} \cdot (\varphi_{\beta_w, W} \cdot \beta_w E_w + \beta_f E_f + \beta_o E_o) S_{WN} - \gamma I_{WN}$$

$$\frac{dI_{SH}}{dt} = (\beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_{SH} - \gamma I_{SH}$$

$$\frac{dI_{SN}}{dt} = \varphi_{\beta, N} \cdot (\beta_w E_w + \beta_f E_f + \beta_o E_o) S_{SN} - \gamma I_{SN}$$

$$\frac{dI_{HN}}{dt} = \varphi_{\beta, N} \cdot (\beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_{HN} - \gamma I_{HN}$$

$$\frac{dI_{WSH}}{dt} = (\varphi_{\beta_w, W} \cdot \beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_{WSH} - \gamma I_{WSH}$$

$$\frac{dI_{WSN}}{dt} = \varphi_{\beta, N} \cdot (\varphi_{\beta_w, W} \cdot \beta_w E_w + \beta_f E_f + \beta_o E_o) S_{WSN} - \gamma I_{WSN}$$

$$\frac{dI_{WHN}}{dt} = \varphi_{\beta, N} \cdot (\varphi_{\beta_w, W} \cdot \beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_{WHN} - \gamma I_{WHN}$$

$$\frac{dI_{SHN}}{dt} = \varphi_{\beta, N} \cdot (\beta_w E_w + \varphi_{\beta_f, H} \cdot \beta_f E_f + \beta_o E_o) S_{SHN} - \gamma I_{SHN}$$

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

$$\frac{dE_w}{dt} = \alpha_w \left(\sum_{S \text{ not in } i} I_i + \varphi_{\alpha,S} \sum_{S \text{ in } i} I_i \right) - \xi_w E_w$$

$$\frac{dE_f}{dt} = \alpha_f \left(\sum_i I_i \right) - \xi_f E_f$$

$$\frac{dE_o}{dt} = \alpha_o \left(\sum_i I_i \right) - \xi_o E_o$$

215 To find the steady state values (denoted by *) for the human compartments in the intervention arm, we set
 216 the above equations to 0 and simplify:

$$0 = (R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_0^* - I_0^*, \#(2)$$

$$0 = (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_W^* - I_W^*,$$

$$0 = (R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_S^* - I_S^*,$$

$$0 = (R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_H^* - I_H^*,$$

$$0 = \varphi_{\beta,N} \cdot (R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_N^* - I_N^*,$$

$$0 = (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{WS}^* - I_{WS}^*,$$

$$0 = (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{WH}^* - I_{WH}^*,$$

$$0 = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{WN}^* - I_{WN}^*,$$

$$0 = (R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{SH}^* - I_{SH}^*,$$

$$0 = \varphi_{\beta,N} \cdot (R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{SN}^* - I_{SN}^*,$$

$$0 = \varphi_{\beta,N} \cdot (R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{HN}^* - I_{HN}^*,$$

$$0 = (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{WSH}^* - I_{WSH}^*,$$

$$0 = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{WSN}^* - I_{WSN}^*,$$

$$0 = \varphi_{\beta,N} \cdot (\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o) S_{WHN}^* - I_{WHN}^*,$$

$$0 = \varphi_{\beta,N} \cdot \left(R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o \right) S_{SHN}^* - I_{SHN}^*,$$

$$0 = \varphi_{\beta,N} \cdot \left(\varphi_{\beta_w,W} \cdot R_{0,w} \bar{E}_w + \varphi_{\beta_f,H} \cdot R_{0,f} \bar{E}_f + R_{0,o} \bar{E}_o \right) S_{WSHN}^* - S_{WSHN}^*,$$

$$\bar{E}_w = \sum_{S \text{ not in } i} I_i^* + \varphi_{\alpha,S} \sum_{S \text{ in } i} I_i^*,$$

$$\bar{E}_f = \bar{E}_o = \sum_i I_i^*.$$

217 Here, $R_{0,j} = \frac{\alpha_j \beta_j}{\xi_j \gamma}$ is the pathway-specific reproduction number for transmission through environment

218 E_i . The variables $\bar{E}_j = \xi_j E_j^* / \alpha_j$ are conveniently scaled environmental steady states. For this model, the
 219 overall basic reproduction number is $R_0 = R_{0,w} + R_{0,f} + R_{0,o}$, denoting the sum of the transmission
 220 potential through each pathway.

221 To solve for the steady states solutions for our 32 state variables (16 exposure groups times two
 222 susceptible/infection states), we solve the nonlinear system of equations (Eqs (3)) subject to the
 223 population constraint given in Eqs (1). The prevalence of disease in the population is denoted $\pi^* = \sum I_i^*$.
 224 The prevalence in an intervention arm π^* is compared to the prevalence π_c^* in the control arm, and the
 225 intervention effectiveness for that arm is defined as $\varepsilon = (\pi_c^* - \pi^*) / \pi_c^*$, namely the fractional reduction in
 226 prevalence in the intervention arm relative to the control arm.

227 In summary, the model includes 18 parameters: i) the overall basic reproduction number R_0 , which
 228 defines the transmission potential in the control arm at baseline, ii) two parameters partitioning R_0 into
 229 the strengths of the drinking water $R_{0,w}$, fomite & hands $R_{0,f}$, and all other transmission pathways $R_{0,o}$,
 230 iii) eight relative reproduction numbers accounting for systematic differences in disease pressure over the
 231 trial time periods (baseline, midline, and endline) and across arms independently, iv) the community
 232 coverage ω , and v) efficacy parameters defining the effect of each intervention (four) or preexisting
 233 WASH condition (two) on the transmission pathways (the W intervention (chlorination) reduces
 234 transmission via the water pathway $\varphi_{\beta_w,W}$; the S intervention (latrine with water seal) reduces shedding
 235 into the shared water environment with different efficacy for preexisting conditions $\bar{\varphi}_{\alpha_w,S}$ and the trial
 236 intervention $\varphi_{\alpha_w,S}$; the H (handwashing with soap and water) reduces transmission via the fomite
 237 pathway with different efficacy for preexisting conditions $\bar{\varphi}_{\beta_f,H}$ and trial intervention $\varphi_{\beta_f,H}$; and the N
 238 intervention (nutrition supplementation) reduces susceptibility to all transmission $\varphi_{\beta,N}$). In the Eqs (3),
 239 the basic reproduction number parameters are adjusted by the time and arm-specific and relative basic

240 reproduction numbers corresponding to the cluster being modeled, and the intervention efficacy
241 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
242 interventions, respectively, and at baseline.

243 When solving for the steady state of these differential equations for a given cluster in a given time period,
244 we use the distribution of interventions and preexisting WASH conditions ρ recorded in the data for those
245 participants and assume that participants not in the study have the same distribution of preexisting
246 conditions as the control arm participants ρ_0 . We solved this system using the `nleqslv` package in R.

247 To fit the model to the trial data, we employed a hybrid sampling-importance resampling and estimation
248 framework to obtain 50,000 parameter combinations that represented a good fit to the diarrheal outcomes
249 of each participant using a Bernoulli statistical likelihood.³³ We resampled, with replacement, from our
250 initial 50,000 parameter combinations, based on their goodness of fit; 3,774 unique parameter
251 combinations were included in the final sample, with varying frequency. These parameters sets are similar
252 to but not exactly the same as in Brouwer et al,³⁰ as they include a small code correction and use the
253 computational improvement described previously.³¹ The fit to the data is given in Figure S1, and the
254 distributions of parameters are given in Figures S2-5.

255 **WASH-B Bangladesh counterfactual analysis.** We conducted two types of counterfactual analysis.

256 First, we estimated the counterfactual intervention effectiveness in each arm across a range of each of the
257 WASH factors starting from the scenario based on the median value of each parameter (which resulted in
258 a fit close to the best-fit and was more representative of the parameter distributions than the specific best-
259 fit parameter set). Because we are using the actual preexisting WASH conditions and intervention
260 compliance recorded for each individual in the trial, there is not a well-defined way to continuously scale
261 these two factors. So, we only compare the actual simulation to a “no preexisting conditions” and “full
262 compliance” counterfactual, respectively.

263 Second, to account for the uncertainty in the parameters underlying the actual intervention outcomes, for
264 each of the 50,000 parameter sets k identified by fitting the model to WASH-B Bangladesh, we defined
265 the corresponding original scenario matching the WASH-B Bangladesh trial outcomes and the
266 corresponding intervention effectiveness ε_0^k . Because we are now accounting for uncertainty across these
267 50,000 parameters sets, it is not possible to succinctly capture changes as we continuously vary the
268 factors. Thus, we considered six specific counterfactual scenarios, detailed in Table 2. Any parameter sets
269 that eliminated disease in the control arm in a counterfactual simulation were censored from the results as
270 they did not provide information on intervention effectiveness.

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

280 The counterfactuals scenarios are not intended to be “plausible” for some specific, real-world changes.
281 Indeed, changing the contextual factors of the preexisting WASH conditions and baseline disease
282 prevalence is not possible. Instead, we can imagine these counterfactuals representing running the same
283 trial in a different location to assess what the results would have been. The intervenable fraction is also
284 not changeable for a given intervention (and underlying set of pathogens) but could be changed by adding
285 additional intervention aspects to reduce transmission along other pathways. More broadly, we believe
286 that investigating a broad range of counterfactual scenarios improves our understanding of the disease–
287 intervention system, so that more effective interventions may be designed in the future.

288 **Results**

289 **WASH-B Bangladesh counterfactual analysis from the median parameters**

290 At the median parameter values (Figs S2-S5), intervention effectivenesses were 8.5% in the W arm,
291 40.8% in the S arm, 37.2% in the H arm, 32.1% in the WSH arm, 35.0% in N arm, and 34.6% in the
292 WSHN arm, reflecting the results of the WASH-B Bangladesh trial. We estimated that removing all
293 preexisting sanitation and hygiene infrastructure would have resulted in a modest reduction in
294 intervention efficacy in all arms except for W, where the reduction was negligible (Figure 2a). Increasing
295 baseline disease prevalence (by changing increasing the R_0 parameter) nonlinearly decreased intervention
296 effectiveness in all arms, with decreasing reductions in effectiveness as baseline prevalence increased
297 (Figure 2b). We estimated that there would be negligible-to-modest improvements in intervention
298 effectiveness if there was full compliance (Figure 2c). Increasing the intervenable fraction (by reducing
299 the strength of the “other” pathway and proportionally increasing the strength of the water and fomite &
300 hands pathways while keeping the overall R_0 constant), increasingly improved the intervention
301 effectiveness (Figure 2d), with the WSH and WSHN interventions nearly achieving disease elimination if
302 transmission were 100% intervenable. Increasing community coverage increased effectiveness

303 approximately linearly, with each intervention achieving disease elimination at a different level of
304 community coverage (Figure 2e; W: 75%, S: 90%, H: 75%, WSH: 35%, N: 55%, WSHN: 30%).
305 Increasing intervention efficacy approximately linearly increased intervention effectiveness in the
306 corresponding arms (Figure 2f-i). Increasing efficacy of the sanitation and hygiene interventions to 100%
307 resulted in approximate disease elimination in the corresponding arms, but elimination would not have
308 been achieved by increasing the efficacy of the W and N interventions.

309 **WASH-B Bangladesh counterfactual analysis accounting for parameter uncertainty**

310 Because we do not know that the median values of the parameters in the scenario investigated in the first
311 counterfactual analysis are accurate, we also considered the distribution of changes in intervention
312 effectiveness when each counterfactual scenario (Table 2) was applied to the distribution of samples that
313 fit the original data well. The median baseline disease prevalence in the original scenario was 7.1% (range
314 5.9–8.2%), decreasing to 5.7% (range 5.2–6.3%) for midline/endline (Figure S1). The median
315 intervention effectivenesses were 8.2% in the W arm, 36.4% in the S arm, 33.1% in the H arm, 30.3% in
316 the WSH arm, 34.0% in N arm, and 34.6% in the WSHN arm (Table 3). The percentage point change
317 varied across arms in each counterfactual scenario. Figure 3 shows the distribution of percentage point
318 change in intervention effectiveness over the 50,000 parameter sets for each arm and counterfactual
319 scenario, and Table 3 gives the median values.

320 *Eliminate preexisting WASH Conditions.* We found that implementing the interventions in a community
321 with no handwashing stations with soap and water or latrines with water seals would have likely resulted
322 in less effective interventions compared to the actual community's higher baseline WASH conditions
323 (e.g., 9.5 percentage points less in the WSH arm; Figure 3a). The W arm is the exception because it had
324 lower effectiveness in the original scenario. The uncertainty in change in intervention effectiveness in
325 each arm was largely driven by uncertainty in what the baseline disease prevalence would have been in
326 the counterfactual scenario (median 8.9%, range 6.4–23.1%).

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

332 *Full compliance.* The impact of increasing intervention adherence was negligible-to-modest (e.g., 4.4
333 percentage points more in the WSH arm; Figure 3c). Note that intervention compliance, as defined by the
334 investigators, was already high.^{14,34}

335 *Half of the other pathway transmission can be intervened on.* We found that intervention effectiveness
336 could have been greater if more of the total disease transmission was via the water and fomite pathways
337 rather than through pathways that were not intervened on (e.g., 20.1 percentage points more in the WSH
338 arm; Figure 3d). There was potential for a substantial increase in intervention effectiveness, as indicated
339 by the distribution of the individual simulation outcomes, but the median impact was modest, with a less
340 than 25 percentage point increase in effectiveness in the multi-intervention arms and a less than 15
341 percentage point increase in the single-intervention arms (Table 3). The uncertainty in the potential
342 impact was largely driven by uncertainty in how much of the disease transmission was through other
343 pathways in the original scenario.

344 *Double intervention efficacy.* We assessed the impact of increasing efficacy—defined as increasing the
345 reduction of transmission along the relevant pathway(s)—of the four interventions. We found that in each
346 of these increased efficacy scenarios, substantial increases in intervention efficacy could have improved
347 intervention effectiveness in the corresponding arms (Figure 3e–h), with median improvements between
348 5% and 20% points.

349 *Increase community coverage.* The median estimated community coverage in the trial was 5.4%, but this
350 estimate was highly uncertain, ranging from nearly 0% to 20% (Figure S5). For our main coverage
351 counterfactual, we increased the community coverage in each simulation to 20.0%, chosen as a
352 substantial but not unreasonable increase in coverage.³⁵ This counterfactual scenario was associated with
353 the greatest median increase in intervention effectiveness (among all households now covered by the
354 intervention) of any of the considered counterfactual scenarios (e.g., 34.0 and 45.5 percentage points
355 more in the WSH and WSHN arms; Figure 4a). Following the results of the single-intervention model
356 that highlighted that the effect of coverage depended on the other WASH factors, we plotted the
357 intervention effectiveness distributions for this community coverage counterfactual by quintiles of the
358 values of the other WASH factors. Effect modification is present if the effect of increased coverage
359 depends on the quintile of the WASH factor. Note that when looking at quintiles of one factor, the values
360 of the other factors may not be evenly distributed across the quintiles if values of the factors are
361 correlated. We found that the increase intervention effectiveness with increased community coverage in
362 the W, S, WSH, and WSHN intervention arms depended partly on the strength of transmission via the
363 water pathway (Figure 4b). The increases in intervention effectiveness in these arms could only reach
364 their full potential if the strength of the water pathway were high. A similar, but more modest effect was
365 seen for the H arm and the strength of the fomite pathway (Figure 4c). The greatest overall effect modifier
366 of the impact of increased coverage on intervention effectiveness is the strength of the other pathways
367 (i.e., the intervenable fraction, Figure 4d). When the strength of other pathways was high, increasing

368 coverage had less of an impact. Intervention efficacy also modified the impact of increased coverage but
369 only in the intervention arms with those interventions (Figure 4e–h).

370 To further understand the joint impact of community coverage and the intervenable fraction (i.e., the
371 strength of the other transmission pathway), we simulated the intervention effectiveness as a function of
372 increased coverage for the highest and lowest quintiles of intervention completeness (Figure 5). The
373 impact of increased coverage on intervention effectiveness depended on the intervenable fraction most
374 strongly for the W, WSH, and WSHN arms, moderately for the S arm, and little for the H and N arms.
375 For example, in the W arm, increasing community coverage to 50% resulted in a median increase of 80%
376 points for samples with the highest intervenable fractions but only 23 percentage points for samples with
377 the lowest intervenable fractions. In contrast, in the H arm, increasing community coverage to 50%
378 resulted in a median increase of 31 percentage points for samples with the highest intervenable fractions
379 compared only 22 percentage points for samples with the lowest intervenable fractions. (Note that the fact
380 that intervenable fraction was relevant for the N arm at all is a result of the correlations between the
381 intervenable fraction and the other parameters in the original parameter sets.)

382 **Discussion**

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

395 The importance of ensuring high community WASH coverage to address health outcomes has been
396 highlighted in multiple context, including latrines,^{29,36} bed nets,³⁷ and chemotherapy for helminths,³⁸
397 among others. Further work is needed to improve our measures of indirect and direct intervention
398 effects^{29,39} to better determine sanitation targets. Our findings support the call for systems-level WASH

399 provisioning and improved universal access, underscoring the fundamental push to achieve the 2030
400 sustainable development targets.⁴⁰

401 Our finding that the intervenable fraction (completeness) was an important effect modifier, emphasizes
402 the need to better understand the sources of exposure not impacted by traditional WASH interventions.
403 For example, contamination of food outside the home or from flies or exposure to feces from animals
404 living near or inside the home may not be reduced by water quality or latrine interventions.^{41–43} Capturing
405 and reducing transmission through additional targeted interventions would increase the fraction of
406 transmission intervened on and thereby make increased community coverage even more effective.

407 Low diarrheal prevalence makes it more difficult to observe a statistically significant reduction in
408 diarrhea.¹⁴ However, from a mechanistic perspective, we found that intervention effectiveness would have
409 been lower had the background disease pressure in the community been higher (i.e., higher baseline
410 disease prevalence) because individual-level interventions can be overwhelmed by higher disease
411 pressure from the community, including those not covered by the intervention. This finding is supported
412 by the outcomes of WASH-B Kenya trial, which had higher disease prevalence (27% in the control arm)
413 and no significant intervention effects on diarrheal prevalence¹⁵ and is consistent with previous literature
414 that has shown that non-pharmaceutical interventions are more effective for less transmissible pathogens
415 or when the population has a higher degree of population immunity.⁴⁴ This is not to say that individual
416 improvements would have no effect but that the effects are blunted if disease pressure in the rest of the
417 community were not also addressed.

418 Similarly, many have suggested that when preexisting WASH conditions are relatively high, interventions
419 do not provide a substantial improvement in efficacy and thus health outcomes.^{18,19,22,45} However, from
420 the transmission system perspective reflected by our results, if the preexisting WASH conditions
421 (particularly among those *not* covered by the intervention) were poorer, the community disease pressure
422 would be greater, and it would be more difficult to protect study participants from infection, even if the
423 people covered by the intervention had a greater improvement in protection.

424 Because enteric pathogens can exploit multiple transmission pathways, many studies have tried to
425 determine whether combined WASH interventions (WSH) are more effective than single interventions
426 (W, S, or H).^{32,46} Whether or not there is an additional effect of combined interventions depends on
427 whether the interventions are complementary, that is, whether they each block some of the transmission
428 that the other interventions would not have blocked.⁴⁷ This complementarity is an assumption in our
429 transmission model framework (as each intervention affects different parts of the disease system), and
430 because the model can fit the data, we find that complementarity is consistent with the observed trial

431 results.³³ Other modeling and empirical studies, support that WASH interventions can complement each
432 other, or even potentially be synergistic.^{48,49} In this work, we found that the combined interventions could
433 have a greater effect than the individual interventions, but that the effects were generally sub-additive,
434 meaning that the effectiveness of the combined WSH intervention was less than the sum of its parts
435 (Table 3). Combined interventions offer a substantially better chance of disease elimination, especially at
436 higher coverage levels (Figure 5).

437 One challenge that WASH RCTs often face is achieving high compliance through both high fidelity
438 (providing the interventions as planned) and high adherence of participants to the use of the intervention.
439 In WASH-B Bangladesh, the intervention compliance, as defined by trial investigators, was high,
440 generally above 90%.^{14,34} Accordingly, our full compliance counterfactual was limited in the impact it
441 could detect.

442 The strength of our approach is underscored by the rich and high-quality data collected by the WASH-B
443 Bangladesh trial (and other RCTs) and in our transmission model framework capturing relative disease
444 prevalence. RCTs provide the gold standard of evidence about intervention effectiveness in a specific
445 context, and our approach allows us to generalize RCT results to other contexts, providing a tool for
446 powerful policy and programmatic guidance. The SISE-RCT model can be customized for local contexts
447 and interventions and then used to support local decision-making (e.g., to determine whether to invest in
448 community coverage vs intervention efficacy). Future work may also develop recommendations for
449 achieving elimination while minimizing costs. One limitation of our study is the high uncertainty in many
450 of the model parameters, especially the intervenable fraction, which propagates into the counterfactual
451 scenarios. These uncertainties stem from potential trade-offs in the model, e.g., a low intervenable
452 fraction and a low intervention efficacy may have similar effects. Fortunately, our framework has the
453 potential to incorporate additional information about parameters like the intervenable fraction and
454 efficacy through our Bayesian sampling-importance resampling approach, allowing us to tailor
455 projections of intervention effectiveness to specific parameter regions based on additional information
456 (e.g., chlorination efficacy above 75%). One limitation of the data was the inability to distinguish whether
457 non-target children were members of the same household as the target child or not, which introduced
458 misspecification into our classification of W and H exposures, likely attenuating the efficacy estimates for
459 those interventions. Also, we accounted for changes in disease pressure between, but not within, survey
460 periods; future work may more directly address seasonal changes in disease pressure and even pathway
461 strength, as a function of precipitation, seasonal flooding, etc. Another limitation of this study is that our
462 results do not directly address some aspects of the Sustainable Development Goal (SDG) Target 6.2.⁴⁰ For
463 example, the sanitation arm did not move households from no or basic sanitation to improved sanitation

464 (as defined by the Joint Monitoring Programme). So, the “sanitation” intervention outcomes we estimated
465 may not directly correspond to the policy-relevant changes required to meet SDG target. Likewise, the
466 “water” intervention focused on water quality improvements (chlorination) but not water quantity. None
467 of these issues are limitations of our modeling framework; rather, they are limitations of our specific
468 application. Applying our methods across other trial datasets could address these limitations by allowing
469 for modeling of other—and perhaps more policy-relevant—WASH exposure parameters.

470 Our work contributes to the robust discussion^{18–24,50} about the future directions of WASH research and
471 programming, and our modeling approach is well-suited to reevaluating current evidence during the
472 “pause for reflection” recommended by a consensus of WASH reserachers.¹⁹ This consensus group said
473 that “the lesson perhaps lies in not seeking to attribute benefits to individual WASH factors but in that the
474 public health dividends are paid when comprehensive services are in place.” Our work underscores this
475 conclusion, not only by emphasizing the importance of coverage and completeness of interventions, but
476 also in its rejection of the hypotheses that greater effectiveness might be found in areas with greater
477 disease prevalence or lower preexisting WASH infrastructure. Indeed, our findings suggest that the effect
478 of individual-level WASH improvements will be blunted the further the community is from achieving
479 herd protection. Accordingly, this analysis provides further evidence supporting community-level
480 interventions seeking to achieve herd protection through high community coverage.

481

482 **Contributors**

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

490 **Acknowledgements**

491 This work was funded by the Bill & Melinda Gates Foundation (grant INV-005081) and the National
492 Science Foundation (grant DMS-1853032). The original WASH-B Bangladesh trial was also funded by
493 the Bill & Melinda Gates Foundation (grant OPPGD759). ANMK’s contributions were directly funded by
494 the Bill & Melinda Gates Foundation and not as part of the foundation grant to the authors. ANMK is an

495 employee of the Bill & Melinda Gates Foundation; however, this study does not necessarily represent the
496 views of the Bill & Melinda Gates Foundation.

497 **Data sharing**

498 The WASH Benefits Bangladesh data is publicly available at <https://osf.io/tprw2/>. The data and code
499 underlying the results of this paper are available at <https://doi.org/10.5281/zenodo.10950560>. Data
500 underlying each of the figures is given in the Excel spreadsheet supplement.

501 **References**

- 502 1. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and
503 territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*.
504 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9
- 505 2. Troeger C, Blacker B, Khalil IA, et al. Estimates of the global, regional, and national morbidity,
506 mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a
507 systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis*.
508 2018;18(11):1191-1210. doi:10.1016/S1473-3099(18)30310-4
- 509 3. Wolf J, Johnston RB, Ambelu A, et al. Burden of disease attributable to unsafe drinking water,
510 sanitation, and hygiene in domestic settings: a global analysis for selected adverse health
511 outcomes. *Lancet*. 2023;401(10393):2060-2071. doi:10.1016/S0140-6736(23)00458-0
- 512 4. Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet*.
513 2009;374(9694):1032-1035. doi:10.1016/S0140-6736(09)60950-8
- 514 5. Julian TR. Environmental transmission of diarrheal pathogens in low and middle income
515 countries. *Environ Sci Process Impacts*. 2016;18(8):944-955. doi:10.1039/C6EM00222F
- 516 6. Wagner EG, Lanoix JN. Excreta disposal for rural areas and small communities. *Monogr Ser*
517 *World Health Organ*. 1958;39:1-182. <http://www.ncbi.nlm.nih.gov/pubmed/13581743>.
- 518 7. Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in
519 infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS):
520 A prospective, case-control study. *Lancet*. 2013;382(9888):209-222. doi:10.1016/S0140-
521 6736(13)60844-2
- 522 8. Liu J, Platts-Mills JA, Juma J, et al. Use of quantitative molecular diagnostic methods to identify
523 causes of diarrhoea in children: a reanalysis of the GEMS case-control study. *Lancet*.
524 2016;388(10051):1291-1301. doi:10.1016/S0140-6736(16)31529-X
- 525 9. Watson SI, Rego RTT, Hofer T, Lilford RJ. Evaluations of water, sanitation and hygiene
526 interventions should not use diarrhoea as (primary) outcome. *BMJ Glob Heal*. 2022;7(5):e008521.
527 doi:10.1136/bmjgh-2022-008521
- 528 10. Wolf J, Hubbard S, Brauer M, et al. Effectiveness of interventions to improve drinking water,
529 sanitation, and handwashing with soap on risk of diarrhoeal disease in children in low-income and
530 middle-income settings: a systematic review and meta-analysis. *Lancet*. 2022;400(10345):48-59.

- 531 doi:10.1016/S0140-6736(22)00937-0
- 532 11. Clasen T, Boisson S, Routray P, et al. Effectiveness of a rural sanitation programme on diarrhoea,
533 soil-transmitted helminth infection, and child malnutrition in Odisha, India: A cluster-randomised
534 trial. *Lancet Glob Heal*. 2014;2(11):e645-e653. doi:10.1016/S2214-109X(14)70307-9
- 535 12. Patil SR, Arnold BF, Salvatore AL, et al. The effect of India's total sanitation campaign on
536 defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled
537 trial. *PLoS Med*. 2015;11(8). doi:10.1371/journal.pmed.1001709
- 538 13. Pickering AJ, Djebbari H, Lopez C, Coulibaly M, Alzua ML. Effect of a community-led sanitation
539 intervention on child diarrhoea and child growth in rural Mali: A cluster-randomised controlled
540 trial. *Lancet Glob Heal*. 2015;3(11):e701-e711. doi:10.1016/S2214-109X(15)00144-8
- 541 14. Luby SP, Rahman M, Arnold BF, et al. Effects of water quality, sanitation, handwashing, and
542 nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised
543 controlled trial. *Lancet Glob Heal*. 2018;6(3):e302-e315. doi:10.1016/S2214-109X(17)30490-4
- 544 15. Null C, Stewart CP, Pickering AJ, et al. Effects of water quality, sanitation, handwashing, and
545 nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised
546 controlled trial. *Lancet Glob Heal*. 2018;6(3):e316-e329. doi:10.1016/S2214-109X(18)30005-6
- 547 16. Rogawski McQuade ET, Platts-Mills JA, Gratz J, et al. Impact of Water Quality, Sanitation,
548 Handwashing, and Nutritional Interventions on Enteric Infections in Rural Zimbabwe: The
549 Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial. *J Infect Dis*. 2020;221(8):1379-1386.
550 doi:10.1093/infdis/jiz179
- 551 17. Knee J, Sumner T, Adriano Z, et al. Effects of an urban sanitation intervention on childhood
552 enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-after trial. *Elife*.
553 2021;10. doi:10.7554/eLife.62278
- 554 18. Cumming O, Curtis V. Implications of WASH Benefits trials for water and sanitation. *Lancet*
555 *Glob Heal*. 2018;6(6):e613-e614. doi:10.1016/S2214-109X(18)30192-X
- 556 19. Cumming O, Arnold BF, Ban R, et al. The implications of three major new trials for the effect of
557 water, sanitation and hygiene on childhood diarrhea and stunting: A consensus statement. *BMC*
558 *Med*. 2019;17(1):1-9. doi:10.1186/s12916-019-1410-x
- 559 20. Levy K, Eisenberg JNS. Moving towards transformational WASH. *Lancet Glob Heal*.
560 2019;7(11):e1492. doi:10.1016/S2214-109X(19)30396-1

- 561 21. Pickering AJ, Null C, Winch PJ, et al. The WASH Benefits and SHINE trials: interpretation of
562 WASH intervention effects on linear growth and diarrhoea. *Lancet Glob Heal*. 2019;7(8):e1139-
563 e1146. doi:10.1016/S2214-109X(19)30268-2
- 564 22. Coffey D, Spears D. Implications of WASH Benefits trials for water and sanitation. *Lancet Glob*
565 *Heal*. 2018;6(6):e615. doi:10.1016/S2214-109X(18)30225-0
- 566 23. Haque SS, Freeman MC. The Applications of Implementation Science in Water, Sanitation, and
567 Hygiene (WASH) Research and Practice. *Environ Health Perspect*. 2021;129(6):65002.
568 doi:10.1289/EHP7762
- 569 24. Amebelu A, Ban R, Bhagwan J, et al. The Lancet Commission on water, sanitation and hygiene,
570 and health. *Lancet*. 2021;398(10310):1469-1470. doi:10.1016/S0140-6736(21)02005-5
- 571 25. Li X, Mukandavire C, Cucunubá ZM, et al. Estimating the health impact of vaccination against ten
572 pathogens in 98 low-income and middle-income countries from 2000 to 2030: a modelling study.
573 *Lancet*. 2021;397(10272):398-408. doi:10.1016/S0140-6736(20)32657-X
- 574 26. World Health Organization (WHO). *Impact of the COVID-19 Pandemic on Seven Neglected*
575 *Tropical Diseases: A Model-Based Analysis*. Geneva; 2021.
576 <https://apps.who.int/iris/handle/10665/343993>.
- 577 27. Lofgren E, Halloran ME, Rivers CM, et al. Mathematical models: A key tool for outbreak
578 response. *Proc Natl Acad Sci*. 2015;112(2):E234-E234. doi:10.1073/pnas.1423846112
- 579 28. Coffeng LE, Vaz Nery S, Gray DJ, Bakker R, de Vlas SJ, Clements ACAA. Predicted short and
580 long-term impact of deworming and water, hygiene, and sanitation on transmission of soil-
581 transmitted helminths. Walker M, ed. *PLoS Negl Trop Dis*. 2018;12(12):1-17.
582 doi:10.1371/journal.pntd.0006758
- 583 29. Fuller JA, Eisenberg JNS. Herd protection from drinking water, sanitation, and hygiene
584 interventions. *Am J Trop Med Hyg*. 2016;95(5):1201-1210. doi:10.4269/ajtmh.15-0677
- 585 30. Brouwer AF, Eisenberg MC, Bakker KM, et al. Leveraging infectious disease models to interpret
586 randomized controlled trials: Controlling enteric pathogen transmission through water, sanitation,
587 and hygiene interventions. Lau EH, ed. *PLOS Comput Biol*. 2022;18(12):e1010748.
588 doi:10.1371/journal.pcbi.1010748
- 589 31. Brouwer AF, Kraay ANM, Zahid MH, Eisenberg MC, Freeman MC, Eisenberg JNS. A
590 mechanistic modeling approach to assessing the sensitivity of outcomes of water, sanitation, and

- 591 hygiene interventions to local contexts and intervention factors. *medRxiv*. 2024.
592 doi:10.1101/2024.03.09.24304020
- 593 32. Arnold BF, Null C, Luby SP, et al. Cluster-randomised controlled trials of individual and
594 combined water, sanitation, hygiene and nutritional interventions in rural bangladesh and Kenya:
595 The WASH benefits study design and rationale. *BMJ Open*. 2013;3(8):1-17.
596 doi:10.1136/bmjopen-2013-003476
- 597 33. Brouwer AF, Eisenberg MC, Bakker KM, et al. Leveraging infectious disease models to interpret
598 randomized controlled trials: controlling enteric pathogen transmission through water, sanitation,
599 and hygiene interventions. *medRxiv*. January 2022:2022.04.28.22274441.
600 doi:10.1101/2022.04.28.22274441
- 601 34. Ercumen A, Pickering AJ, Kwong LH, et al. Do Sanitation Improvements Reduce Fecal
602 Contamination of Water, Hands, Food, Soil, and Flies? Evidence from a Cluster-Randomized
603 Controlled Trial in Rural Bangladesh. *Environ Sci Technol*. 2018;52(21):12089-12097.
604 doi:10.1021/acs.est.8b02988
- 605 35. Garn J V., Sclar GD, Freeman MC, et al. The impact of sanitation interventions on latrine
606 coverage and latrine use: A systematic review and meta-analysis. *Int J Hyg Environ Health*.
607 2017;220(2):329-340. doi:10.1016/j.ijheh.2016.10.001
- 608 36. Harris M, Alzua ML, Osbert N, Pickering A. Community-Level Sanitation Coverage More
609 Strongly Associated with Child Growth and Household Drinking Water Quality than Access to a
610 Private Toilet in Rural Mali. *Environ Sci Technol*. 2017;51(12):7219-7227.
611 doi:10.1021/acs.est.7b00178
- 612 37. Teklehaimanot A, Sachs JD, Curtis C. Malaria control needs mass distribution of insecticidal
613 bednets. *Lancet*. 2007;369(9580):2143-2146. doi:10.1016/S0140-6736(07)60951-9
- 614 38. Anderson R, Truscott J, Hollingsworth TD. The coverage and frequency of mass drug
615 administration required to eliminate persistent transmission of soil-transmitted helminths. *Philos*
616 *Trans R Soc B*. 2014;369(1645):20130435. doi:10.1098/rstb.2013.0435
- 617 39. Halloran ME, Auranen K, Baird S, et al. Simulations for designing and interpreting intervention
618 trials in infectious diseases. *BMC Med*. 2017;15(1):1-8. doi:10.1186/s12916-017-0985-3
- 619 40. World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). *Progress*
620 *on Household Drinking Water, Sanitation and Hygiene 2000-2020: Five Years into the SDGs*.

- 621 Geneva; 2021.
- 622 41. Penakalapati G, Swarthout J, Delahoy MJ, et al. Exposure to Animal Feces and Human Health: A
623 Systematic Review and Proposed Research Priorities. *Environ Sci Technol*. 2017;51(20):11537-
624 11552. doi:10.1021/acs.est.7b02811
- 625 42. Khamesipour F, Lankarani KB, Honarvar B, Kwenti TE. A systematic review of human pathogens
626 carried by the housefly (*Musca domestica* L.). *BMC Public Health*. 2018;18(1):1049.
627 doi:10.1186/s12889-018-5934-3
- 628 43. DuPont HL. The Growing Threat of Foodborne Bacterial Enteropathogens of Animal Origin. *Clin*
629 *Infect Dis*. 2007;45(10):1353-1361. doi:10.1086/522662
- 630 44. Baker RE, Saad-Roy CM, Park SW, Farrar J, Metcalf CJE, Grenfell BT. Long-term benefits of
631 nonpharmaceutical interventions for endemic infections are shaped by respiratory pathogen
632 dynamics. *Proc Natl Acad Sci*. 2022;119(49). doi:10.1073/pnas.2208895119
- 633 45. Arnold BF, Null C, Luby SP, Colford JM. Implications of WASH Benefits trials for water and
634 sanitation – Authors’ reply. *Lancet Glob Heal*. 2018;6(6):e616-e617. doi:10.1016/S2214-
635 109X(18)30229-8
- 636 46. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM. Water, sanitation, and
637 hygiene interventions to reduce diarrhoea in less developed countries: A systematic review and
638 meta-analysis. *Lancet Infect Dis*. 2005;5(1):42-52. doi:10.1016/S1473-3099(04)01253-8
- 639 47. Eisenberg JNS, Scott JC, Porco T. Integrating disease control strategies: Balancing water
640 sanitation and hygiene interventions to reduce diarrheal disease burden. *Am J Public Health*.
641 2007;97(5):846-852. doi:10.2105/AJPH.2006.086207
- 642 48. Eisenberg JNS, Trostle J, Sorensen RJD, Shields KF. Toward a Systems Approach to Enteric
643 Pathogen Transmission: From Individual Independence to Community Interdependence. *Annu Rev*
644 *Public Health*. 2012;33(1):239-257. doi:10.1146/annurev-publhealth-031811-124530
- 645 49. Esrey SA. Water, waste, and well-being: A multicountry study. *Am J Epidemiol*. 1996;143(6):608-
646 623. doi:10.1093/oxfordjournals.aje.a008791
- 647 50. D’Mello-Guyett L, Yates T, Bastable A, et al. Setting priorities for humanitarian water, sanitation
648 and hygiene research: A meeting report. *Confl Health*. 2018;12(1):10-14. doi:10.1186/s13031-
649 018-0159-8

651 **Tables**

652 **Table 1: Parameters of the SISE-RCT model.** The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission
 653 through environmental (E) compartments and simulated to steady state to approximate an RCT. The median value column denotes the median values for the
 654 multi-intervention model. Parameters ρ_0 and ρ do not have median values because they are determined by the data.

Parameter	Definition	Median value
ρ_0	Baseline WASH conditions (fraction of individuals in the community with intervention-level WASH infrastructure)	—
ρ	Compliance (fraction of individuals in intervention arm using intervention)	—
R_0	Transmission potential (basic reproduction number)	1.10
π_c^*	Baseline disease prevalence (note: this technically a model output, but we treat it as a parameter because it has a 1-to-1 correspondence with R_0 given the other parameters)	6.9%
$R_{0,w}/R_0$	Fraction of transmission along the water pathway	
$R_{0,f}/R_0$	Fraction of transmission along the fomite & hands pathway	
$R_{0,f}/R_0$	Fraction of transmission on pathways not intervened on (i.e., 1 minus the intervenable fraction)	0.48
$1 - \varphi_\alpha$	Intervention efficacy for reducing shedding (S=sanitation intervention)	0.23 (S),
$1 - \varphi_\beta$	Intervention efficacy for reducing transmission (W= water, H=hygiene, N=nutrition interventions)	0.44 (W), 0.33 (H), 0.16 (N)
ω	Community coverage fraction (fraction of community included in the intervention trial)	5.4%

655

656

657

658

659

660

661

662 **Table 2: WASH-B Bangladesh counterfactual scenarios and implementations.** WASH = water, sanitation, & hygiene. *indicates parameters in the
 663 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. (Setting ρ_0^* to the vector $(1, 0, \dots, 0)$).
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. (Adjusting ρ^* appropriately, e.g., moving the fraction of the population from S to HS in the hygiene arm).
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%. ($\omega^* = 0.2, \dots, 1.0$)

664

665

666

667

668 **Table 3. Median intervention effectiveness and median percent point change in intervention effectiveness in each intervention arm for each**669 **counterfactual scenario compared to the original scenario.** Intervention effectiveness (ϵ) in intervention effectiveness is 1 minus the relative risk of diarrhea670 in the intervention arm vs the control arm in each scenario, expressed as a percentage. The column $\Delta\epsilon$ gives the median change in intervention effectiveness in671 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	$\Delta\epsilon$	ϵ	$\Delta\epsilon$	ϵ	$\Delta\epsilon$	ϵ	$\Delta\epsilon$	ϵ	$\Delta\epsilon$	ϵ	$\Delta\epsilon$
Original scenario	8.1%	—	36.3%	—	33.0%	—	30.2%	—	33.9%	—	34.5%	—
No WASH baseline conditions	8.0%	-0.2%	21.7%	-14.6%	24.6%	-8.6%	20.6%	-9.5%	26.5%	-7.3%	25.7%	-8.4%
Double baseline disease prevalence	6.8%	-1.3%	14.7%	-21.6%	15.7%	-17.5%	18.6%	-11.9%	17.7%	-16.0%	24.1%	-10.3%
Full compliance	11.6%	+3.6%	36.4%	+0.1%	34.2%	+1.1%	34.7%	+4.4%	34.7%	+0.9%	38.6%	+4.1%
Half of other pathway transmission can be intervened on	14.2%	+6.0%	48.4%	+11.8%	43.7%	+10.6%	50.8%	+20.1%	37.8%	+3.8%	53.0%	+18.2%
Double efficacy of chlorination	15.4%	+7.1%	—	—	—	—	36.4%	+6.1%	—	—	40.6%	+5.9%
Double efficacy of latrine water seal	—	—	55.4%	+18.9%	—	—	46.0%	+15.7%	—	—	48.9%	+14.2%
Double efficacy of handwashing	—	—	—	—	54.6%	+21.3%	50.2%	+19.8%	—	—	52.8%	+18.3%
Double efficacy of nutrition	—	—	—	—	—	—	—	—	45.3%	+11.3%	44.3%	+10.1%
Increase community coverage to 20%	22.9%	+14.6%	43.3%	+6.8%	41.4%	+8.3%	63.8%	+34.0%	49.2%	+15.3%	79.6%	+45.5%

672

673

674

675

676

677

678 **Figure captions**

679 **Figure 1: Single-intervention SISE-RCT model diagram with one attenuated exposure population and a regular exposure population interacting**
680 **through shared environments.** The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through
681 environmental (E) compartments and simulated to steady state to approximate an RCT. The black lines denote infection and recovery, the blue lines denote
682 shedding from infectious individuals into environmental compartments, the grey lines denote pick-up of pathogens from the environment by susceptible
683 individuals, and the orange lines denote environmental pathogen decay. S_i and I_i denote susceptible and infectious fraction of the i th attenuated exposure
684 population, and S_0 and I_0 denote susceptible and infectious fraction of the regular exposure population. E_w , E_f , and E_o denote environmental pathways for water,
685 fomites & hands, and all other pathways. This figure was adapted from Figure 2 of Brouwer et al (2022).³⁰

686 **Figure 2: Intervention effectiveness as a function of WASH intervention factors.** The SISE-RCT model is a compartmental susceptible-infectious-
687 susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. The model was
688 simulated at the median values of the model parameters when fit to the WASH Benefits Bangladesh trial and across ranges of counterfactual values of six
689 contextual and intervention WASH factors. The six WASH factors are A) preexisting WASH conditions (fraction of individuals not enrolled in the intervention
690 arm that are using preexisting WASH infrastructure), B) baseline disease prevalence (a function of the basic reproduction number R_0), C) compliance (fraction of
691 individuals enrolled in the intervention arm that are using the intervention), D) intervenable fraction of transmission (how much of the transmission could be
692 prevented in a perfect intervention), E) the community coverage fraction (fraction of the population enrolled in the trial), and F-I) the intervention efficacy
693 (fraction reduction in transmission or shedding when using the intervention) of each intervention. The underlying data are provided in Excel Table S1. WASH =
694 water, sanitation, & hygiene.

695 **Figure 3: Percentage point change in intervention effectiveness compared to the original scenario in each counterfactual scenario.** The SISE-RCT model
696 is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E) compartments and simulated to steady state to
697 approximate an RCT. Here, we applied it to data from the WASH-B Bangladesh trial, selecting 50,000 parameter sets consistent with the trial outcomes. We
698 simulated each parameter set under each counterfactual scenario (Table 2). The violin plots give the distribution of values across the 50,000 simulations, with
699 median points. The underlying data are provided in Excel Table S2A-H. W = water, S = sanitation, H = hygiene, N = nutrition.

700 **Figure 4: Percentage point change in intervention effectiveness compared to the original scenario in the 20% coverage counterfactual scenario (a)**
701 **overall and (b) considering other parameters as potential effect modifiers (b-h).** The SISE-RCT model is a compartmental susceptible-infectious-susceptible
702 (SIS) model with transmission through environmental (E) compartments and simulated to steady state to approximate an RCT. Here, we applied it to data from

703 the WASH-B Bangladesh trial, selecting 25,000 parameter sets consistent with the trial outcomes. We simulated each parameter set under the 20% coverage
704 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
705 plots give the distributions of the intervention effectiveness across quintiles of the listed potential effect modifier. The underlying data are provided in Excel
706 Table S3A-H. W = water, S = sanitation, H = hygiene, N = nutrition.

707 **Figure 5: Percentage point change in intervention effectiveness compared to the original scenario in each arm for the lowest and highest quintiles of**
708 **intervenable fraction.** The SISE-RCT model is a compartmental susceptible-infectious-susceptible (SIS) model with transmission through environmental (E)
709 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
710 sets consistent with the trial outcomes. We simulated each parameter set for coverage counterfactual scenarios ranging from 20% to 100% (Table 2). The violin
711 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
712 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
713 = water, S = sanitation, H = hygiene, N = nutrition.

Attenuated exposure population i









