Accepted for Publication, Published online September 6, 2016; doi:10.4269/ajtmh.15-0677.
The latest version is at http://ajtmh.org/cgi/doi/10.4269/ajtmh.15-0677

Herd Protection from Drinking Water, Sanitation, and Hygiene Interventions

James A. Fuller[1] and Joseph N. S. Eisenberg[1]*

[1]*Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan*

* Address correspondence to Joseph N. S. Eisenberg, Department of Epidemiology, University of Michigan School of Public Health, 1415 Washington Heights, Ann Arbor, MI 48109. E-mail: jnse@umich.edu

Abstract.

Herd immunity arises when a communicable disease is less able to propagate because a substantial portion of the population is immune. Nonimmunizing interventions, such as insecticide-treated bednets and deworming drugs, have shown similar herd-protective effects. Less is known about the herd protection from drinking water, sanitation, and hand hygiene (WASH) interventions. We first constructed a transmission model to illustrate mechanisms through which different WASH interventions may provide herd protection. We then conducted an extensive review of the literature to assess the validity of the model results and identify current gaps in research. The model suggests that herd protection accounts for a substantial portion of the total protection provided by WASH interventions. However, both the literature and the model suggest that sanitation interventions in particular are the most likely to provide herd protection, since they reduce environmental contamination. Many studies fail to account for these indirect effects and thus underestimate the total impact an intervention may have. Although cluster-randomized trials of WASH interventions have reported the total or overall efficacy of WASH interventions, they have not quantified the role of herd protection. Just as it does in immunization policy, understanding the role of herd protection from WASH interventions can help inform coverage targets and strategies that indirectly protect those that are unable to be reached by WASH campaigns. Toward this end, studies are needed to confirm the differential role that herd protection plays across the WASH interventions suggested by our transmission model.

INTRODUCTION

Enteric pathogens are a major source of disease burden worldwide leading to diarrhea, subclinical environmental enteropathy, malnutrition, and death.[1] These pathogens are largely transmitted via the fecal-oral route with a variety of environmental intermediaries, such as drinking water, soil, food, fomites, and hands. Drinking water, sanitation, and hand hygiene (WASH) interventions have been shown to be protective against enteric infections,[1] but little attention is given to the herd-protective effects that they may provide. Herd protection occurs when an infectious disease intervention provides indirect protection to nonrecipients. Studies that fail to account for herd protection will lead to an underestimate of the total protective effectiveness of the intervention.

Halloran and Struchiner[2] provided a useful framework for conducting studies to measure these indirect effects (Figure 1). These study designs require different populations with varying levels of intervention coverage. The direct effect is the protective efficacy of the intervention, and is measured by comparing intervention groups within population A, as is done in a simple randomized controlled trial where individuals are the unit of randomization. The direct effect can be estimated at each level of intervention coverage by the equation $1–D_1/D_0$, where D_1 represents the disease risk in the intervention group and D_0 represents the disease risk in the

nonintervention or control group in the same population. The indirect effect represents the herd protection provided by the intervention and is measured through cross-population studies. It is estimated at each level of coverage by the equation $1-D_0/D_{0*}$, where D_{0*} represents the disease risk in a population where the intervention is entirely absent (coverage = 0%). The total effect is the combination of the direct and indirect effects. The overall effect is the risk in population A (the weighted average of D_1 and D_0) compared with the risk in population B (D_{0*}). The majority of cluster-randomized trials measure either the overall effect or the total effect but without disaggregating the direct and indirect effects.[3]

The framework of Halloran and Struchiner[2] has been used to show evidence of herd protection for a variety of infectious disease interventions, such as vaccines[4,5] and insecticide-treated bednets to prevent malaria,[6–8] mass deworming drugs to treat helminth infection,[9,10] highly active retroviral therapy to treat human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS),[11,12] and antibiotics to treat trachoma.[13,14] A few studies[15–18] have shown that WASH technologies may have community-wide benefits, the practices of one household having a beneficial effect on neighbors. Use of this framework, however, assumes comparability (no confounding) between populations with different levels of intervention, which can be difficult to achieve or prove. Although a few of these studies were rigorously done, many others have serious methodological limitations, such as failure to adjust for this group-level confounding or failure to account for the individual level of the exposure in question.

To encourage more empirical research on the indirect effects of WASH interventions, we present a stochastic mathematical model of enteric pathogen transmission in a community. This model serves to illustrate the different mechanisms through which WASH interventions can provide community-wide protective effects, that is, herd protection. We then compare the findings of our model to the existing literature to highlight current gaps in empirical research.

METHODS

Model structure.

We simulated a community of 500 individuals, nested within 100 households. Individuals are modeled as discrete entities using a stochastic framework. The quantity of pathogens in the environment is modeled as continuous using ordinary differential equations.[19] Individuals are categorized as susceptible, infectious, or immune, and immunity is assumed to be permanent (SIR model). Many enteric pathogens do not provide permanent immunity, but during short-term outbreaks, such as those simulated in this model, the difference between permanent and temporary immunity is insignificant. All transmission of pathogens occurs via the environment.[20] Infectious individuals can transmit pathogens to susceptible individuals by either of two pathways (Figure 2). First, the infectious individual sheds pathogens into their household environment at rate δ. Susceptible individuals pick up pathogens from their household environment at rate ρ. This household environment represents household surfaces, stored drinking water, or any other pathogen-harboring area located within the household. Second, infectious individuals shed pathogens in the community environment at rate ϕ. All susceptible individuals in the community pick up pathogens from the community environment at rate α. The community environment can represent an unprotected source of drinking water such as a pond, common or shared areas such as schools, or any other pathogen-harboring area accessed by people from multiple households. Pathogen survival in both the household and community

environment is determined by the parameter μ. The model was coded in R version 3.0.2.[21] For details, see the Supplemental Appendix A.

Simulation analysis.

Each simulation begins with a population that is entirely susceptible, except for one infectious individual; this is representative of a new pathogen strain being introduced into a community with no prior protection. We then simulate an epidemic. The primary outcome of interest is the cumulative incidence, defined as the proportion infected or immune at the end of the epidemic.

To estimate the effect of intervention coverage, we simulate these epidemics at varying levels of intervention coverage in the community. In the first scenario, no households in the community are using the intervention, that is, everyone is in the control group. In subsequent scenarios, we increase the percentage of households using the intervention by increments of 10%, until coverage reaches 100%. At each level of coverage, the model is simulated 100 times, for a total of 1,100 simulations. The median cumulative incidence at each level of coverage is used to calculate the protective efficacy of the intervention.

To estimate the amount of herd protection, the direct, indirect, total, and overall effects of the intervention are measured using the framework presented in Figure 1.[2] The direct effect is the protective efficacy of the intervention that a simple randomized controlled trial would measure. It is estimated at each level of intervention coverage by the equation $1-D_1/D_0$, where D_1 represents the disease risk in the intervention group at a given level of coverage and D_0 represents the disease risk in the control group (those not receiving the intervention) at that same level of coverage. The indirect effect represents the herd protection provided by the intervention. It is estimated at each level of coverage by the equation $1-D_0/D_{0*}$, where D_{0*} represents the disease risk in a population where the intervention is entirely absent (coverage = 0%). The total effect is the combination of the direct and indirect effects, and is estimated at each level of coverage by the equation $1-D_1/D_{0*}$. Lastly, the overall effect is the percentage reduction in risk in all of population A due to the intervention, and it is estimated at each level of coverage by the equation $1-(w_0D_0 + w_1D_1)/D_{0*}$, where w_0 and w_1 are weights equal to the proportion of individuals in population A that are in the control and intervention groups, respectively.

WASH interventions.

We separately model a drinking water intervention, a sanitation intervention, and a hand hygiene intervention. Each of these interventions is applied at the household level with no within-household heterogeneity and compliance is assumed to be 100%. We conceptualize these interventions in our model as resulting in a reduction of the rate of pathogen shedding into the environment and/or a reduction of the rate of pathogen pickup from the environment. Specifically, the drinking water intervention reduces the value of α, the rate of pathogen pickup from the community environment. The sanitation intervention reduces the value of ϕ, the rate of shedding pathogens into community environment. Finally, the hand hygiene intervention reduces the value of δ and ρ, the rate of shedding into the household environment and the rate of pathogen pickup from the household environment, respectively. These interventions are shown in Figure 2 as grey boxes, suggesting that they partially block these pathways.

Parameter values for each of these interventions along with the baseline practices of the control group are shown in Table 1. The goal of this model is to provide a conceptual framework for herd protection and not to estimate the magnitude of direct or indirect effects. Therefore, parameter values were chosen to generate a baseline risk of approximately 70% and a total effect ($1-D_1/D_{0*}$ at 100% coverage) of approximately 35% based on estimates from systematic reviews.[22–32]

Literature review.

To identify studies that investigated herd protection from WASH interventions, we searched PubMed and Google Scholar for articles using the following search criteria: "herd protection," "herd effect," "indirect effect," "mass effect," "community effect," "externalities," or "neighborhood," as well as key words for each intervention: "sanitation," "toilet," or "latrine," "water" and "hygiene," or "handwashing." After selecting relevant articles from these search results, we also considered articles in their citation network (those articles cited by or citing the original article). In our literature review, we present only those studies that included drinking water, sanitation, or hand hygiene coverage as an explanatory variable for a health outcome.

RESULTS

Drinking water.

When the entire community is among the control group, the median risk of infection in the community across 100 simulations is 69.6%. As coverage of the drinking water intervention increases, the risk in both the intervention group and the control group decreases (Figure 3, Panel A). When coverage reaches 100%, the median risk in the community is 44.5%, providing a total protective efficacy of 36.1% (0.361 = 1 – 0.445/0.696). At all levels of coverage, the risk in the intervention group is lower than that of the control group (Figure 3, Panel A). The direct protective efficacy of the drinking water intervention is about 15%, and is relatively constant across all levels of coverage (Figure 3, Panel B). As coverage increases, however, the indirect effect becomes more pronounced and approaches 20% at high levels of coverage.

Several studies have attempted to show evidence of herd protection from drinking water (Table 2 and Supplemental Table 1). All of these studies used cross-sectional national household surveys such as the Demographic and Health Surveys (dhsprogram.org) and typically used the percentage of households in a survey cluster with an improved source of drinking water as an independent variable. Six studies from three publications[17,33,34] assessed the effect of community-level drinking water on child or infant mortality by reconstructing cohorts of children based on 5- or 10-year birth histories, two studies[35,36] assessed child malnutrition (height for age), and one study assessed low birth weight.[37] Most showed a protective effect, though the results were not always consistent.

These studies had important methodological limitations. First, many did not adjust for community-level confounding variables. Just as socioeconomic status is an important confounder at the household level,[38] communities with differences in WASH coverage are likely different in many other important ways. Failure to account for potential confounding variables at the community level makes causal inference dubious. Second, many did not account for WASH access at the household level. Failure to account for household-level drinking water makes the assessment of herd protection difficult. Household-level access may be confounding the

protective effect of community-level access, so the protective effect at the community level is actually a mixture of the direct and indirect effects, with no ability to disentangle to two. A strength of our simulation approach is that it eliminates any potential confounding by providing perfect experimental conditions. Finally, some studies were often not focused on estimating the unbiased protective effect of drinking water, but simply attempted to show that community-level variables in general have explanatory power, with little discussion or understanding of the pathogen transmission process. Only one study[17] did not include any of these limitations, and it showed no protective effect of community-level coverage of improved drinking water.

Sanitation.

As coverage of the sanitation intervention increases, the risk of infection in both the intervention group and the control group decreases (Figure 4, Panel A). When coverage of the sanitation intervention reaches 100%, the median risk is 44.7 per 100 persons, providing a total protective efficacy of 35.8% (0.358 = 1 – 0.447/0.696). This total effect is entirely attributable to the indirect effect, as the direct effect is negligible (Figure 4, Panel B). In other words, sanitation provides no direct benefit to the user, but protects the entire community equally.

Several studies have suggested that sanitation provides herd protection to neighboring households (Table 3 and Supplemental Table 2). The majority of these were also from cross-sectional national household surveys. Outcomes in these studies varied from diarrhea,[15,39–41] prevalence of parasitic infection,[42] infant or child mortality,[17,34,43] child stunting,[16,35,44] and child wasting.[45] The majority of these studies had the same limitations mentioned for the drinking water studies, namely insufficient adjustment for community-level confounders, no adjustment for household-level sanitation, and no discussion of interrupting pathogen transmission. There were, however, a few exceptions.[15–17,42] Barreto and others[15] conducted two cohort studies in Salvador, Brazil, one before a city-wide sanitation campaign and one after. The study attributed the entirety of the 21% reduction in diarrhea prevalence to the increase in sewer-connected toilet coverage in the neighborhood (26–80%). Increases in household-level toileting did not explain the reduction. Our model results confirm these findings. Within the same study setting, Barreto and others[42] also assessed the prevalence of various intestinal parasites before and after the city-wide sanitation campaign. For the three parasites studied, the prevalence of each dropped substantially, and 25–40% of the reduced prevalence was attributable to the increase in sewer-connected toilet coverage in the neighborhood. Fuller and others[16] used a cohort study in rural Ecuador to show that the total effect of sanitation was a 72% lower prevalence of stunting, and 94% of this total effect was attributable to the indirect effect of sanitation coverage. Finally, Van de Poel and others[17] used data from six Demographic and Health Surveys in sub-Saharan Africa, and found that sanitation coverage at the community level was surprisingly a risk factor for infant mortality in urban areas; sanitation coverage in rural areas was not associated with infant mortality.

Hand hygiene.

As coverage of the hand hygiene intervention increases, the risk in both the intervention group and the control group decreases (Figure 5, Panel A). When coverage reaches 100%, the median risk in the community is 43.1%, providing a total protective efficacy of 38.1% (0.381 = 1 – 0.431/0.696). At all levels of coverage, the risk in the intervention group is lower than that of the control group (Figure 5, Panel A). The direct protective efficacy of hand hygiene is about

15%, and is relatively constant across all levels of coverage (Figure 5, Panel B). As coverage increases, however, the indirect effect becomes more pronounced, becoming roughly equal to the direct effect at 70% coverage.

We found no studies showing that hand hygiene in one household provides a benefit to neighboring households. One study[46] showed that when the food preparer practices handwashing, then children in the household are less likely to have diarrhea. This highlights that the indirect effects from WASH interventions may occur not only between households as shown by our model and literature review but also within households. It also raises an important point of within-household heterogeneity of WASH practices. Although these interventions are often applied at the household level, household members may comply to differing degrees, particularly young children that often use diapers instead of defecating directly into a pit latrine.

DISCUSSION

WASH interventions can provide both direct protection to the user and indirect (herd) protection to the entire community. Household-level drinking water interventions aim to reduce the number of pathogens that the user ingests, providing a clear direct benefit to that user. Individuals from neighboring households, however, may receive an indirect benefit because there are fewer total infections in the community, which will result in less cumulative environmental contamination, though the rates of shedding are unchanged. This type of indirect effect should arise in any communicable disease system whenever an intervention reduces the prevalence of infection, as susceptible persons in the community will be less likely to have contact with an infectious individual, whether that contact is direct or environmentally mediated.

Our model simulations and literature review suggest that the indirect effects from a sanitation intervention are a larger component of the total effect compared with those from a drinking water or hand hygiene intervention. If sanitation can reduce the rate of shedding pathogens into the community environment, all surrounding households will benefit regardless of their own sanitation practices. In our model, sanitation provides no direct benefit to the household. This is in some ways similar to the transmission-blocking malaria vaccines currently in development, wherein the recipient of the vaccine would not be immune to infection or symptoms, but the parasite would be unable to be transmitted to a susceptible mosquito.[47,48] The only benefit to the vaccinated person is the indirect effect shared by everyone in the community. Sanitation also has the potential to provide a direct benefit to the household if the intervention can reduce contamination of the household environment.

Hand hygiene is unique in that it can both reduce the contagiousness of an infected individual and protect a susceptible individual. The direct effect occurs because the person washing their hands is reducing the number of pathogens that they will ingest. The indirect effect occurs for two reasons. The first is similar to the mechanism of the household drinking water intervention in that by preventing new infections, it reduces the overall number of pathogens excreted into the environment. The second is like that of sanitation, namely a reduced shedding rate into the environment.

There are many studies reporting the results of cluster-randomized trials of drinking water,[49,50] sanitation,[51,52] and hand hygiene interventions,[46,53,54] which we have not included in our review of the literature. There are several motivations for the use of this study design, one of which is the potential to account for indirect effects or herd protection.[3] Published articles, however, have reported the results in terms of the total effect or the overall effect, without

disaggregating the direct and indirect effects. Although these studies do not underestimate the total impact of an intervention, they do not provide a complete picture of the type of impact WASH may have. Specifically, understanding the indirect effects can be essential for disease control strategies, especially if the indirect effect is nonlinear with increases in coverage, as is the case with a herd immunity threshold. Also, identifying herd protection and not just a total effect suggests that these interventions may protect households unreached by intervention campaigns. It is possible that these cluster-randomized trials could be reanalyzed to assess herd protection from various WASH interventions,[4] though the benefits of randomization would likely be lost.

As in all modeling exercises, our findings could be sensitive to a relaxation of our simplifying assumptions. First, WASH-related practices vary substantially across communities, between individuals and even within individuals. Our model assumes 100% compliance with the intervention, though recent studies have shown that efficacious WASH interventions yield no benefit if high levels of coverage are not accompanied by high levels of compliance.[51,52,55] Second, many enteric pathogens are transmitted person to person, but our model does not explicitly capture this pathway. Although we did not model transmission as a traditional mass action, environmentally mediated transmission partially captures this phenomenon, as some person-to-person transmission may actually occur via fomites and other objects. Third, due to very few quality studies on the subject, our analysis did not seek to estimate the actual amount of herd protection from a given intervention. Also, this will vary substantially across different settings. For example, open defecation in a rural setting may occur in the bush, far from human dwellings. Such a practice may create little to no risk for other individuals in the village.[56] The degree of environmental connectivity[57] between households will vary by setting and pathogen. The magnitude of the indirect effect largely depends on the relative proportion of transmission that occurs between households compared with transmission within households. Drinking water interventions, for example, will provide a much smaller indirect effect if contamination from one household is less able to reach a shared environment. The degree of herd protection will also depend on the type of pathogens circulating in the community, as different pathogens are able to exploit different pathways to varying degrees. Cholera, for example, has a high infectious dose and thrives in surface water enhancing its ability to survive in water-based community environment. *Shigella* is less able to survive in the environment and has a much lower infectious dose, allowing it to be readily transmitted via food and hands and possibly increasing within-household transmission. Also, we have oversimplified what are often complex interventions. For example, hand hygiene (handwashing) interventions are only a subset of what could be considered hygiene. The same could be said for drinking water and sanitation interventions, which often focus on infrastructure, but could also include behavior change. This simplification, however, serves to highlight the important mechanisms through which each class of intervention may provide protection to nonrecipients. Finally, we have simulated short-term epidemics of enteric pathogens. In reality, these pathogens are often endemic or become endemic after an outbreak.

Our model was constructed with the intent of illustrating the mechanisms through which WASH interventions may provide indirect protection. To date, data illustrating herd protection for water sanitation and hygiene are still relatively limited. As more high-quality empirical studies are produced, it will be possible to obtain more robust estimates of model's parameters and outcomes, and thereby address more detailed questions pertaining to the mechanisms driving

herd protection. Data may also provide insight into how generalizable herd protection is across different environmental and social conditions as well as to different pathogens.

Our study has important implications for both future research and program delivery. Assessment of WASH interventions has largely ignored community interdependence.[58] Studies assessing the health benefits of WASH interventions without accounting for herd protection will likely underestimate the total efficacy, particularly for sanitation interventions whose benefits may be largely due to their indirect effects. Future studies should seek to quantify these additional benefits, which would alter the cost-effectiveness calculations of decision makers. Also, WASH programs often aim to achieve 100% coverage in communities, but 100% coverage, not to mention 100% compliance, is elusive.[59] Immunization policy is often based on reaching a threshold of vaccination coverage, often less than 100%, at which transmission will be interrupted and the disease eliminated. It is unclear, however, whether such a threshold exists for WASH interventions. Interventions for the control of infectious diseases can provide indirect protection to nonusers. Although the mechanism behind herd protection varies by pathogen and transmission cycle, the goal of providing sufficient coverage to interrupt transmission transcends these differences.

Received September 17, 2015.

Accepted for publication July 20, 2016.

Note: Supplemental appendices and tables appear at www.ajtmh.org.

Acknowledgments:

We would like to thank Marisa Eisenberg, Rafael Meza, James Koopman, and Manish Desai for their insightful comments on the manuscript.

Financial support: This work was supported by the National Institute of Allergy and Infectious Diseases (grant no. RO1-AI050038), the Ecology of Infectious Disease Program of the Fogarty International Center of the National Institutes of Health and the National Science Foundation (grant no. 0811934), and the National Institute of General Medical Sciences of the National Institutes of Health (grant no. U01GM110712).

Authors' addresses: James A. Fuller and Joseph N. S. Eisenberg, Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, E-mails: jafuller@umich.edu and jnse@umich.edu.

REFERENCES

<jrn>1. Bartram J, Cairncross S, 2010. Hygiene, sanitation, and water: forgotten foundations of health. *PLoS Med 7:* e1000367.</jrn>

<jrn>2. Halloran ME, Struchiner CJ, 1991. Study designs for dependent happenings. *Epidemiology 2:* 331–338.</jrn>

<jrn>3. Hayes RJ, Alexander ND, Bennett S, Cousens SN, 2000. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Methods Med Res 9:* 95–116.</jrn>

<jrn>4. Ali M, Emch M, von Seidlein L, Yunus M, Sack DA, Rao M, Holmgren J, Clemens JD, 2005. Herd immunity conferred by killed oral cholera vaccines in Bangladesh: a reanalysis. *Lancet 366:* 44–49.</jrn>

<jrn>5. Musher DM, 2006. Pneumococcal vaccine: direct and indirect ("herd") effects. *N Engl J Med 354:* 1522–1524.</jrn>

<jrn>6. Hawley WA, Phillips-Howard PA, ter Kuile FO, Terlouw DJ, Vulule JM, Ombok M, Nahlen BL, Gimnig JE, Kariuki SK, Kolczak MS, Hightower AW, 2003. Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *Am J Trop Med Hyg 68:* 121–127.</jrn>

<jrn>7. Hii JL, Smith T, Vounatsou P, Alexander N, Mai A, Ibam E, Alpers MP, 2001. Area effects of bednet use in a malaria-endemic area in Papua New Guinea. *Trans R Soc Trop Med Hyg 95:* 7–13.</jrn>

<jrn>8. Howard SC, Omumbo J, Nevill C, Some ES, Donnelly CA, Snow RW, 2000. Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Trans R Soc Trop Med Hyg 94:* 357–360.</jrn>

<jrn>9. Miguel E, Kremer M, 2004. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72:* 159–217.</jrn>

<jrn>10. Bundy DA, Wong MS, Lewis LL, Horton J, 1990. Control of geohelminths by delivery of targeted chemotherapy through schools. *Trans R Soc Trop Med Hyg 84:* 115–120.</jrn>

<jrn>11. Attia S, Egger M, Muller M, Zwahlen M, Low N, 2009. Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. *AIDS 23:* 1397–1404.</jrn>

<jrn>12. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG, 2009. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet 373:* 48–57.</jrn>

<jrn>13. Chidambaram JD, Melese M, Alemayehu W, Yi E, Prabriputaloong T, Lee DC, Cevallos V, Zhou Z, Whitcher JP, Gaynor BD, Lietman TM, 2004. Mass antibiotic treatment and community protection in trachoma control programs. *Clin Infect Dis 39:* e95–e97.</jrn>

<jrn>14. House JI, Ayele B, Porco TC, Zhou Z, Hong KC, Gebre T, Ray KJ, Keenan JD, Stoller NE, Whitcher JP, Gaynor BD, Emerson PM, Lietman TM, 2009. Assessment of herd protection against trachoma due to repeated mass antibiotic distributions: a cluster-randomised trial. *Lancet 373:* 1111–1118.</jrn>

<jrn>15. Barreto ML, Genser B, Strina A, Assis AMO, Rego RF, Teles CA, Prado MS, Matos SM, Santos DN, dos Santos LA, 2007. Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort studies. *Lancet 370:* 1622–1628.</jrn>

<jrn>16. Fuller JA, Villamor E, Cevallos W, Trostle J, Eisenberg JNS, 2016. I get height with a little help from my friends: herd protection from sanitation on child growth in rural Ecuador. *Int J Epidemiol 45:* 460–469.</jrn>

<jrn>17. Van de Poel E, O'donnell O, Van Doorslaer E, 2009. What explains the rural-urban gap in infant mortality: household or community characteristics? *Demography 46:* 827–850.</jrn>

<jrn>18. Alderman H, Hentschel J, Sabates R, 2003. With the help of one's neighbors: externalities in the production of nutrition in Peru. *Soc Sci Med 56:* 2019–2031.</jrn>

<jrn>19. Eisenberg JN, Scott JC, Porco T, 2007. Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. *Am J Public Health 97:* 846–852.</jrn>

<jrn>20. Li S, Eisenberg JN, Spicknall IH, Koopman JS, 2009. Dynamics and control of infections transmitted from person to person through the environment. *Am J Epidemiol 170:* 257–265.</jrn>

<bok>21. R Development Core Team, 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.</bok>

<jrn>22. Aiello AE, Coulborn RM, Perez V, Larson EL, 2008. Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis. *Am J Public Health 98:* 1372–1381.</jrn>

<jrn>23. Arnold BF, Colford JM Jr, 2007. Treating water with chlorine at point-of-use to improve water quality and reduce child diarrhea in developing countries: a systematic review and meta-analysis. *Am J Trop Med Hyg 76:* 354–364.</jrn>

<jrn>24. Cairncross S, Hunt C, Boisson S, Bostoen K, Curtis V, Fung IC, Schmidt WP, 2010. Water, sanitation and hygiene for the prevention of diarrhoea. *Int J Epidemiol 39 (Suppl 1):* i193–i205.</jrn>

<jrn>25. Clasen T, Schmidt WP, Rabie T, Roberts I, Cairncross S, 2007. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ 334:* 782.</jrn>

<jrn>26. Curtis V, Cairncross S, 2003. Effect of washing hands with soap on diarrhoea risk in the community: a systematic review. *Lancet Infect Dis 3:* 275–281.</jrn>

<jrn>27. Esrey SA, Potash JB, Roberts L, Shiff C, 1991. Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. *Bull World Health Organ 69:* 609–621.</jrn>

<jrn>28. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM Jr, 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis 5:* 42–52.</jrn>

<jrn>29. Gundry S, Wright J, Conroy R, 2004. A systematic review of the health outcomes related to household water quality in developing countries. *J Water Health 2:* 1–13.</jrn>

<jrn>30. Hunter PR, 2009. Household water treatment in developing countries: comparing different intervention types using meta-regression. *Environ Sci Technol 43:* 8991–8997.</jrn>

<bok>31. Waddington H, Snilstveit B, White H, Fewtrell L, 2009. *Water, Sanitation and Hygiene Interventions to Combat Childhood Diarrhoea in Developing Countries*. New Delhi, India: International Initiative for Impact Evaluation.</bok>

<jrn>32. Wolf J, Prüss-Ustün A, Cumming O, Bartram J, Bonjour S, Cairncross S, Clasen T, Colford JM, Curtis V, France J, 2014. Systematic review: assessing the impact of drinking water and sanitation on diarrhoeal disease in low-and middle-income settings: systematic review and meta-regression. *Trop Med Int Health 19:* 928–942.</jrn>

<jrn>33. Adedini SA, Odimegwu C, Imasiku EN, Ononokpono DN, Ibisomi L, 2015. Regional variations in infant and child mortality in Nigeria: a multilevel analysis. *J Biosoc Sci 47:* 165–187.</jrn>

<jrn>34. Sastry N, 1996. Community characteristics, individual and household attributes, and child survival in Brazil. *Demography 33:* 211–229.</jrn>

<jrn>35. Adekanmbi VT, Kayode GA, Uthman OA, 2013. Individual and contextual factors associated with childhood stunting in Nigeria: a multilevel analysis. *Matern Child Nutr 9:* 244–259.</jrn>

<jrn>36. Chikhungu LC, Madise NJ, Padmadas SS, 2014. How important are community characteristics in influencing children's nutritional status? Evidence from Malawi population-based household and community surveys. *Health Place 30:* 187–195.</jrn>

<jrn>37. Kayode GA, Amoakoh-Coleman M, Agyepong IA, Ansah E, Grobbee DE, Klipstein-Grobusch K, 2014. Contextual risk factors for low birth weight: a multilevel analysis. *PLoS One 9:* e109333.</jrn>

<jrn>38. Fuller JA, Westphal JA, Kenney B, Eisenberg JN, 2015. The joint effects of water and sanitation on diarrhoeal disease: a multicountry analysis of the Demographic and Health Surveys. *Trop Med Int Health 20:* 284–292.</jrn>

<bok>39. Andres LA, Briceño B, Chase C, Echenique JA, 2014. *Sanitation and Externalities: Evidence from Early Childhood Health in Rural India*. Washington, DC: World Bank.</bok>

<jrn>40. Root GP, 2001. Sanitation, community environments, and childhood diarrhoea in rural Zimbabwe. *J Health Popul Nutr 19:* 73–82.</jrn>

<jrn>41. Santos CA, Strina A, Amorim LD, Genser B, Assis AM, Prado MS, Barreto ML, 2012. Individual and contextual determinants of the duration of diarrhoeal episodes in preschool children: a longitudinal study in an urban setting. *Epidemiol Infect 140:* 689–696.</jrn>

<jrn>42. Barreto ML, Genser B, Strina A, Teixeira MG, Assis A, Rego RF, Teles CA, Prado MS, Matos S, Alcântara-Neves NM, 2011. Impact of a city-wide sanitation programme in northeast Brazil on intestinal parasites infection in young children. *Environ Health Perspect 118:* 1637–1642.</jrn>

<other>43. Geruso M, Spears D, 2015. *Neighborhood Sanitation and Infant Mortality*. Working Paper 21184. Cambridge, MA: National Bureau of Economic Research.</other>

<conf>44. Bateman OM, Smith S, 1991. *A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala*. The Demographic and Health Surveys World Conference, Volume 2, August 5–7, 1991, Columbia, MD.</conf>

<jrn>45. Buttenheim AM, 2008. The sanitation environment in urban slums: implications for child health. *Popul Environ 30:* 26–47.</jrn>

<jrn>46. Luby SP, Halder AK, Huda T, Unicomb L, Johnston RB, 2011. The effect of handwashing at recommended times with water alone and with soap on child diarrhea in rural Bangladesh: an observational study. *PLoS Med 8:* 798.</jrn>

<jrn>47. Nunes JK, Woods C, Carter T, Raphael T, Morin MJ, Diallo D, Leboulleux D, Jain S, Loucq C, Kaslow DC, 2014. Development of a transmission-blocking malaria vaccine: progress, challenges, and the path forward. *Vaccine 32:* 5531–5539.</jrn>

<jrn>48. Delrieu I, Leboulleux D, Ivinson K, Gessner BD; Malaria Transmission Blocking Vaccine Technical Consultation Group, 2015. Design of a phase III cluster randomized trial to assess the efficacy and safety of a malaria transmission blocking vaccine. *Vaccine 33:* 1518–1526.</jrn>

<jrn>49. Mausezahl D, Christen A, Pacheco GD, Tellez FA, Iriarte M, Zapata ME, Cevallos M, Hattendorf J, Cattaneo MD, Arnold B, 2009. Solar drinking water disinfection (SODIS) to reduce childhood diarrhoea in rural Bolivia: a cluster-randomized, controlled trial. *PLoS Med 6:* 882.</jrn>

<jrn>50. McGuigan KG, Samaiyar P, du Preez M, Conroy RM, 2011. High compliance randomized controlled field trial of solar disinfection of drinking water and its impact on childhood diarrhea in rural Cambodia. *Environ Sci Technol 45:* 7862–7867.</jrn>

<jrn>51. Clasen T, Boisson S, Routray P, Torondel B, Bell M, Cumming O, Ensink J, Freeman M, Jenkins M, Odagiri M, 2014. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *Lancet Glob Health 2:* e645–e653.</jrn>

<jrn>52. Patil SR, Arnold BF, Salvatore AL, Briceno B, Ganguly S, Colford JM Jr, Gertler PJ, 2014. 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 11:* e1001709.</jrn>

<jrn>53. Luby SP, Agboatwalla M, Painter J, Altaf A, Billhimer WL, Hoekstra RM, 2004. Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan: a randomized controlled trial. *JAMA 291:* 2547–2554.</jrn>

<jrn>54. Luby SP, Agboatwalla M, Feikin DR, Painter J, Billhimer W, Altaf A, Hoekstra RM, 2005. Effect of handwashing on child health: a randomised controlled trial. *Lancet 366:* 225–233.</jrn>

<jrn>55. Enger KS, Nelson KL, Rose JB, Eisenberg JN, 2013. The joint effects of efficacy and compliance: a study of household water treatment effectiveness against childhood diarrhea. *Water Res 47:* 1181–1190.</jrn>

<jrn>56. Luby S, 2014. Is targeting access to sanitation enough? *Lancet Glob Health 2:* e619–e620.</jrn>

<jrn>57. Emch M, Ali M, Root ED, Yunus M, 2009. Spatial and environmental connectivity analysis in a cholera vaccine trial. *Soc Sci Med 68:* 631–637.</jrn>

<jrn>58. Eisenberg JN, Trostle J, Sorensen RJ, Shields KF, 2012. Toward a systems approach to enteric pathogen transmission: from individual independence to community interdependence. *Annu Rev Public Health 33:* 239–257.</jrn>

<jrn>59. Barnard S, Routray P, Majorin F, Peletz R, Boisson S, Sinha A, Clasen T, 2013. Impact of Indian Total Sanitation Campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation. *PLoS One 8:* e71438.</jrn>

FIGURE 1. The framework proposed by Halloran and Struchiner[2] for measuring the herd protection from an intervention, where D_1 represents the disease risk in the intervention group, D_0 represents the disease risk in the control group, and D_{0*} represents the disease risk in a separate population where there is no intervention.

FIGURE 2. Schematic of household transmission model. Conceptualization of drinking water, sanitation, and hygiene interventions are also shown, where δ and ϕ represent the rate at which infectious individuals shed pathogens into their household and community environments, respectively, ρ and α represent the rate at which susceptible individuals pick up pathogens from their household and community environments respectively, and μ represents the rate at which pathogens die off in both the household and community environment (see Table 1 and text for details).

FIGURE 3. Model simulations showing the effect of a drinking water intervention across different levels of intervention coverage in the community. At each level of coverage of the intervention, the stochastic model was simulated 100 times. Panels A and B show the median values for the cumulative incidence (proportion infected) and protective efficacy, respectively, where D_1 represents the disease risk in the intervention group, D_0 represents the disease risk in the control group, and D_{0*} represents the disease risk in a separate population where there is no intervention (see Figure 1 and text for more details).

FIGURE 4. Model simulations showing the effect of a sanitation intervention across different levels of intervention coverage in the community. At each level of coverage of the intervention, the stochastic model was simulated 100 times. Panels A and B show the median values for the cumulative incidence (proportion infected) and protective efficacy, respectively, where D_1 represents the disease risk in the intervention group, D_0 represents the disease risk in the control group, and D_{0*} represents the disease risk in a separate population where there is no intervention (see Figure 1 and text for details).

FIGURE 5. Model simulations showing the effect of a hand hygiene intervention across different levels of intervention coverage in the community. At each level of coverage of the intervention, the stochastic model was simulated 100 times. Panels A and B show the median values for the cumulative incidence (proportion infected) and protective efficacy, respectively, where D_1 represents the disease risk in the intervention group, D_0 represents the disease risk in the control group, and D_{0*} represents the disease risk in a separate population where there is no intervention (see Figure 1 for details).

TABLE 1

Parameter values and their description used in the analysis

Parameter	Description	Value
General parameters		
–	The number of households in the community	100
–	The number of individuals per household	~N (5, SD = 2)*
μ	The rate at which pathogens die in the environment	1/10
γ	The rate at which infectious individuals recover from infection	1/3
Control group parameters		
α	The rate at which susceptible individuals in the control group pick up pathogens from the community environment	$1/10^6$
ϕ	The rate at which infectious individuals in the control group shed pathogens into the community environment	78
δ	The rate at which infectious individuals in the control group shed pathogens into their own household environment	395
ρ	The rate at which susceptible individuals in the control group pick up pathogens from their household environment	$1/10^5$
Intervention group parameters		
α_1	The rate at which susceptible individuals using a drinking water intervention pick up pathogens from the community environment	$0.73/10^6$
ϕ_1	The rate at which infectious individuals using a sanitation intervention shed pathogens into the community environment	56.94
δ_1	The rate at which infectious individuals using a hygiene intervention shed pathogens into their own household environment	173.8
ρ_1	The rate at which susceptible individuals using a hygiene intervention pick up pathogens from their household environment	$0.44/10^5$

SD = standard deviation. All rates are per day.

* Normally distributed, with a mean of 5 and a standard deviation of 2.

TABLE 2

Summary of studies assessing herd protection from drinking water grouped by outcome

Setting	Specific outcome	Conceptualization of drinking water coverage	Effect on outcome	Limitations*
Mortality				
6 Sub-Saharan African countries: urban[17]	Infant mortality	Proportion of households in survey cluster with water from a tap, protected well, bottle, or vendor	Probit regression coefficient = 0.064 ($P > 0.05$)	–
6 Sub-Saharan African countries: rural[17]	Infant mortality	Proportion of households in survey cluster with water from a tap, protected well, bottle, or vendor	Probit regression coefficient = 0.007 ($P > 0.05$)	–
Nigeria: national[33]	Infant mortality	% of households in survey cluster with piped water (low, medium, high)	Medium vs. low, HR = 0.67 ($P < 005$); high vs. low, HR = 0.87 ($P < 0.05$)	b, c
Nigeria: national[33]	Child mortality	% of households in survey cluster with piped water (low, medium, high)	Medium vs. low, HR = 1.01 ($P > 0.05$); high vs. low, HR = 1.42 ($P < 0.05$)	b, c
Brazil: northeast[34]	Child mortality	% of households in municipality with regular network water or % with well water	Network water: HR = 0.30 ($P > 0.05$); well water: HR = 0.34 ($P < 0.05$)	b
Brazil: south and southeast[34]	Child mortality	% of households in municipality with regular network water	Network water: HR = 30.9 ($P > 0.05$); well water: HR = 4.54 ($P > 0.05$)	b
Nutrition				
Nigeria: national[35]	Stunting	Safe water, community level (no vs. yes, undefined)	OR = 1.08 (95% CL = 0.74–1.15)	b, c
Malawi: national[36]	Stunting	% of households in survey cluster with protected water source (> 73% vs. ≤ 73%)	OR = 0.83 (95% CL = 0.55–1.01)	a, c
Ghana: national[37]	Low birth weight	% of households in survey cluster with access to safe water (high vs. low, undefined)	OR = 0.74 (95% CL = 0.57–0.96)	b, c

HR = hazard ratio; OR = odds ratio; WASH = drinking water, sanitation, and hand hygiene. No studies could be related to the Halloran[2] framework. Studies are grouped by outcome. Additional study details can be found in the Supplemental Appendix B.

* Potential limitations include: a—did not adjust or insufficiently adjusted for community-level socioeconomic status, b—did not include the corresponding WASH exposure at the household or individual level, c—discussion of social context but not of disease transmission process.

TABLE 3

Summary of studies assessing herd protection from sanitation and their relation to the Halloran[2] framework where relevant

Setting	Specific outcome	Conceptualization of sanitation coverage	Effect on outcome	Relation¶ to the Halloran framework				Limitations*
				Overall effect (protective efficacy)	Total effect (protective efficacy)	Indirect effect (protective efficacy)	% of overall/total effect due to the indirect effect	
Diarrhea								
Brazil: urban[15]	Diarrhea prevalence	Primary exposure is the intervention (after vs. before). Potential mediator is proportion of households in area with sewer connection	PR (after vs. before intervention) = 0.78; 100% mediated by area sewer coverage	22%	–	22%	100%	–
India: rural[39]	Diarrhea prevalence	Proportion of households in a village with access to improved sanitation	Estimated with quadratic	–	47%	35%	75%	a
Zimbabwe: rural[40]	No. of episodes of diarrhea in 45 weeks	Among children with no sanitation, whether the nearest neighbor had improved latrine	Mean difference: 1.13 fewer episodes when nearest neighbor has improved latrine	–	–	–	–	a
Brazil: urban[41]	Duration (days) of diarrheal episodes	% of households with sewer connection: very low (< 34.1%), low (34.147.1%), normal (47.2–55.0%), good (> 55.0%)	Mean difference (days): normal vs. good 0.07, low vs. good 0.23, very low vs. good 0.47	–	–	–	–	a, b
Parasitic infection								
Brazil: urban[42]	Prevalence of *Ascaris lumbricoides*	Primary exposure is the intervention (after vs. before). Potential mediator is proportion of households in area with sewer connection	PR (after vs. before) = 0.57; 40% mediated by area sewer coverage	43%	–	17%	40%	–
Brazil: urban[42]	Prevalence of *Trichuris trichiura*	Primary exposure is the intervention (after vs. before). Potential mediator is proportion of households in area with sewer connection	PR (after vs. before) = 0.38; 30% mediated by area sewer coverage	62%	–	19%	30%	–
Brazil: urban[42]	Prevalence of *Giardia duodenalis*	Primary exposure is the intervention (after vs. before). Potential mediator is proportion of households in area with sewer connection	PR (after vs. before) = 0.41; 25% mediated by area sewer coverage	59%	–	15%	25%	–
Mortality								
6 Sub-Saharan African countries: urban[17]	Infant mortality	Proportion of households in survey cluster with any toilet facility	Probit regression coefficient: 0.314 ($P < 0.01$)	–	–	–	–	–
6 Sub-Saharan African countries: rural[17]	Infant mortality	Proportion of households in survey cluster with any toilet facility	Probit regression coefficient: 0.008 ($P > 0.05$)	–	–	–	–	–
India: national[43]	Infant mortality	Proportion of households in survey cluster NOT practicing open defecation	27.1 fewer deaths per 1,000 children	–	35†§	27†§	77%	a
Brazil: northeast[34]	Child mortality	% of households in municipality with sewage connection or % with any sanitation	Sewage connection: HR = 0.08 ($P > 0.05$); any sanitation: HR = 0.29 ($P > 0.05$)	–	–	–	–	b
Brazil: south and southeast[34]	Child mortality	% of households in municipality with sewage connection or % with any sanitation	Sewage connection: HR = 0.67 ($P > 0.05$); any sanitation: HR = 0.33 ($P > 0.05$)	–	–	–	–	b
Nutrition								
Ecuador: rural[16]	Stunting	Proportion of households within 500 m that have improved sanitation	PR = 0.32 (95% CL = 0.15–0.69)	–	72%§	68%§	94%	–
Nigeria: national[35]	Stunting	Proper sanitation (no definition provided), community level (yes vs. no)	OR = 0.83 (95% CL = 0.74–1.12)	–	–	–	–	b, c
Guatemala: urban[44]	Stunting	Binary: 0 if proportion of children < 5 in survey cluster with flush toilet was < 75%, 1 if proportion was ≥ 75%	OR = 0.47 ($P < 0.01$)	–	75%§	53%§	71%	a
Guatemala: rural[44]	Stunting	Binary: 0 if proportion of children < 5 in survey cluster with flush toilet was < 75%, 1 if proportion was ≥ 75%	OR = 0.34 ($P < 0.05$)		85%§	66%§	78%	a
Bangladesh: urban[45]	Wasting	Percentage of households with children < 4 in community with improved sanitation	z-score difference (100% coverage vs. 0% coverage): 1.2	–	1.4‡	1.2‡	85%	a

HR = hazard ratio; OR = odds ratio; PR = prevalence ratio. Studies are grouped by outcome. Additional study details can be found in the Supplemental Appendix B.

* Potential limitations include: a—did not adjust or insufficiently adjusted for community-level socioeconomic status, b—did not include the corresponding WASH exposure at the household or individual level, c—discussion of social context but not of disease transmission process.

† Protective efficacy reported as absolute difference in number of deaths per 1,000 children.

‡ Protective efficacy reported as absolute difference in weight-for-height z-score.

§ The study did not stratify by household-level access or include an interaction between household access and coverage, so the reported indirect effect is not a true indirect effect. Assuming no interaction, however, the effect of sanitation coverage can be interpreted as an indirect effect. We estimated the total effect for these studies as the sum or product of the effects of household access and coverage.

¶ We have not shown these values for studies that did not account for household- or individual-level access to sanitation, since the reported association cannot be interpreted as an indirect effect.

SUPPLEMENTAL APPENDIX A

R CODE USED TO GENERATE FIGURE 3

```
#########################################################################
#########################
#This File Runs the Models to Create Figure 3
#The User only Needs to change the working directory where files created by
this script will be saved
setwd("C:/…")
#########################################################################
#########################
#Clear the Environment
rm(list = ls())
#####################################################################
#Set up the General Parameters of the Model (Values from Table 1)
parm <- c('gamma'=1/3, 'pi'=1, 'delta.0'=395, 'delta.1'=395,
'phi.0'=78, 'phi.1'=78, 'rho.0'=1e-5, 'rho.1'=1e-5,
'alpha.0'=1e-6, 'alpha.1'=0.73*1e-6, 'mu.HH.0'=1/10, 'mu.HH.1'=1/10,
'mu.E'=1/10)
#Gamma=Duration of Infection
#pi=scaling parameter
#delta=Pathogen shedding into Household Environment
#phi=Pathogen shedding into Community Environment
#rho=Pathogen Pickup from Household Environment
#alpha=Pathogen Pickup from Community Environment
#mu=Rate of Pathogen die-off in the environment
#####################################################################
#Set up the Initial Conditions of the Model
set.seed(1) #Set the seed for the random number generator, so results can be
replicated
N.HH <- 100 #Number of Household in the Community
N.Pop <- 500 #Size of Population (Number of Individuals)
HH.Size <- round(rnorm(N.HH, N.Pop/N.HH, 1.5)) #Create Vector of HH Sizes
N.Pop.Actual <- sum(HH.Size) #Actual Population Size, should be close to
N.Pop (500)
N.Pop.Actual
Io <- 1 #Number of Initial Infecteds
#####################################################################
#Create a vector the different coverage levels to be used
coverage.levels <- c(0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0)
#####################################################################
#How many times to run the model at each level of coverage?
N.runs <- 100
#####################################################################
#Create Empty Data Frame to be filled
data.total <- data.frame(coverage=NA, run=NA,
risk.overall=NA, risk.0=NA, risk.1=NA,
Inf.HH=NA, Inf.E=NA, Inf.Total=NA, prop.HH=NA, prop.E=NA,
rr.direct=NA, rr.indirect=NA, rr.total=NA, rr.overall=NA)
#####################################################################
coverage.n <- 0 #A counter for the level of coverage
#####################################################################
for (coverage in coverage.levels) { #Loop across levels of coverage (0-100%,
by 10%)
#Running counts
coverage.n <- coverage.n+1
```

```
for (run in 1:N.runs) {
#Report overall progress of the simulations
percent <- ((((coverage.n-
1)*N.runs)+run)/(length(coverage.levels)*N.runs))*100
cat('Coverage =',coverage,"\nRun =",run,"\nOverall Progress
=",round(percent,1),"%\n","\n")
#Create a vector indicating if HH has improved (1) or unimproved (0)
sanitation
sanitation <- vector(mode='numeric', length=N.HH)
for (i in 1:N.HH) {
if (i <= round(coverage*N.HH)) sanitation[i] <- 1
}
#Summarize Population according to sanitation
#(Number of individuals using improved/unimproved sanitation)
N.San.0 <- 0
N.San.1 <- 0
for (i in 1:N.HH) {
if (sanitation[i]==0) N.San.0 <- N.San.0 + HH.Size[i]
if (sanitation[i]==1) N.San.1 <- N.San.1 + HH.Size[i]
}
#Create Matrix of Susceptibles, Infected, Recovered, and HH Environments
#1st Column is the current event, 2nd column is for the next event
HHS <- matrix(nrow=N.HH, ncol=2)
HHI <- matrix(nrow=N.HH, ncol=2)
HHR <- matrix(nrow=N.HH, ncol=2)
HHE <- matrix(nrow=N.HH, ncol=2)
for (i in 1:N.HH) {
HHS[i,1] <- HH.Size[i] #Begin with everyone susceptible
HHI[i,1] <- 0 #Begin with no infected
HHR[i,1] <- 0 #Begin with no immune
HHE[i,1] <- 0 #Begin with a clean HH environment
}
#Create vector of the Community Environment
E <- c(0, NA) #Begin with a clean Community environment
#Create Vectors of Parameters for Each Household
gamma <- vector()
delta <- vector()
phi <- vector()
rho <- vector()
alpha <- vector()
mu.HH <- vector()
mu.E <- parm['mu.E']
for (i in 1:N.HH) {
gamma[i] <- parm['gamma']
if (sanitation[i]==0) {
delta[i] <- parm['delta.0']
phi[i] <- parm['phi.0']
rho[i] <- parm['rho.0']
alpha[i] <- parm['alpha.0']
mu.HH[i] <- parm['mu.HH.0']
}
if (sanitation[i]==1) {
delta[i] <- parm['delta.1']
phi[i] <- parm['phi.1']
rho[i] <- parm['rho.1']
alpha[i] <- parm['alpha.1']
mu.HH[i] <- parm['mu.HH.1']
```

```
}
}
#Randomly place initial infected in the population
infHouseholds <- sample(1:N.HH, Io, replace=TRUE, prob=HH.Size)
for (i in infHouseholds) {
HHI[i,1] <- HHI[i,1]+1
HHS[i,1] <- HHS[i,1]-1
}
#Create Vectors to keep track of Cumulative S, I, R
time <- 0
S.0 <- 0
I.0 <- 0
R.0 <- 0
S.1 <- 0
I.1 <- 0
R.1 <- 0
for (i in 1:N.HH) { #Input the Initial Values for Each
if (sanitation[i]==0) {
S.0 <- S.0 + HHS[i]
I.0 <- I.0 + HHI[i]
R.0 <- R.0 + HHR[i]
}
if (sanitation[i]==1) {
S.1 <- S.1 + HHS[i]
I.1 <- I.1 + HHI[i]
R.1 <- R.1 + HHR[i]
}
}
Cum.Inf.HH <- 0 #For tracking cumulative infections from the HH Environment
Cum.Inf.E <- 0 #For tracking cumulative infections from the shared
Environment
time.i <- 0 #Start at time=0
iter <- 0
max.time <- 500 #Maximum Number of Days to Run the Simulation
while (time.i < max.time & sum(HHS[,1])>0 & (sum(HHI[,1])>0 | sum(HHE[,1])>0
| E[1]>0)) {
iter <- iter + 1
#Create Vector of Rates
recovery <- vector()
newInfectionHH <- vector()
newInfectionE <- vector()
for (i in 1:N.HH) {
recovery[i] <- gamma[i] * HHI[i]
newInfectionHH[i] <- HHS[i]*parm['pi']*rho[i]*HHE[i,1]
newInfectionE[i] <- HHS[i]*parm['pi']*alpha[i]*E[1]
}
#Create Vector of the rates of all events
all.events <- c(recovery, newInfectionHH, newInfectionE)
#When does the next event happen?
tau <- rexp(1, rate=sum(all.events))
#Update Time
time.i <- time.i + tau
#What Event Occured?
event <- sample(length(all.events),1,prob=all.events)
#Update States of People
#Create a vector of State Transitions based on events
#Corresponds with the vector 'all.events'
```

```
HHS[,2] <- HHS[,1]
HHI[,2] <- HHI[,1]
HHR[,2] <- HHR[,1]
if (event >=1 & event <= N.HH) { #Recovery
HHI[event,2] <- HHI[event,1] - 1
HHR[event,2] <- HHR[event,1] + 1
}
if (event >= N.HH+1 & event <= N.HH*2) { #New Infection from HH Environment
HHS[event-N.HH,2] <- HHS[event-N.HH,1] - 1
HHI[event-N.HH,2] <- HHI[event-N.HH,1] + 1
Cum.Inf.HH <- Cum.Inf.HH + 1
}
if (event >= N.HH*2+1 & event <= N.HH*3) { #New Infections from Comm
Environment
HHS[event-2*N.HH,2] <- HHS[event-2*N.HH,1] - 1
HHI[event-2*N.HH,2] <- HHI[event-2*N.HH,1] + 1
Cum.Inf.E <- Cum.Inf.E + 1
}
#Update Values for Environmental Compartments
for (i in 1:N.HH) {
HHE[i,2] <- HHE[i,1]*exp(-mu.HH[i]*tau) + (delta[i]*HHI[i]/mu.HH[i])*(1-exp(-
mu.HH[i]*tau))
}
E[2] <- E[1]*exp(-mu.E*tau) + (sum(phi*HHI[,1])/mu.E)*(1-exp(-mu.E*tau))
#Update Cumulative S, I, R Tracker
time[iter+1] <- time.i
S.0[iter+1] <- 0
I.0[iter+1] <- 0
R.0[iter+1] <- 0
S.1[iter+1] <- 0
I.1[iter+1] <- 0
R.1[iter+1] <- 0
for (i in 1:N.HH) { #Input the Initial Values for Each
if (sanitation[i]==0) {
S.0[iter+1] <- S.0[iter+1] + HHS[i,2]
I.0[iter+1] <- I.0[iter+1] + HHI[i,2]
R.0[iter+1] <- R.0[iter+1] + HHR[i,2]
}
if (sanitation[i]==1) {
S.1[iter+1] <- S.1[iter+1] + HHS[i]
I.1[iter+1] <- I.1[iter+1] + HHI[i]
R.1[iter+1] <- R.1[iter+1] + HHR[i]
}
}
HHS[,1] <- HHS[,2]
HHI[,1] <- HHI[,2]
HHR[,1] <- HHR[,2]
HHE[,1] <- HHE[,2]
E[1] <- E[2]
}
#Create data frame of cumulative incidence for each run
data.row = c(coverage=coverage, run=run,
risk.overall=(S.0[1]+S.1[1]-tail(S.0+S.1,n=1))/(S.0[1]+S.1[1]),
risk.0=(S.0[1]-tail(S.0,n=1))/S.0[1],
risk.1=(S.1[1]-tail(S.1,n=1))/S.1[1],
Inf.HH=Cum.Inf.HH,
Inf.E=Cum.Inf.E,
```

```
Inf.Total=Cum.Inf.HH+Cum.Inf.E,
prop.HH=Cum.Inf.HH/(Cum.Inf.HH+Cum.Inf.E),
prop.E=Cum.Inf.E/(Cum.Inf.HH+Cum.Inf.E),
rr.direct=NA,
rr.indirect=NA,
rr.total=NA,
rr.overall=NA)
data.total <- rbind(data.total, data.row)
} #Close 'Number of Runs' Loop
} #Close Coverage Level Loop
data.total <- data.total[-1,]
#Calculate Relative Risks
data.total$rr.direct <- data.total$risk.0 / data.total$risk.1
for (c in coverage.levels) {
for (r in 1:N.runs) {
data.total$rr.indirect[data.total$coverage==c & data.total$run==r] <-
data.total$risk.0[data.total$coverage==c & data.total$run==r] /
data.total$risk.0[data.total$coverage==0 & data.total$run==r]
data.total$rr.total[data.total$coverage==c & data.total$run==r] <-
data.total$risk.1[data.total$coverage==c & data.total$run==r] /
data.total$risk.0[data.total$coverage==0 & data.total$run==r]
data.total$rr.overall[data.total$coverage==c & data.total$run==r] <-
data.total$risk.overall[data.total$coverage==c & data.total$run==r] /
data.total$risk.overall[data.total$coverage==0 & data.total$run==r]
}
}
write.csv(data.total,file="Figure 3 - Water (Total).csv")
#Calculate Medians
data.medians <- data.frame(coverage=NA, run=NA,
risk.overall=NA, risk.0=NA, risk.1=NA,
Inf.HH=NA, Inf.E=NA, Inf.Total=NA, prop.HH=NA, prop.E=NA,
rr.direct=NA, rr.indirect=NA, rr.total=NA, rr.overall=NA)
obs <- 0
for(c in 1:length(coverage.levels)) {
obs <- obs+1
data.medians[obs,] <- NA
data.medians$coverage[obs] <- coverage.levels[c]
data.medians$risk.overall[obs] <-
median(data.total$risk.overall[which(data.total$coverage==coverage.levels[c])
], na.rm=T)
data.medians$risk.0[obs] <-
median(data.total$risk.0[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
data.medians$risk.1[obs] <-
median(data.total$risk.1[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
data.medians$Inf.HH[obs] <-
median(data.total$Inf.HH[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
data.medians$Inf.E[obs] <-
median(data.total$Inf.E[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
data.medians$Inf.Total[obs] <-
median(data.total$Inf.Total[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
```

```
data.medians$prop.HH[obs] <-
median(data.total$prop.HH[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
data.medians$prop.E[obs] <-
median(data.total$prop.E[which(data.total$coverage==coverage.levels[c])],
na.rm=T)
}
data.medians$rr.direct <- data.medians$risk.1 / data.medians$risk.0
for (c in coverage.levels) {
data.medians$rr.indirect[data.medians$coverage==c] <-
data.medians$risk.0[data.medians$coverage==c] / data.medians$risk.0[1]
data.medians$rr.total[data.medians$coverage==c] <-
data.medians$risk.1[data.medians$coverage==c] / data.medians$risk.0[1]
data.medians$rr.overall[data.medians$coverage==c] <-
data.medians$risk.overall[data.medians$coverage==c] /
data.medians$risk.overall[1]
}
#Save Output Data as Local Files
write.csv(data.medians,file="Figure 3 - Water (Medians).csv")
save(data.total, data.medians,Io, max.time, N.HH, N.Pop, N.Pop.Actual,
N.runs,
parm, sanitation,
file="Figure 3 - Water.RData")
#Make Plot Window
rm(list = ls())
load("Figure 3 - Water.RData")
dev.new(width=7.5, height=3)
par(mfrow=c(1,2))
par(mar=c(3,3,1,1))
lw <- 1.5
#Plot of the Median Risk
plot(0,0,pch=NA, ylim=c(0,1), xlim=c(0,100), xlab="", ylab="", main="")
lines(data.medians$coverage*100, data.medians$risk.0, col='black', lty=2,
lwd=lw)
lines(data.medians$coverage*100, data.medians$risk.1, col='black', lty=1,
lwd=lw)
points(data.medians$coverage[1]*100, data.medians$risk.0[1], pch=17, cex=1.5)
mtext(side=1, text="Intervention Coverage (%) in the Community", line=2,
cex=0.75)
mtext(side=2, text="Proportion Infected", line=2, cex=0.75)
text(100,1,"A",adj=c(1,1),cex=2)
legend("bottomleft",c(expression(paste("D" ["0*"])),
expression(paste("Control Group (D" ["0"], ")")),
expression(paste("Intervention Group (D" ["1"], ")"))),
col='black', lty=c(NA,2,1), lwd=c(NA,1.2,1.2), pch=c(17,NA,NA),
pt.cex=c(1.5,NA,NA), cex=0.8, bty="n")
#Plot the Protective Efficacy
plot(data.medians$coverage*100, (1-data.medians$rr.total)*100, col='black',
type='l', lty=2, lwd=lw, ylab=NA, xlab=NA, xlim=c(0,100), ylim=c(0,50))
lines(data.medians$coverage*100, (1-data.medians$rr.direct)*100, col='black',
lty=1, lwd=lw)
lines(data.medians$coverage*100, (1-data.medians$rr.indirect)*100,
col='gray60', lty=1, lwd=lw)
lines(data.medians$coverage*100, (1-data.medians$rr.overall)*100,
col='gray60', lty=2, lwd=lw)
mtext(side=1, text="Intervention Coverage (%) in the Community", line=2,
cex=0.75)
```

```
mtext(side=2, text="Protective Efficacy (%)", line=2, cex=0.75)
text(100,50,"B",adj=c(1,1),cex=2)
legend("topleft",c(expression(paste("Overall Effect (1-(w" ["0"],"D"
["0"],"+w" ["1"],"D" ["1"],")/D" ["0*"],")")),
expression(paste("Total Effect (1-D" ["1"],"/D" ["0*"], ")")),
expression(paste("Direct Effect (1-D" ["1"],"/D" ["0"], ")")),
expression(paste("Indirect Effect (1-D" ["0"],"/D" ["0*"], ")"))),
col=c('gray60','black','black','gray60'), lty=c(2,2,1,1), lwd=1.2, cex=0.8,
bty="n")
```

SUPPLEMENTAL APPENDIX B

SUPPLEMENTAL TABLES FROM THE LITERATURE REVIEW

SUPPLEMENTAL TABLE 1

Details of studies assessing herd protection from drinking water

Setting (reference)	Study design	Outcome	Range of drinking water coverage included in the study
Mortality			
6 Sub-Saharan African countries: urban[1]	Retrospective cohort form 10-year birth history	Infant mortality (< 1 year of age)	0–100%
6 Sub-Saharan African countries: rural[1]	Retrospective cohort form 10-year birth history	Infant mortality (< 1 year of age)	0–100%
Nigeria: national[2]	Retrospective cohort from 5-year birth history	Infant mortality (0–11 months of age)	Not provided
Nigeria: national[2]	Retrospective cohort from 5-year birth history	Child mortality (12–59 months of age)	Not provided
Brazil: northeast[3]	Retrospective cohort from 5-year birth history	Child mortality (< 5 years of age)	Not provided, but mean is 30.1% and standard deviation is 24.5
Brazil: south and southeast[3]	Retrospective cohort from 5-year birth history	Child mortality (< 5 years of age)	Not provided, but mean is 54.5% and standard deviation is 26.7
Nutrition			
Nigeria: national[4]	Cross-sectional	Childhood stunting (height-for-age z-score < –2)	Not provided
Malawi: national[5]	Cross-sectional	Childhood stunting (height-for-age z-score < –2)	Not provided
Ghana: national[6]	Cross-sectional	Low birth weight (smaller than average)	Not provided

Studies are grouped by outcome and presented in the same order as Table 2.

SUPPLEMENTAL TABLE 2

Details of studies assessing herd protection from sanitation

Setting (reference)	Study design	Outcome	Range of sanitation coverage included in the study
Diarrhea			
Brazil: urban[7]	Before and after intervention (2 distinct cohorts)	Diarrhea prevalence in the previous 3–4 days among children < 3 years of age	26–80%
India: rural[8]	Cross-sectional	Diarrhea prevalence in the previous 2 weeks among children < 4 years of age	0–100%
Zimbabwe: rural[9]	Cohort	No. of episodes of diarrhea in 45 weeks among children < 5 years of age at baseline	Not applicable
Brazil: urban[10]	Cohort	Duration in days of diarrheal episodes	Not reported, though at least 34.1–55.1%
Parasitic infection			
Brazil: urban[11]	Before and after interventions (2 distinct cross sections)	Prevalence of *Ascaris lumbricoides*	26–80%
Brazil: urban[11]	Before and after interventions (2 distinct cross sections)	Prevalence of *Trichuris trichiura*	26–80%
Brazil: urban[11]	Before and after interventions (2 distinct cross sections)	Prevalence of *Giardia duodenalis*	26–80%
Mortality			
6 Sub-Saharan African countries: urban[1]	Retrospective cohort form 10-year birth history	Infant mortality (< 1 year of age)	0–100%
6 Sub-Saharan African countries: rural[1]	Retrospective cohort form 10-year birth history	Infant mortality (< 1 year of age)	0–100%
India: national[12]	Retrospective cohort from 10-year birth history	Infant mortality (< 1 year of age)	0–100%
Brazil: northeast[3]	Retrospective cohort from 5-year birth history	Child mortality (< 5 years of age)	Not provided, but mean is 4.23 with standard deviation of 10.4
Brazil: south and southeast[3]	Retrospective cohort from 5-year birth history	Child mortality (< 5 years of age)	Not provided, but mean is 29.3% with standard deviation of 28.2
Nutrition			
Ecuador: rural[13]	Cohort study	Stunting (height-for-age z-score < −2) among children < 5 years of age	0–100%
Nigeria: national[4]	Cross-sectional	Stunting (height-for-age z-score < −2) among children < 5 years of age	Unknown
Guatemala: urban[14]	Cross-sectional	Stunting (length-for-age z-score < −2) among children 6–36 months of age	0–100%
Guatemala: rural[14]	Cross-sectional	Stunting (length-for-age z-score < −2) among children 6–36 months of age	0–100%
Bangladesh: urban[15]	Cohort	Wasting (weight-for-height z-score) in children 0–35 months of age at baseline	About 18% to about 93%

Studies are grouped by outcome and presented in the same order as Table 3.

SUPPLEMENTAL REFERENCES

<jrn>1. Van de Poel E, O'donnell O, Van Doorslaer E, 2009. What explains the rural-urban gap in infant mortality: household or community characteristics? *Demography 46:* 827–850.</jrn>

<jrn>2. Adedini SA, Odimegwu C, Imasiku EN, Ononokpono DN, Ibisomi L, 2015. Regional variations in infant and child mortality in Nigeria: a multilevel analysis. *J Biosoc Sci 47:* 165–187.</jrn>

<jrn>3. Sastry N, 1996. Community characteristics, individual and household attributes, and child survival in Brazil. *Demography 33:* 211–229.</jrn>

<jrn>4. Adekanmbi VT, Kayode GA, Uthman OA, 2013. Individual and contextual factors associated with childhood stunting in Nigeria: a multilevel analysis. *Matern Child Nutr 9:* 244–259.</jrn>

<jrn>5. Chikhungu LC, Madise NJ, Padmadas SS, 2014. How important are community characteristics in influencing children's nutritional status? Evidence from Malawi population-based household and community surveys. *Health Place 30:* 187–195.</jrn>

<jrn>6. Kayode GA, Amoakoh-Coleman M, Agyepong IA, Ansah E, Grobbee DE, Klipstein-Grobusch K, 2014. Contextual risk factors for low birth weight: a multilevel analysis. *PLoS One 9:* e109333.</jrn>

<jrn>7. Barreto ML, Genser B, Strina A, Assis AMO, Rego RF, Teles CA, Prado MS, Matos SM, Santos DN, dos Santos LA, 2007. Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort studies. *Lancet 370:* 1622–1628.</jrn>

<bok>8. Andres LA, Briceño B, Chase C, Echenique JA, 2014. *Sanitation and Externalities: Evidence from Early Childhood Health in Rural India*. Washington, DC: World Bank.</bok>

<jrn>9. Root GP, 2001. Sanitation, community environments, and childhood diarrhoea in rural Zimbabwe. *J Health Popul Nutr 19:* 73–82.</jrn>

<jrn>10. Santos CA, Strina A, Amorim LD, Genser B, Assis AM, Prado MS, Barreto ML, 2012. Individual and contextual determinants of the duration of diarrhoeal episodes in preschool children: a longitudinal study in an urban setting. *Epidemiol Infect 140:* 689–696.</jrn>

<jrn>11. Barreto ML, Genser B, Strina A, Teixeira MG, Assis A, Rego RF, Teles CA, Prado MS, Matos S, Alcântara-Neves NM, 2011. Impact of a city-wide sanitation programme in northeast Brazil on intestinal parasites infection in young children. *Environ Health Perspect 118:* 1637–1642.</jrn>

<other>12. Geruso M, Spears D, 2015. *Neighborhood Sanitation and Infant Mortality*. Working paper 21184. Cambridge, MA: National Bureau of Economic Research.</other>

<jrn>13. Fuller JA, Villamor E, Cevallos W, Trostle J, Eisenberg JNS, 2016. I get height with a little help from my friends: herd protection from sanitation on child growth in rural Ecuador. *Int J Epidemiol 45:* 460–469.</jrn>

<bok>14. Bateman OM, Smith S, 1991. *A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala*. Anonymous WASH field report. WASH.</bok>

<conf>15. Bateman OM, Smith S, 1991. *A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala*. The Demographic and Health Surveys World Conference, Volume 2, August 5–7, 1991, Columbia, MD.</conf>

Figure 1

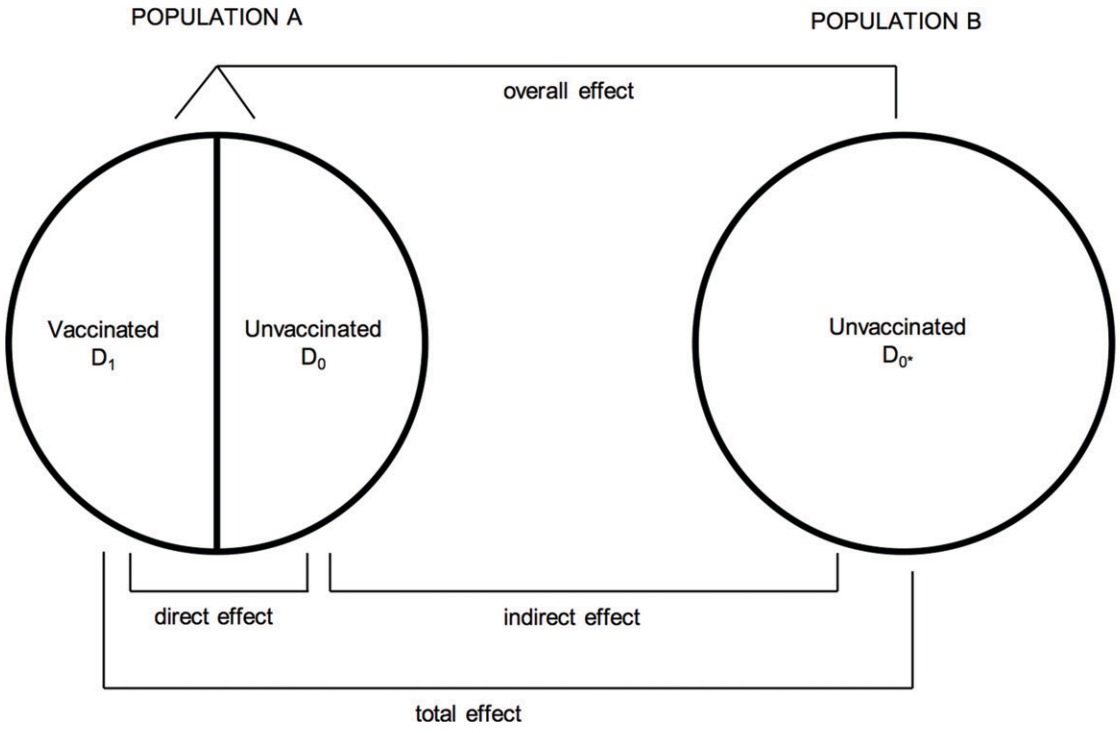

POPULATION A
POPULATION B
overall effect
Vaccinated
D_1
Unvaccinated
D_0
Unvaccinated
D_{0^*}
direct effect
indirect effect
total effect

Figure 2

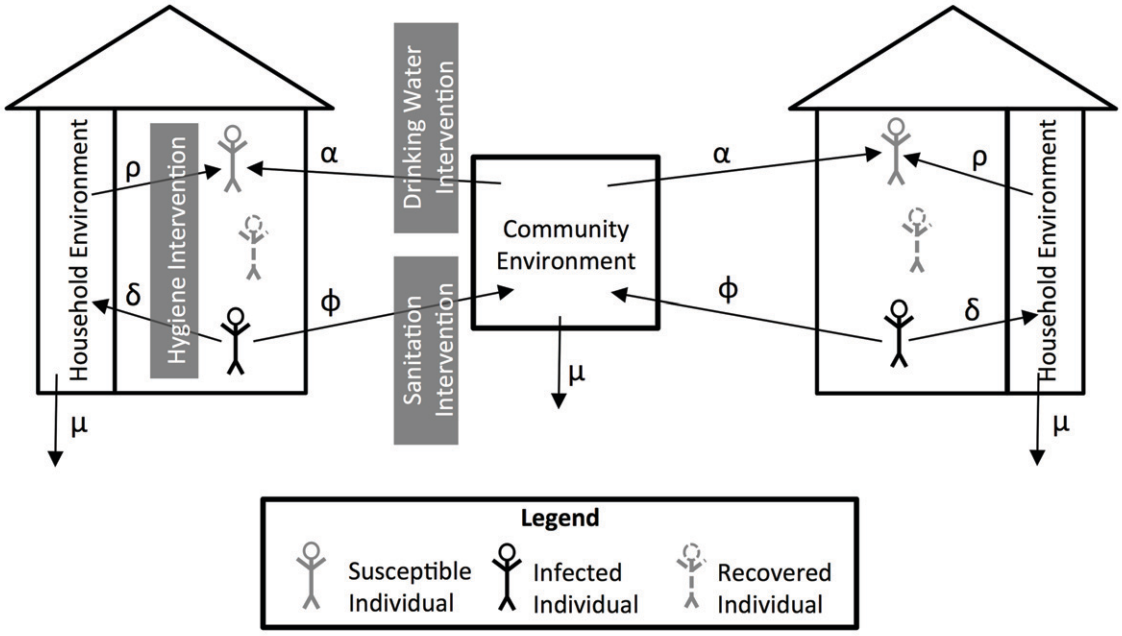

Household Environment
Hygiene Intervention
Drinking Water Intervention
Sanitation Intervention
Community Environment
Household Environment
ρ
α
δ
ϕ
μ
Legend
Susceptible Individual
Infected Individual
Recovered Individual

Figure 3

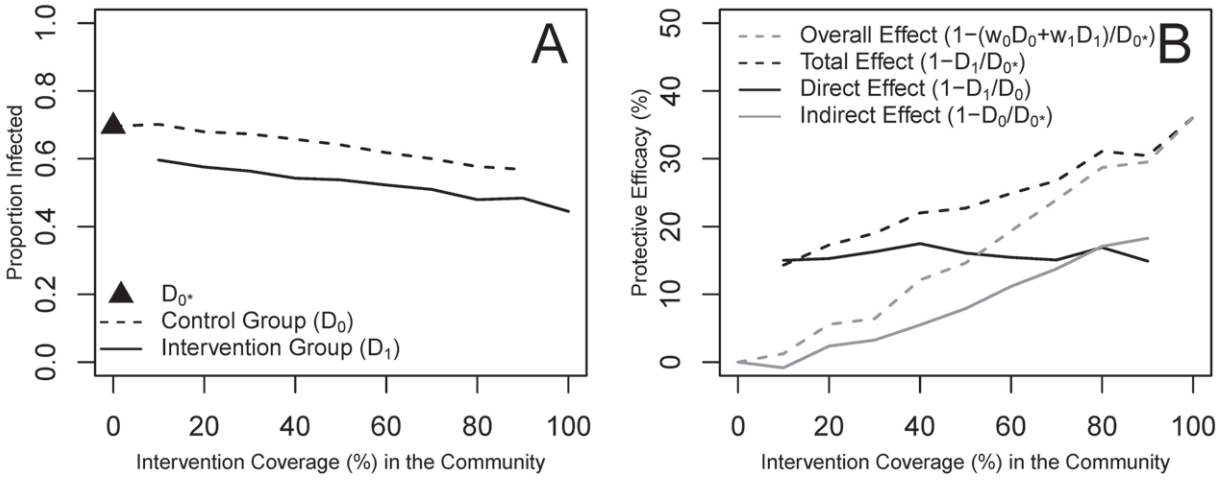

A
Proportion Infected
0.0
0.2
0.4
0.6
0.8
1.0
D_{0^*}
Control Group (D_0)
Intervention Group (D_1)
0
20
40
60
80
100
Intervention Coverage (%) in the Community
B
Protective Efficacy (%)
0
10
20
30
40
50
Overall Effect ($1-(w_0D_0+w_1D_1)/D_{0^*}$)
Total Effect ($1-D_1/D_{0^*}$)
Direct Effect ($1-D_1/D_0$)
Indirect Effect ($1-D_0/D_{0^*}$)
0
20
40
60
80
100
Intervention Coverage (%) in the Community

Figure 4

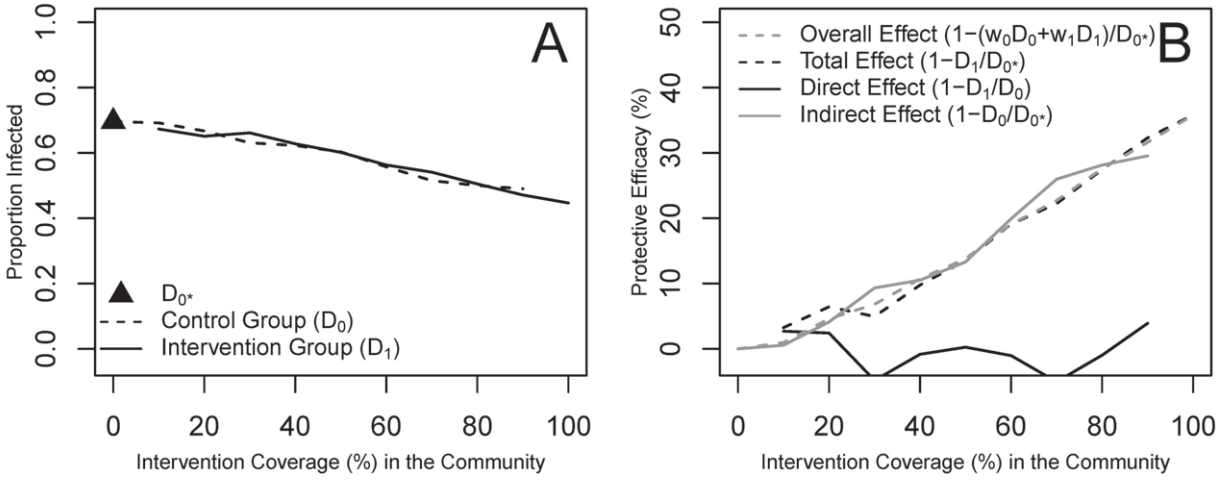
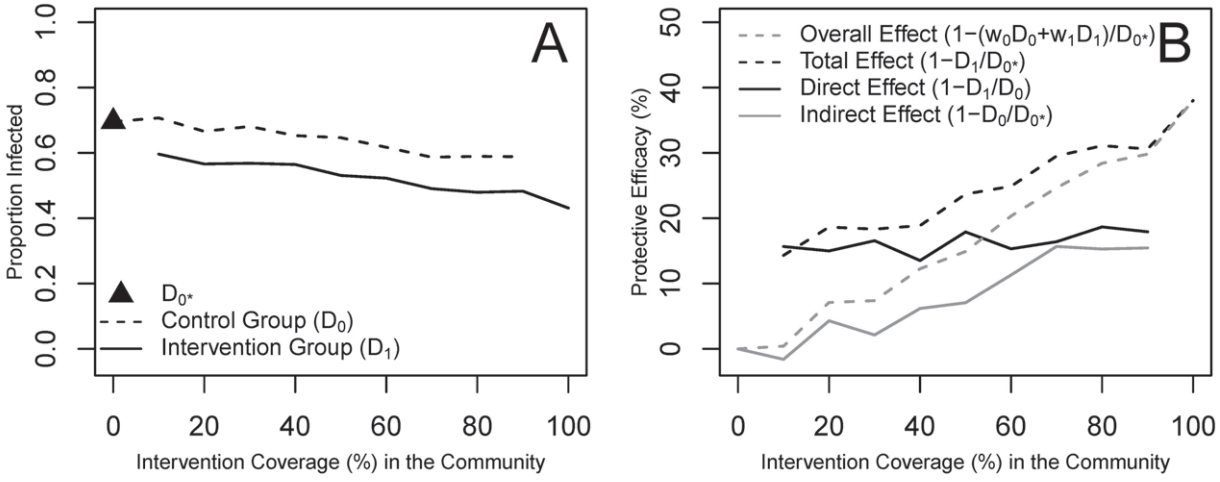

A
Proportion Infected
0.0
0.2
0.4
0.6
0.8
1.0
D_{0^*}
Control Group (D_0)
Intervention Group (D_1)
0
20
40
60
80
100
Intervention Coverage (%) in the Community
B
Protective Efficacy (%)
0
10
20
30
40
50
Overall Effect ($1-(w_0D_0+w_1D_1)/D_{0^*}$)
Total Effect ($1-D_1/D_{0^*}$)
Direct Effect ($1-D_1/D_0$)
Indirect Effect ($1-D_0/D_{0^*}$)
0
20
40
60
80
100
Intervention Coverage (%) in the Community

Figure 5


A
Proportion Infected
0.0
0.2
0.4
0.6
0.8
1.0
D_{0^*}
Control Group (D_0)
Intervention Group (D_1)
0
20
40
60
80
100
Intervention Coverage (%) in the Community
B
Protective Efficacy (%)
0
10
20
30
40
50
Overall Effect ($1-(w_0D_0+w_1D_1)/D_{0^*}$)
Total Effect ($1-D_1/D_{0^*}$)
Direct Effect ($1-D_1/D_0$)
Indirect Effect ($1-D_0/D_{0^*}$)
0
20
40
60
80
100
Intervention Coverage (%) in the Community