Title

Innovative Randomized Controlled Trial Design and Analysis Methods to Account for and Examine Variations in Population Health in Low and Middle-Income Countries

Author

Butzin-Dozier, Zachary

Publication Date

2022

Peer reviewed|Thesis/dissertation

Innovative Randomized Controlled Trial Design and Analysis Methods to Account for and Examine Variations in Population Health in Low and Middle-Income Countries

By

Zachary Butzin-Dozier

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Colford, Chair
Professor Alan Hubbard
Professor Jay Graham
Dr. Andrew Mertens

Summer 2022

Abstract

Innovative Randomized Controlled Trial Design and Analysis Methods to Account for and Examine Variations in Population Health in Low and Middle-Income Countries

by

Zachary Butzin-Dozier

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor John Colford, Chair

Reliance on traditional research methods in situations where innovative methodologies are superior can hinder inference. In global health research, where wasted resources can be equated with lives lost, non-adaptation of promising methods has dire consequences. Given the high cost of randomized controlled trials (RCTs), one major obstacle in these studies is the Type II error, in which an investigator fails to reject the null hypothesis when the null hypothesis is false. Unexpected disease transmission rates may lead studies to be underpowered and yield Type II errors. Furthermore, even if an RCT correctly rejects or fails to reject the null, data may not be fully utilized if investigators only analyze and report the intervention effect in the total study sample. Only focusing on quantifying the overall impact of the intervention of interest may neglect investigation into important observational relationships between covariates of interest or heterogeneity of the intervention effect in population subgroups, although these observational analyses and subgroups of interest should be pre-specified to maintain transparency and replicability. As participants were not randomized to receive exposures represented by covariate values, additional analytic rigor and inferential caution must be exercised in order to enable meaningful interpretation. My dissertation aims to explore three examples of potential methodological solutions to these challenges to inference in RCTs: a systematic review of the ring trial design, a targeted learning analysis of treatment heterogeneity, and an observational analysis of multiple biomarkers using trial data.

Chapter 1. Spatiotemporal clustering: ring trials. For cluster RCTs in emergent and elimination disease settings, unpredictable spatiotemporal clustering often leads to imbalanced clusters, or even clusters with zero incident cases, which can reduce study power and limit inference. Ring trials, a trial design in which the units of randomization are responsively-defined clusters of individuals in social or physical proximity to an index case, may improve investigators' ability to make inferences in these settings. Despite this potential utility, this RCT design remains under-examined and under-utilized. Investigators conducted a systematic review of the ring trial design to examine the existing applications of this study design as well as its benefits and drawbacks. We identified 26 ring trials, 15 cluster-randomized trials that used ring interventions, five trials that used ring recruitment and randomized within rings, and one individually-randomized trial that used a ring intervention. Ring trial designs require strong disease surveillance and contact tracing mechanisms, rapid intervention delivery systems, and a treatment with a strong post-exposure prophylactic effect. In these settings, ring trials can retain power despite unpredictable spatiotemporal clustering of the outcome of interest.

Chapter 2. Heterogeneity in treatment effect: targeted learning analysis of treatment heterogeneity. Even if there is an effect of the intervention on the outcome among certain individuals, a study may fail to detect this relationship if the effect is heterogeneous in the study sample. Investigators can gain additional insight through analysis of the conditional average treatment effect, which is the treatment effect based on individual covariate status. Using data from the WASH Benefits study, which enrolled pregnant mothers and young children in rural Bangladesh, analysis of treatment heterogeneity can improve our understanding of child growth in low and middle-income countries. Despite the widespread use of water, sanitation, hygiene (WSH), nutrition (N), and combined (N+WSH) interventions, investigators have found mixed evidence regarding these interventions' impact on child growth and development. Insufficient reduction of pathogens may explain WSH's lack of impact, and environmental enteric dysfunction (EED), a condition of impaired intestinal permeability and inflammation, may modify the impact of WSH and N interventions on child growth. This study applied targeted machine learning methods to assess treatment heterogeneity of N+WSH, WSH, and N interventions on child growth by pathogen and EED biomarker status. We found that children with greater levels of myeloperoxidase, a gut inflammation biomarker associated with EED, and *Campylobacter*, a genus of bacteria that is associated with EED onset, had a greater effect of all treatments on growth. These results contribute to the body of literature characterizing individual predictors of N+WSH, WSH, and N intervention effectiveness as well as our understanding of EED.

Chapter 3. Maximizing data utilization: observational analysis of high-dimensional data nested within a randomized controlled trial. Trials devote enormous resources to evaluating the effect of the randomized intervention in the study sample, but limiting analyses to only include this intervention may neglect the wealth of data that these RCTs can provide. In addition to analysis of the effect of the intervention, investigators can conduct observational analyses nested within an RCT, although these analyses will require additional methodological rigor in order to limit confounding and bias to enable meaningful inference. Data from the WASH Benefits study provide an opportunity to assess the relationship between stress neurobiology, an exposure that could not be ethically randomized, and child development. Stress has been implicated as a key pathway by which adverse circumstances can lead to developmental impairment, and prior studies have indicated a possible link between stress and subsequent development. This study evaluated the relationship between stress and development through an observational analysis nested within an RCT. We assessed physiologic measures of stress using measures of the hypothalamic-pituitary-adrenal (HPA) axis, the sympathetic-adrenal-medullary (SAM) system, and oxidative status. We constructed generalized additive models to compare development outcomes of children at the 75th and 25th percentiles of stress biomarker distributions while adjusting for potential confounders. We found that measures of HPA axis activity were associated with poor development outcomes. These observations support the use of HPA axis biomarkers, particularly cortisol and glucocorticoid receptor methylation, to indicate children who are at risk of poor developmental outcomes.

This study explores and provides applied examples of RCT design and analysis methods that may improve efficiency in global epidemiologic research. This research serves to 1) improve our understanding of a neglected trial design and explore innovative methods of analyzing trial data; 2) identify which characteristics define amenability to N+WSH, WSH, and N interventions, providing insights that EED may be associated with treatment effectiveness; 3) evaluate the

relationship between stress biomarkers and child development through observational analyses nested within an RCT, which supports the use of HPA axis biomarkers to indicate children at risk for poor developmental outcomes.

TABLE OF CONTENTS

I. DEDICATION

For my mom, my dad, and Lauren.

II. ACKNOWLEDGEMENTS

I want to thank my mom, who taught me to never question what I could accomplish, and my dad, who taught me to love numbers and question the world around me. I could not have completed this without my partner, Lauren, who supported me through this work.

I am deeply grateful for the support and guidance of my advisor, Jack, who has been an incredible mentor through my graduate studies. He has shaped my outlook as a researcher and has shown me how my research can build upon my impact as an educator.

I thank Audrie Lin, Andrew Mertens, and Jade Benjamin-Chung for their guidance through the conceptualization, design, analysis, and interpretation of these chapters.

I would like to thank my dissertation committee members, Jack Colford, Alan Hubbard, Jay Graham, and Andrew Mertens, for their support, feedback, and mentorship.

This research would not have been possible without the support and dedication of a robust team of experts across a range of disciplines who bolstered my belief in the importance of team science. I would like to thank Jade Benjamin-Chung, Tejas S. Athni, Yunwen (Wendy) Ji, Ivana Malenica, Jeremy Coyle, Jess Grembi, Liz Rogawski, James A Platt-Mills, Md. Ziaur Rahman, Shahjahan Ali, Jie Liu, Abul K Shoab, Syeda Luthfa Famida, Mami Taniuchi, Mahbubur Rahman, Eric Houpt, Mohammad Alauddin, Sarah T. Alauddin, Palash Mutsuddi, Md Ohedul Islam, Md Rana Miah, Salma Akther, Jack Colford, Jay Graham, Alan Hubbard, Andrew Mertens, Audrie Lin, Sophia Tan, Douglas A. Granger, Helen O. Pitchik, Dora Il'yasova, Fahmida Tofail, Ivan Spasojevic, Amy Li, Idan Shalev, Benjamin F. Arnold, Lisa Kim, Mohammed Rabiul Karim, Sunny Shahriar, Gabrielle Shuman, Md. Saheen Hossen, Leanne Unicomb, Kishor K. Das, Liying Yan, Ann Meyer, Patricia Kariger, Christine P. Stewart, Ruchira Tabassum Naved, Kausar Parvin, Md. Mahfuz Al Mamun, Stephen P. Luby, and Lia C. H. Fernald for their contributions to and support of this work.

Finally, I am grateful to the participants of the WASH Benefits Bangladesh study and the participants of each reviewed randomized controlled trial.

III. INTRODUCTION

Randomized controlled trials: benefits and challenges

Randomized controlled trials are justifiably considered the gold standard for evidence in epidemiologic research (1). A primary goal in public health research is causal inference. Investigators aim to approximate the counterfactual, which is the outcome an individual or population would have experienced if they, contrary to fact, had not experienced the exposure of interest, but all other factors remained the same (2). This is not possible to observe, as it requires observing two parallel realities for a single individual or population over the same time period. But, randomized controlled trials enable investigators to approximate the counterfactual outcomes for a population through randomized assignment to intervention. The unique benefit that randomized controlled trials confer, in comparison to observational study designs, is that adequate randomized assignment will balance all measured and unmeasured confounders between treated and untreated groups (3). While other design phase (e.g., restriction and stratification) or analysis phase (e.g., multivariate adjustment) techniques may control for confounding by observed characteristics, only randomization is able to control for unobserved characteristics. Despite this major benefit, randomized controlled trials have several limitations, including their high cost, limited generalizability, and inability to randomize certain exposures due to ethical or feasibility constraints (1).

Randomly assigning participants to receive an intervention or control requires that equipoise is met (4). Equipoise refers to genuine uncertainty regarding whether one intervention is preferable to another (or control). This presents an inferential challenge to many epidemiologic research questions, where observational or biologic evidence points to a deleterious effect of an exposure, violating equipoise, but further investigation is needed to understand the magnitude or pathways of these effects. Even if equipoise is met, some exposures are inherently nonmodifiable. For example, it would be impossible to assign a participant to have certain genetic characteristics. In these settings, nested observational analyses within randomized trials can utilize the rigorous data collection of randomized trials and leverage variations in sample characteristics due to randomized intervention.

A key component of randomized trials is the unit of randomization, and improper definition of this unit may lead to failed randomization. In an individually-randomized trial, investigators will assign each participant to intervention or control one by one. In some settings, particularly for infectious diseases, those assigned to control may experience benefits of the intervention, also known as contamination (5). For example, a child assigned to receive deworming medication (intervention) will likely provide some protection against worms for all children in their household, even if other household members were assigned to receive no treatment (control). Cluster-randomized trials can address this challenge by using clusters (e.g., households, schools, villages, etc.) as the unit of randomization rather than the individual (5–7). Appropriately-defined clusters can reduce the possibility of contamination between intervention and control groups (5). The use of buffer zones, or a required distance between included clusters, can further reduce the risk of contamination (6–8). The assessment of spillover effects, or the impact of an intervention on untreated individuals, can quantify potential contamination and the impact of an intervention on the broader population, which may be relevant to policymakers (9).

In traditional cluster-randomized trials, clusters are defined prior to randomization of participants. One limitation of this approach is that disease incidence may be imbalanced between treatment arms at baseline (prior to the intervention), and some clusters may even have zero incidence at baseline (10). This spatiotemporal clustering of disease, which is particularly common for infectious diseases in emergent and elimination settings, can severely limit study power (11,12). An underpowered trial is unable to make meaningful inference regarding the effectiveness of the intervention and leads to wasted resources. The ring trial design can overcome this challenge by reactively defining the units of randomization as clusters of contacts surrounding an index case (an individual who experienced an incident case of the outcome of interest). For example, the Ebola Vaccine Trial, which was the first applied example of a ring trial design, defined the unit of randomization (the ring) as all individuals who were in close contact with an index case of Ebola Virus Disease as well as these contacts' contacts and assigned each ring to receive Ebola Virus vaccination or control (delayed vaccination) (13,14).

Targeted learning and optimal individualized treatment regime analysis

The goal of public health research is typically to identify population-level, rather than individual, causes of disease (15). With rising interest in targeted interventions, where susceptible individuals can be targeted for precision interventions, public health officials aim to utilize individual characteristics to maximize population health outcomes (16,17). In order to evaluate the importance of individual characteristics, we can estimate the conditional average treatment effect to assess treatment heterogeneity based on these factors (18). Although there has been rising interest in these targeted interventions, previous work has frequently depended on parametric assumptions (19–24). These parametric assumptions may be violated, leading to biased estimates, but targeted learning methods may reduce this vulnerability (25). Targeted maximum likelihood estimation can estimate the maximum likelihood of a parameter of interest while reducing bias and variance in finite samples, and cross validation can prevent overfitting (26). These estimates are optimally efficient, and the methods are doubly-robust, so estimates will be unbiased if either the outcome regression or treatment mechanism is estimated consistently (26,27). In Chapter 2, we use the targeted maximum likelihood estimator for mean child growth under optimal individualized treatment in order to attain valid inference with minimal parametric assumptions (25).

Child health in Bangladesh: growth, development, and stress

Globally, approximately 156 million experience growth faltering, and more than 250 million children in low and middle-income countries (LMIC) are at risk of failing to meet their developmental potential (28,29). Children who experience growth impairment are more likely to have low educational attainment and low income in adulthood, and are more likely to have children with these characteristics, perpetuating the cycle of poverty (30,31). In addition, poor development in early life can have lifelong implications (30,32,33). Investigators have identified early childhood as an important time to intervene on risk factors for poor developmental status (34). In order to limit the incidence of developmental and growth impairment in LMICs, investigators can evaluate the causes of these outcomes and develop methods to identify children who are at risk of these outcomes.

Stress neurobiology is a key component of child development. Chronic exposure to stress early in life can impact multiple biological systems related to memory, sleep, metabolism, and mental health (35,36). Children in rural areas of LMICs suffer disproportionate exposure to these stressors, as children in poor, rural communities endure additional biological, environmental, and psychosocial stressors, which are risk factors for poor developmental status (37).

Water, sanitation, hygiene, and nutrition

Lack of access to adequate water, sanitation, hygiene (WSH) and nutrition is another risk factor associated with poor growth and development, and global health experts have aimed to improve child health by intervening on these factors.

As one of its Sustainable Development Goals, the United Nations hopes that WSH access will be universal by the year 2030 (38). Despite the broad promotion of these practices, the relationship between WSH and child growth is unclear. While observational studies indicated a positive relationship between household WSH and child growth, the WASH Benefits study, a randomized controlled trial that enrolled pregnant mothers in Bangladesh and Kenya, and SHINE, a trial in Zimbabwe, found that household WSH interventions did not improve child growth (39,29,31,40,41). Randomized controlled trials in LMICs have found that nutrition (N) interventions early in life can improve child growth (31,42,43). The WASH Benefits study found that its N intervention led to a modest improvement in child growth, but that the combined nutrition and WSH (N+WSH) intervention conferred no additional benefit with respect to child growth (40).

One key hypothesized mechanism of WSH intervention effectiveness is through reduction of exposure to pathogens. The WASH Benefits study found that 99% of children in its sample were infected with at least one enteropathogen (44). Investigators found that the WSH intervention reduced viral infection at 14 months of age, relative to control, although they did not detect a significant reduction in bacteria or parasites (44). Environmental enteric dysfunction (EED) is a condition of increased gut permeability and systemic inflammation that may be caused by chronic exposure to pathogens, and this condition has been hypothesized to be associated with the effectiveness of nutrition interventions (45,46). Investigators hypothesize that WSH intervention may reduce EED by reducing pathogenic exposure (45,46).

Summary

In this dissertation, I will provide applied examples that address and overcome challenges to inference in RCTs. In the first chapter, I will review the applications and utility of a novel trial design, the ring trial design, which can improve power in settings with high spatiotemporal variation of disease transmission (e.g., infectious diseases in emergent or elimination settings). Given the high cost of trials, ensuring adequate power prevents unnecessarily wasting resources. In the second chapter, I apply targeted learning methods to assess treatment heterogeneity of N+WSH, WSH, and N interventions on growth based on pathogen and EED biomarker status. The identification of factors associated with treatment heterogeneity highlight individual characteristics that define N+WSH, WSH and N treatment effectiveness and strengthen our understanding of EED. In the third chapter, I evaluate the relationship between stress biomarkers and child development in a prospective cohort nested within a randomized controlled trial of

young children in rural Bangladesh. As it would be unethical and unfeasible to randomize children to receive varying levels of stress, this observational analysis of trial data provides an example of how nested analyses can maximize the inference gained from randomized trials. This analysis explores stress biomarkers can indicate children who are at risk for poor development. In addition to subject matter contributions, each chapter shares the goal of providing replicable methods that can maximize the inferences that can be gained from RCT data.

1. Chapter 1. A Review of the Ring Trial Design for Evaluating Ring Interventions for Infectious Diseases

1.1 Abstract

In trials of infectious disease interventions, rare outcomes and unpredictable spatiotemporal variation can introduce bias, reduce statistical power, and prevent conclusive inferences. Spillover effects can complicate inference if individual randomization is used to gain efficiency. Ring trials are a type of cluster-randomized trial that may increase efficiency and minimize bias, particularly in emergency and elimination settings with strong clustering of infection. They can be used to evaluate ring interventions, which are delivered to individuals in proximity to or contact with index cases. Here we review ring trials, compare them to other trial designs for evaluating ring interventions, and describe strengths and weaknesses of each design. We conducted a systematic review to identify trials and trial protocols evaluating ring interventions. Of 849 articles and 322 protocols screened, we identified 26 ring trials, 15 cluster-randomized trials, five trials that randomized households or individuals within rings, and one individually randomized trial. The most common interventions were post-exposure prophylaxis (n = 23) and focal mass drug administration and screening and treatment (n = 7). Ring trials require robust surveillance systems and contact tracing for directly transmitted diseases. For rare diseases with strong spatiotemporal clustering, they may have higher efficiency and internal validity than cluster-randomized designs in part because they ensure that no clusters are excluded from analysis due to zero cluster incidence. Though further research is needed to compare them to other types of trials, ring trials hold promise as a design that can increase trial speed and efficiency while reducing bias.

1.2 Introduction

Infectious disease transmission is inherently heterogenous, with a minority of the population responsible for the majority of transmission (47). This is especially the case in emerging infectious disease and disease elimination settings, where diseases are rare and strongly clustered within space or contact networks (48–50). These epidemiologic features can pose challenges in randomized trials (48).

Strong spatial clustering and unpredictable timing of outbreaks can compromise baseline balance between trial arms, reducing statistical power and face validity (11,12). This is particularly true in cluster-randomized controlled trials (CRCTs), which are commonly used to evaluate infectious disease interventions and enroll fewer units than individually randomized trials. While adjusting for baseline covariates may address baseline imbalance, substantive differences in adjusted and unadjusted estimates may undermine trial credibility and replicability (12).

In addition, in studies at the early or waning stages of an outbreak or in elimination settings, rare, clustered outcomes require large numbers of clusters to minimize false negative results, which may be infeasible and cost-prohibitive (48,51). Individually randomized trials are more efficient than CRCTs, but contamination can prevent valid estimation of the estimand of interest – the effect of individual treatment vs. control (5,6). CRCTs are often used when contamination is a concern (52), among other reasons (e.g., to evaluate group-level interventions or to increase compliance or feasibility) (6). When buffer zones are established between clusters to maintain independence, CRCTs can minimize contamination (6–8). However, when disease is highly

clustered in space or time, disease cases may only occur in a subset of pre-defined clusters, which may compromise statistical power in CRCTs (51).

Diseases that can be subclinical or asymptomatic pose another challenge to trials (48). For example, malaria and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) can be transmitted without symptoms (53), and asymptomatic Zika infection in pregnant women may result in birth defects. For such diseases, it is critical to evaluate asymptomatic infections, but doing so requires outcome measurement in population-based samples instead of or in addition to routine surveillance, which can be difficult and costly.

Ring trials are a type of CRCT that may increase efficiency and minimize bias in emerging infection and disease elimination settings (48). This design is well-suited for evaluations of ring interventions (e.g., "case-area targeted interventions" (54–56), "targeted interventions" (57), "focal interventions" (58–61), "reactive interventions"), which are delivered to individuals in proximity to or contact with index cases. Ring interventions have been proposed or implemented for a wide range of diseases, including smallpox (50), malaria (62), and coronavirus disease 2019 (COVID-19) (63). In ring trials, as index cases are detected, each "ring" of individuals around the index case is randomized. This design was used to evaluate ring vaccination for the Ebola vaccine (13) and may be effective for ring interventions for other infectious diseases with asymptomatic transmission and high spatiotemporal transmission heterogeneity.

Here we review ring trial designs, compare them to traditional trial designs, and discuss optimal settings for their use. We also report the findings of a systematic review of ring trials and trials of ring interventions, including published studies and protocols for ongoing studies.

1.3 Methods

We conducted a narrative review of articles related to ring trials and ring interventions, focusing on methodological papers and simulation studies. To identify empirical studies, we conducted a systematic review to identify all published studies and registered study protocols reporting trials of ring interventions, including ring trials and other types of trials (PROSPERO registration: CRD42021238932). The remainder of this section focuses on the methods we used in the systematic review.

Inclusion and Exclusion Criteria

We included studies that: 1) were reported as a research article or trial protocol; 2) used a ring trial design or other randomized design to evaluate a ring intervention; 3) measured disease or health-related outcomes; 4) evaluated public health intervention(s); 5) enrolled humans; 6) were reported in English; and 7) were published or registered before August 23, 2021.

We defined ring interventions as interventions delivered to neighbors, contacts of index cases, or contacts of contacts of index cases. Index cases may be detected through passive surveillance, in which cases present at health facilities, or active surveillance, in which cases are detected through population screening. Typically, interventions are delivered within a relatively short period after index case detection, when onward transmission to ring members is expected. We distinguished ring interventions from reactive interventions, which are delivered in response to an outbreak but are not restricted to individuals in proximity to or contact with index cases (64–71). We defined a ring trial as a study that enrolled rings of individuals or households in physical proximity or in contact with an index case and randomly allocated each ring to study groups. We did not consider interventions to be ring interventions if a single contact of an index case was

enrolled or if contacts were enrolled who were possibly exposed to an index case, but the trial investigators made no attempt to identify or confirm index cases.

Search Strategy

We searched PubMed (MEDLINE) and ClinicalTrials.gov in August 2021. We included the search terms "ring trial," "responsive target population," "ring vaccine," "ring intervention," "ring vaccination," "ring treatment," "ring vaccine," "responsive target population," "case area targeted intervention," "permuted locus," "reactive case detection," "reactive focal," "ring prophylaxis," "focal mass drug administration," "targeted mass drug administration," "household contact," and "post exposure prophylaxis" independently and in combination with the terms "trial," "randomized trial," "randomized controlled trial," "randomized control trial," "controlled trial," and "control trial." See additional details in Supplemental Material 1.

Article Selection

Two investigators independently assessed title, abstract, and full text eligibility. Investigators logged inclusion and exclusion criteria during abstract and full text review and resolved discordant classifications between each stage; for discordant classifications during title and abstract review, we erred on the side of including records in the full text review. For trial registrations, two investigators reviewed registration eligibility in a single stage.

Data Extraction

We extracted the following data from each selected publication: country, year, primary and secondary outcomes, intervention(s), comparison group(s), study design, rationale for the study design, ring definition, randomization unit, randomization type (e.g., stratified randomization), index case definition, buffer zones, planned study size, power calculation assumptions, and eligibility criteria. For completed studies, we also extracted results (e.g., study size, compliance, mean response time, parameter estimated, analysis method, outcomes per group, measures of effect).

Risk of Bias Assessment

Investigators independently assessed risk of bias using the revised Cochrane risk-of-bias tool for cluster randomized trials (72). For publications that reported multiple analyses, we focused on the primary analysis. We classified the risk of bias in each domain and overall as "low risk," "some concerns," or "high risk." We resolved discordant classifications through consensus.

1.4 Results and Discussion

Below, we summarize the findings of studies identified in our systematic review. We highlight features of ring trial design and contrast them with alternative designs, drawing on relevant methods and simulation studies.

Trial Selection

We performed title review on all 849 publications, abstract review on 238 publications, and full text review on 73 publications (Figure 1). We reviewed 322 ClinicalTrials.gov registrations and 20 met inclusion criteria. Initial concordance between investigators was 90% following title review, 92% following abstract review, and 93% following full text review; we resolved all discordances through consensus. Concordance for ClinicalTrials.gov registrations was 96%. In total, 52 trials (50 publications and 20 registrations) met inclusion criteria.

Trial Characteristics

Thirty-one trials were completed, 16 were in progress, and three were registered but had not started, and two had withdrawn their registrations (Table 1). Twenty-five trials used a ring design (Figure 2a), seven trials individually randomized contacts of index cases (Figure 2b), and 15 others were CRCTs (Figure 2c). Twenty trials were located in low- or middle-income countries, and 31 studies were located in high-income countries. Studies measured infectious diseases in emergency, outbreak, and emerging infection settings (n = 18), epidemics (n = 15), endemic settings (n = 12), and elimination settings (n = 7).

Interventions

The most common type of interventions were post-exposure prophylaxis or preventive chemotherapy delivered to household members or nearby residents of index cases (Table 1). These included post-exposure prophylaxis for SARS-CoV-2 (n = 12), influenza (n = 9), common cold (n = 1), meningococcal meningitis (n = 1), cholera (n = 1), tuberculosis (n = 1), pertussis (n = 1), and leprosy (n = 1). Studies also applied focal mass drug administration or focal screening and treatment for malaria (n = 7), focal indoor residual spraying for malaria (n = 2), contact or community-based screening and treatment for tuberculosis (n = 2), and household decolonization for *Staphylococcus aureus* (n = 1). Two studies evaluated vaccines for Ebola in contacts of index cases and Hepatitis A in household contacts. A small number of studies evaluated non-pharmaceutical interventions, including handwashing promotion for contacts of cholera or diarrhea cases (n = 2), masks and preventive behavior education for household members of influenza (n = 2) or tuberculosis cases (n = 1), conditional cash transfers for household contacts of tuberculosis cases (n = 1), and notification of partners of chlamydia or HIV cases (n = 2). Several trials compared different types of ring interventions (n = 10); two studies compared ring interventions to population-wide interventions for malaria because the latter are unsustainable, costly, and/or may contribute to drug and insecticide resistance. Most interventions were delivered to all ring members regardless of infection status (n = 49), while some were only delivered to ring members who tested positive for disease (n = 4).

Trial designs

Three types of randomized designs were used to evaluate ring interventions (Figure 2). Twenty-six studies used a ring trial design (Table 1, Figure 2a). Five trials enrolled individuals or households in rings around index cases and then randomly allocated units in each ring to intervention or control, stratifying by ring (i.e., a "ring-stratified trial") (Figure 2b). Fifteen studies were CRCTs of ring interventions, which randomized geographic clusters (e.g., health facility catchment areas) that were defined before index case presentation (Figure 2c). Five trial registrations and one published trial did not include sufficient information to determine trial design. CRCTs were the only design used in elimination settings; ring trials were more common in epidemic and emergency or outbreak settings (Table 2).

In trials where clusters are solely comprised of ring members exposed to index cases, ring trials and CRCTs are equivalent. This was the case in many ring trials that defined rings as household contacts of index cases. On the other hand, several CRCTs defined clusters based on administrative geographic areas, and rings composed a subset of these areas; in this case, ring trials are a subset of a cluster-randomized design. To make this distinction clear, hereafter, we use "CRCT" to refer to traditional cluster-randomized trials that enroll clusters before index case presentation.

In ring trials, the ring was the unit of randomization and the unit of intervention (Figure 2a); in ring-stratified randomized trials, the unit of randomization and intervention was the individual, and randomization was stratified by rings for each index case (Figure 2b); in CRCTs, the units of randomization and intervention were clusters, and a single cluster sometimes contained multiple rings that overlapped in location but not in time (Figure 2c); and in individually randomized trials, the unit of randomization was the individual, and randomization did not consider ring membership.

Index case ascertainment

All but two studies identified index cases through passive surveillance, in which index cases presented at health care facilities, where infection was confirmed with laboratory tests and then reported to surveillance systems (Supplemental Material 2). Passive surveillance effectiveness depends on the extent of healthcare utilization and the robustness of the case reporting system (73–75). Two trials used active surveillance to identify index cases. In one study, health workers tested all individuals in study communities for malaria with rapid diagnostic tests, and treated positive individuals; in the intervention arm, in household members with any positive tests, all individuals were offered treatment regardless of test results (59). A second trial includes an arm in which non-household contacts of leprosy cases who test positive for a serological marker of infection will receive treatment (other arms deliver post-exposure prophylaxis) (76). In principle, active surveillance could also use serologic surveys to detect prior infections, but if prior infections occurred long before serologic assays, interventions may fail to prevent transmission (65).

Ring enrollment

The most common type of ring was household contacts (or nursing home contacts) of index cases, especially in endemic, epidemic, and emergency settings (Table 2, Supplemental Material 2). Studies of Ebola, influenza, COVID-19, chlamydia, and HIV defined rings of contacts and/or contacts of contacts of index cases or household members of an index case. The only trials that defined rings based on geographic proximity (e.g., 100-500m) of index cases used cluster-randomized designs (Table 2). Future ring trials of environmentally-transmitted or vector-borne disease could define rings based on geographic proximity to index cases (51). In studies enrolling contacts of index cases, it may be difficult to identify contacts within the desired response window if the contact tracing system is not robust. Complete contact tracing and enrollment may be more difficult for stigmatized diseases, such as HIV and Ebola (77). In trials defining rings by geographic proximity, it may be difficult to enroll ring members in a timely fashion in the absence of a baseline geographic census identifying the location of all households.

Observation period

Studies measured outcomes within pre-defined observation periods based on the disease incubation period and the expected duration of intervention effectiveness (Supplemental Material 2). For example, most influenza and COVID-19 studies used observation periods of 10-14 days, and malaria interventions used observation periods of 35 days or longer for ring mass drug administration interventions and up to 24 months for reactive indoor residual spraying, which is expected to have a longer effect duration. In some trials, intervention effects were expected to be transient, and participants could be enrolled in ring interventions more than one time. For

example, two trials of reactive focal mass drug administration defined observation periods of 5-8 weeks following the start of an intervention; after this period, if additional index cases occurred in the same area, the intervention was repeated around the new index case (60,61).

A simulation study of ring trials showed the importance of carefully defining observation periods (78). Starting the observation period before the intervention is effective may attenuate effect estimates towards the null. This is because cases occurring soon after index case presentation may result from transmission prior to intervention. Longer follow-up periods will capture initial intervention effects on recipients as well as reductions in secondary transmission, which may be desired. For interventions with short-lived effects, ending the observation period too late could also attenuate effects towards the null because effects on onwards transmission would be expected to be smaller. In vaccine trials, intention-to-treat effects are estimated according to randomized intervention assignment and define the observation period from the time of randomization, which may include the incubation period and time in which vaccinated individuals develop an immune response; per-protocol effects are estimated according to vaccination status, and the observation period starts after the incubation period and development of an immune response (10,79). For example, in its primary analysis, a ring trial of the Ebola vaccine used a per protocol approach that included outcomes 10 days or more following randomization (13,80).

Response time

For infectious diseases with short serial intervals, rapid intervention delivery following index case detection is crucial to ring intervention effectiveness. Trials of influenza and SARS-CoV-2 post-exposure prophylaxis typically had response times close to one day, while response times were longer in other trials (Supplemental Material 3). Longer than planned response times can result in secondary and possibly tertiary transmission before interventions take effect (81,82). For example, two malaria trials reported that longer than planned response times in some clusters might have limited intervention effectiveness and that response time differed between intervention arms (60,61). If response time differs between arms, effect estimates may be biased.

Parameter of interest

CRCTs are amenable to estimation of "total effects," "spillover effects" (i.e., "indirect effects"), and "overall effects," each of which provides different information (9,83,84). Total effects make inferences about effects on intervention recipients, and spillover effects make inferences on untreated individuals in proximity to interventions and may reflect impacts on disease transmission in the study population. Overall effects make inferences about effects on the general population and average across total effects and spillover effects. All completed CRCTs and ring-stratified RCTs estimated "overall effects," comparing all individuals in treatment clusters (including those outside of rings) to all individuals in control clusters (57,60,61,85,86,87, p.). In the ring trial of the Ebola vaccine, the primary analysis estimated "total effects," comparing outcomes among vaccinated individuals in immediate vs. delayed vaccinated groups; a secondary analysis estimated "overall effects" among all eligible individuals in each arm, including unvaccinated individuals (13). No trials estimated "spillover effects" among untreated individuals in treatment vs. control clusters (9).

In CRCTs, it is common to estimate an overall effect, comparing cluster-level outcomes in treatment vs. control clusters. However, when ring members comprise a small proportion of study clusters, the overall effect can differ substantially from the total effect because study

clusters include a large number of untreated individuals. This may be more likely in elimination and emergency settings, where index cases typically occur in spatiotemporal clusters. For example, in two CRCTs in malaria elimination settings that estimated overall effects, the proportion of cluster members that participated in ring interventions ranged from 2% (60) to 27% (61). On the other hand, in endemic settings, index cases may be more evenly distributed within the study population, and ring members may comprise a larger proportion of the study population; in this case, overall effects and total effects may be more similar. Future trials of ring interventions may benefit from estimating each type of effect (total effect, overall effect, and spillover effect if possible) to shed light on intervention impacts in different subpopulations.

Internal validity

We assessed the risk of bias of 33 completed studies; there was a low risk of bias in 22 studies, some concerns about eight studies, and high risk of bias in one study (Supplemental Material 4). Below, we highlight potential risks of bias specific to ring trials and CRCTs of ring interventions, some of which were not identified during the formal risk of bias assessment.

Blinding. The most common threat to internal validity identified in the risk of bias assessment was due to lack of blinding; 11 trials blinded participants to their intervention status (88–98); the remainder were unblinded, typically due to the nature of the interventions. Unblinded studies are often more susceptible to measurement bias, particularly if outcome measurement is subjective, and may have lower retention or compliance (99).

Baseline balance. In all but two ring trials (97,100), all ring-stratified trials, and all but one CRCT (60), baseline characteristics were balanced between study arms. Twenty-two trials used stratified randomization to support baseline balance (Table 1). In CRCTs, if the number of clusters is relatively small, it can be difficult to account for baseline imbalances, even in covariate-adjusted analyses (101). Ring trials and ring-stratified trials are likely to have better baseline balance than CRCTs because randomization occurs after index case detection. This implicitly stratifies study arms by both location and by time, both of which may strongly influence disease incidence. Performing randomization after ring definition was typical in ring trials and ring-stratified trials with some exceptions (97), but in all CRCTs, cluster randomization was performed prior to ring enrollment. In settings with strong spatiotemporal clustering, ring trials and ring-stratified trials can deliver interventions in the same geographic area, which can improve balance, while CRCTs are more vulnerable to imbalances in the number of index cases that occur during follow-up (Figure 2).

Contamination. If there is inadequate social or physical distance between individuals with different treatment assignments, contamination may bias effect estimates towards the null. No ring trials or ring-stratified trials included social or physical buffer zones; however, contamination between rings may be unlikely in household or facility-based ring trials. Ring-stratified randomized trials (Figure 2b) may be particularly vulnerable to contamination because individuals in the same ring may have different treatment assignments. In rings defined as households or nursing homes, contamination is more likely. Three of fifteen CRCTs included geographic buffer zones between clusters to minimize contamination; no studies included buffer zones inside rings or clusters (Supplemental Material 2). One CRCT that did not include buffer zones between clusters assessed possible contamination and did not find evidence of it (60), and one ring trial that enrolled households with index cases reported contamination in which control households adopted intervention behaviors (102). It may be more feasible to include buffers in

ring trials than in CRCTs using fixed geographic areas because ring trials are conducted in a small geographic footprint around index cases, leaving more space for buffers. On the other hand, researchers conducting CRCTs in rare disease settings may need to enroll participants within very large geographic areas to obtain sufficient statistical power, leaving minimal space for buffers. The same principles apply to studies enrolling rings based on contact networks: rings need to be separated by a reasonable number of network nodes to prevent contamination.

Non-compliance. In practice, compliance with random intervention assignment is often imperfect. In trials of ring interventions, non-compliance included: 1) eligible index cases did not trigger interventions (incomplete index case coverage); 2) ring members did not receive their assigned intervention (incomplete target population coverage); and 3) ring members received the incorrect intervention. In CRCTs, the level of non-compliance (cluster vs. individual) affects the magnitude of bias (103). Index case coverage ranged from 58% to 100%, and target population coverage ranged from 27% to 100%; most studies reported that both types of coverage were at least 80% (Supplemental Material 3). In three trials, some participants or clusters received the incorrect intervention, but in two of these cases, the proportion receiving the incorrect intervention was very small (60,104,105). Ring trials require nimble implementation teams to deliver interventions for either arm to any study site location within a short response time. In CRCTs that define clusters in existing administrative areas, it may be easier establish intervention delivery infrastructure within each cluster, increasing compliance. Even so, compliance can remain a challenge in CRCTs: for example, one trial stated that staffing and transportation limitations reduced compliance (60).

When non-compliance depends on participant characteristics or is correlated with loss to follow-up, intention-to-treat estimates that ignore non-compliance are biased (106,107). Two trials investigated this possibility (60,61); one found that disease incidence was inversely associated with target population coverage (61). In any trial, when non-compliance occurs, analysis methods must account for post-treatment measures of compliance (103).

External validity

A common critique of trials is that they have poor external validity (108). Indeed, some included trials cited the need to evaluate ring interventions in multiple sites since benefits of ring interventions may differ between populations (57,58,60). While CRCTs are often considered to have higher external validity than individual RCTs (109), this is not necessarily the case for trials of ring interventions since they are delivered to high-risk individuals. External validity of ring trials may be high when the majority of the study population is susceptible and eligibility criteria are inclusive, as was the case in most of the household post-exposure prophylaxis studies and other ring trials, such as the Ebola vaccine trial (80). On the other hand, CRCTs of malaria ring interventions were predominantly conducted in low transmission elimination settings, where infection occurred in hot spots driven by environmental factors and migration (57,58,60,61). Thus, these trials' findings may generalize only to populations with similar spatiotemporal infection patterns, environmental risk factors, and proportions of immune individuals.

Publication bias

All study registrations for studies that had been completed for at least one year had published a corresponding preprint or manuscript, suggesting that publication bias was not present.

Statistical power

Factors that affect statistical power of CRCTs are well-established (110). Here, we focus on factors that affect power of ring trials and CRCTs of ring interventions. Several included studies described insufficient statistical power (60,86–88,98,111,112). CRCTs commonly estimate the number of clusters required per arm based on the assumed baseline incidence, intra-class correlation (ICC), and true intervention efficacy. In ring trials, additional factors that affect statistical power include the starting day of the follow-up period, probability of case detection, intervention response time, and the force of infection from individuals outside of the ring (78). In a simulation study that used a mathematical model to investigate sample size requirements for immediate vs. delayed Ebola ring vaccination, the factors that had the strongest effect on sample size were the baseline attack rate and the follow-up start day (78).

Probability of case detection. If all detected cases trigger interventions, increasing probabilities of case detection require larger sample sizes because more frequent intervention will cause incidence to decline if the intervention is effective (78). However, increasing case detection probabilities may not require larger sample sizes in trials that do not repeat interventions if subsequent index cases occur during the observation period (60,61).

Baseline incidence. The assumed baseline incidence in sample size calculations were low for elimination settings and studies during the late stage of an outbreak and higher in other settings (Supplemental Material 5). Some studies of Ebola, influenza, and SARS-CoV-2 used the illness rate of contacts rather than the baseline incidence rate in their power calculations (13,100,102,113,114). An advantage of using ring trials over CRCTs when evaluating ring interventions in low incidence settings is that they enroll individuals with the highest incidence in a population, which may translate to greater statistical power. Four trials, including three CRCTs, stated that statistical power was low due to lower than expected incidence during the study period (86,88,98,111). A shared feature of these trials, in contrast to ring trials, is that they enrolled a fixed number of individuals or clusters at baseline instead of at the time of index case presentation. By enrolling rings as index cases occur, ring trials are less susceptible to reductions in statistical power resulting from unexpected decreases in incidence. Simulation studies are needed to investigate whether there is a certain incidence level above which a CRCT is more efficient than a ring trial design.

Compliance. Incomplete intervention coverage, longer than intended response time, and incorrect intervention delivery may compromise statistical power (115). Even in analyses that account for non-compliance (e.g., as treated, per protocol, instrumental variables), higher levels of non-compliance reduce statistical power (103). One study cited unexpectedly low intervention coverage as a potential explanation for limited statistical power (60). A modeling study found that a rapid response time was critical to ring intervention efficacy, especially at higher values of R_0 (82).

Ring size. As in any CRCT, for ring trials, the number of clusters (rings) has a larger impact on statistical power than the number of individuals recruited per ring (6,78). One consideration unique to ring trials is that increasing the ring size (e.g., diameter around the index case or degree of contact network connections enrolled per ring) may reduce the average risk in ring members. If so, increasing the ring size may have little to no benefit to statistical power. To our knowledge, this has not been formally investigated in simulation studies. Further research is needed to evaluate the effect of ring size and membership on statistical power.

Intra-class correlation (ICC). In CRCTs, the extent of clustering can have a large influence on required sample sizes (116). Accurate ICC estimates are often difficult to obtain during trial planning, especially in emerging infection or emergency settings (78,109). For example, in the Ebola ring vaccine trial, the observed ICC of 0.14 was substantially higher than the expected ICC of 0.05 (13). For ring trials, ICCs within the ring around index cases are most relevant, which may be especially difficult to obtain. Observational studies that estimate ICCs in populations adjacent to index cases would support the design of future ring intervention trials (117,118).

Network structure within and between rings. No studies considered transmission network structure in sample size calculations, but one simulation study showed that it can strongly affect statistical power in CRCTs (119). Statistical power reached zero as the proportion of network connections shared between treatment and control clusters approached 50% (119). These findings may apply to ring trials as well, particularly for directly transmitted diseases. This underscores the value of collecting data on spatial and network structure to support sample size calculations. In addition, future studies may benefit from using simulations to inform sample size selection since using ICCs alone may overestimate statistical power when individuals share contacts between rings (120).

Ring trials vs. CRCTs. We note three critical differences between ring trials and CRCTs of ring interventions that we would expect to influence study power. First, the number of interventions and ring members per arm is balanced by design in ring trials but may be imbalanced in CRCTs when there is high spatiotemporal clustering and unpredictable fluctuations in incidence (e.g., emergency and elimination settings). In four CRCTs of ring interventions that enrolled village or health facility clusters, the number of interventions per arm was not balanced because the number of index cases varied between arms (60,61,85,98). Two of these studies noted limited statistical power (60,98). On the other hand, ring trials tended to have balanced numbers of index cases in study arms.

Second, as noted above, ring members may comprise a much smaller proportion of the study population in CRCTs compared to ring trials. This is especially the case when index cases cluster spatiotemporally, as is common in emergency and elimination settings. For example, in a ring trial of the Ebola vaccine, the proportion of ring members was 76%, whereas in three CRCTs conducted in malaria elimination and meningitis outbreak settings, the proportion ranged from 2% to 27% (60,61,85).

Third, by definition, all clusters in ring trials include an index case and are included in analyses; in CRCTs, because clusters are randomized before index case detection, some clusters may have zero index cases during follow-up and must be excluded from analyses. This can reduce power in any setting with a rare outcome. For example, in a CRCT in a malaria elimination setting, only 61% of clusters had at least one index case, limiting statistical power (60). In a CRCT of influenza in nursing homes, the small number of outbreaks resulted in many clusters having no index cases, increasing the length of the study and reducing study power (98). We did not identify any simulation studies that directly compared statistical power of ring trials vs. CRCTs of ring interventions, and this is an important topic for future research.

Ethics

The ethical guidelines for CRCTs largely apply to ring trials (121); below we outline some ethical considerations unique to ring trials.

Informed consent. In community-based CRCTs, consent is often required both at the cluster and individual levels (121,122). In low- and middle-income countries, when the cluster is a community, obtaining group-level consent can be difficult, particularly if there is not an elected community leader to provide consent (123). In ring trials, this is complicated by the fact that ring members around each index case often do not comprise an extant group, such as a school or village. Of the eight completed trials that enrolled village clusters, four obtained both group and individual consent (13,57,61,85). The Ebola ring vaccine trial obtained consent to administer ring vaccination in potential ring sites from local leaders prior to enrolling ring members (80). While it may still be important to obtain the support of local leaders to perform a trial, whether it is ethical to obtain consent from them depends on study circumstances. In addition, in CRCTs, individuals often provide consent to participate after clusters have been randomized for logistical reasons, so it is not possible to obtain consent for randomization (121,123). In the Ebola ring vaccine trial, investigators sought informed consent from ring members after randomization, and notified participants of their treatment assignment after consent was given (13).

Beneficence. In the process of enrolling ring members, ring trials must balance the risk of potentially disclosing index case infection status, which could be harmful for stigmatized diseases, with the potential benefits of the ring intervention. This may be particularly difficult for ring trials that identify ring members through contract tracing. In addition, ring interventions that involve presumptive treatment of individuals without confirmed infection status (e.g., reactive focal mass drug administration) must weigh the potential risk of adverse side effects against benefits, considering that some participants who experience such side effects may be otherwise healthy. Investigators frequently cited minimization of adverse outcomes as a potential benefit of ring interventions in comparison to interventions delivered to an entire population, and several studies monitored adverse effects as a secondary outcome.

Equipoise. The comparison group for a ring intervention must be chosen to ensure equipoise, especially when there is evidence of intervention effectiveness if it is delivered at the individual level in a clinical setting. For example, three malaria trials investigated whether treating all individuals near index cases was more effective than treating individuals near index cases who tested positive using a rapid diagnostic test (the standard of care) (58,60,61). Even though prior trials demonstrated the effectiveness of antimalarials delivered to individuals (124,125) or through mass drug administration (126,127), there was not clear evidence about the effectiveness of the potential ring intervention relative to the standard of care, thus equipoise was present.

Equity. Since ring trials are particularly useful in emergency settings, investigators may consider offering all participants interventions shown to be effective following trial completion to ensure equity. This consideration is particularly important in trials in low- and middle-income countries, where participants may have lower access to care and cutting edge therapies (122). The Ebola ring vaccine trial provided delayed vaccination to the control group to assuage potential concerns of withholding treatment (13,128). Future ring trials could be used in concert with stepped wedge designs to ensure equity in study populations. Another potential advantage of ring trials, particularly for outbreak and emergency settings, is that ring interventions can immediately be implemented following trial discontinuation as was done following the Ebola vaccine ring trial (129).

Extensions

Alternative designs. Ring trials are amenable to additional design modifications, such as adaptive designs (130), as were used in the Ebola ring vaccine trial (80), and stepped wedge designs (131).

Non-communicable diseases. Although we only identified ring intervention trials with infectious disease endpoints, in principle, ring trials could also be appropriate for non-communicable diseases or health behaviors that diffuse through networks (e.g., gun violence (132,133)). Offering interventions to individuals connected to index cases could be particularly useful for outcomes that are stigmatized or underreported (e.g., opioid use disorders (134)). In addition, ring trials could be used for non-communicable vector-borne or environmentally transmitted diseases that tend to cluster spatially or temporally (e.g., Lyme disease, *Coccidioidomycosis*). The design could be particularly useful for studying interventions in populations where climate change results in the introduction or reintroduction of diseases with environmental risk factors.

Limitations

Our search strategy may not have included all possible terms used to describe ring interventions, so our results may not encompass all prior trials of ring interventions. In our narrative review, we only identified a small number of simulation studies investigating ring interventions; only one investigated a ring trial design (78). We consider the paucity of research on this topic an important finding in itself that motivates future research.

1.5 Conclusion

Ring interventions are well-suited to infectious diseases with asymptomatic and heterogeneous transmission. We identified multiple potential advantages of ring trials over ring-stratified trials and CRCTs for evaluating ring interventions. While each type of trial has its limitations, overall, our review identified more potential threats to validity and statistical power in CRCTs of ring interventions and ring-stratified trials than in ring trials, especially in settings with rare and strongly clustered infections. Additional simulation studies are needed to formally compare design features and statistical power of these trial designs. We believe that ring trials hold promise, particularly for evaluations of ring interventions during public health emergencies, seasonal outbreaks, early or waning stages of an epidemic, and disease elimination or eradication settings. To date, novel trial designs have been adopted slowly, particularly in low- and middle-income countries (135). The COVID-19 pandemic has further underscored the urgent need for novel designs, such as the ring trial, that have the potential to maximize investments, reduce cost, and produce rapid, robust results (136).

1.6 Tables and Figures

Table 1. Characteristics of trials identified in the systematic review

First Author, [a] Year (Reference No.)	Registration	Country	Publication Type [b]	Publication Status	Intervention	Control	Study Design	Unit of Randomization	Stratification	Primary Outcome Disease	Study Setting
Barnabas, 2020, 2021 (92,137)	NCT 04328961	United States	Study protocol; article	Completed	Hydroxy-chloroquine as prophylactic	Ascorbic acid	Ring trial	Household	Study site and type of contact (household member vs. healthcare worker)	COVID-19	Emergency / emerging infection
Bath, 2021 (57)	NCT 02556242	South Africa	Article	Completed	Reactive, targeted indoor residual spraying	Standard indoor residual spraying	Cluster RCT	Census ward	Historical malaria and indoor residual spraying coverage, population size and density, and length of waterways	Malaria	Elimination
Bridges, 2017 (58,138)	NCT 02654912	Zambia	Study protocol; trial registration	Recruitment complete	Presumptive antimalarial treatment of population within 140m of index cases	Testing and treatment of positives within 140m of index cases	Cluster RCT	Health facility catchment area	None	Malaria	Elimination
Coldiron, 2017, 2018 (85,139)	NCT 02724046	Niger	Protocol; article	Completed	Ciprofloxacin treatment of index case household or village	Standard of care	Cluster RCT	Village	None	Meningitis	Outbreak
Cowling, 2008 (112)	NCT 00425893	Hong Kong	Article (preliminary results)	Completed	1) Health education plus mask intervention 2) Health education plus handwashing intervention	General health education	Ring trial	Household	None	Influenza	Seasonal epidemic
Echevarria, 1995 (88)	N/A	Peru	Article	Completed	Single-dose ciprofloxacin	Placebo	RCT	Not indicated	None	Cholera	Endemic infection
Egsmose, 1965 (140)	N/A	Kenya	Article	Completed	One year course of isoniazid	Placebo	Ring trial	Household	None	Pulmonary tuberculosis	Endemic infection
Eisele, 2015, 2016, 2020 (59,111,141)	NCT 02329301	Zambia	Study protocol; article; article	Completed	Household-level focal mass drug administration	Community-level mass drug administration	Cluster RCT	Health facility catchment area	Low vs. moderate transmission	Malaria	Elimination setting
Fritz, 2012 (142)	NCT 00731783	United States	Article	Completed	Household infection decolonization	Decolonization of infected individual	Ring trial	Household (intervention) or	None	Staphylococcus aureus infection	Epidemic infection

								individual (control)			
George, 2020; Masud, 2020 (143,144)	NCT 04008134	Bangladesh	Article; article	Completed	Mobile health program focused on handwashing promotion, or mobile health program plus home visits	Standard message on oral rehydration	Ring trial	Household	Study site, hospital ward, and treatment location	Diarrhea	Endemic infection
Halperin, 1999 (89)	N/A	Canada	Article	Completed	Erythromycin estolate for 10 days	Placebo	Ring trial	Household	None	Bordetella pertussis infection	Outbreak
Hayden, 2000 (97)	N/A	United States, Canada, United Kingdom, Finland	Article	Completed	Inhaled zanamivir as prophylactic	Placebo administered through inhaler	Ring trial	Family	None	Influenza	Seasonal epidemic
Hayden, 2004 (105)	N/A	United States, Estonia, United Kingdom	Article	Completed	Oseltamivir as prophylactic	No household treatment except for the index case	Ring trial	Household	Presence of an infant or a second case in the household	Influenza	Seasonal epidemic
Henao-Restrepo, 2015, 2015, 2017 (13,14,80)	PACTR 2015-03001 057193	Guinea	Article; article; study protocol	Completed	Ebola Virus vaccination of contacts and contacts of contacts of index cases	Delayed Ebola Virus vaccination of contacts and contacts of contacts of index cases after 21 days	Ring trial	Contacts and contacts of contacts of index cases	Location (urban vs rural), ring size (<21 vs >20)	Ebola Virus Disease	Emergency / emerging infection
Herzog, 1986 (91)	N/A	Switzerland	Article	Completed	Low-dose intranasal recombinant leucocyte IFN-αA, Ro 22-8181 as prophylactic	Placebo	Ring trial	Family	None	Common cold	Seasonal epidemic
Hsiang, 2020; Medzihradsky, 2018 (61,145,146)	NCT 02610400	Namibia	Article; study protocol; trial registration	Completed	1) Presumptive antimalarial treatment of population within 500m of index cases; 2) Indoor residual spraying within 500m of index cases	1) Testing and treatment of positives within 500m of index cases; 2) Indoor residual spraying within 500m of index cases	Cluster RCT with factorial design	Census enumeration area	Historical malaria incidence, population size and density, and distance from household to healthcare facility	Malaria	Elimination
Ikematsu, 2020 (93)	JapicCTI-184180	Japan	Article	Completed	Baloxavir as prophylactic	Placebo	Ring-stratified RCT	Individuals	Time from illness onset to enrollment; treatment of	Influenza	Seasonal epidemic

									index patient ; participant age		
Iturriaga, 2021 (147)	NCT 04552379	Chile	Study protocol	Recruitment ongoing	Pegylated IFN ß-1a subcutaneous treatment as prophylactic	Standard of care	Ring trial	Household	Number of people in household	COVID-19	Emergency / emerging infection
Kashiwagi, 2013 (94)	JapicCTI-111647	Japan	Article	Completed	Inhaled laninamivir octanoate as prophylactic	Placebo	Ring-stratified RCT	Individuals	Institution; index patient infection with influenza A or B	Influenza	Seasonal epidemic
Kashiwagi, 2016 (95)	JapicCTI-142679	Japan	Article	Completed	Inhaled laninamivir octanoate as prophylactic	Placebo	Ring-stratified RCT	Individuals	Virus type of index case; participants' influenza vaccination status in 2014-2015 influenza season	Influenza	Seasonal epidemic
Low, 2006 (148)	NCT 00112255	England	Article	Completed	Partner notification immediately initiated by practice nurse	Referral to specialist clinic	Ring trial	Sexual partners of index case	Medical practice	Chlamydia	Endemic infection
Mitjá, 2021 (104)	NCT 04304053	Spain	Article	Completed	Hydroxy-chloroquine as prophylactic	Usual care	Ring trial	Ring (e.g., household contacts, healthcare workers, nursing home residents)	None	COVID-19	Emergency / emerging infection
Murphy, 1983 (149)	N/A	United States	Article	Completed	Rifampin as prophylactic	Placebo	Ring trial	Contact unit (members of index household and nonresident contacts)	None	Influenza	Seasonal epidemic
Nakano, 2016 (96)	N/A	Japan	Article	Completed	Inhaled laninamivir octanoate as prophylactic	Placebo	Ring-stratified RCT	Individuals	Virus types for the index case; subjects' influenza vaccination status	Influenza	Seasonal epidemic
Nanni, 2020 (150)	NCT 04363827	Italy	Study protocol	Trial ongoing	1) Hydroxy-chloroquine treatment for one month; 2) hydroxy-chloroquine treatment for 5-7 days as prophylactic	Observation	Ring trial	Household members and/or contacts)	Province COVID-19 incidence; index case is healthcare worker; index case COVID-19 treatment	COVID-19	Emergency / emerging infection
Okebe, 2021 (86)	NCT 02878200	Gambia	Article	Completed	Presumptive dihydro-	Screening of compound	Cluster RCT	Village	Previous leprosy incidence	Malaria	Endemic infection

					artemisinin-piperaquine treatment for all compound members of index case	members of index case					
Ortuno-Gutierrez, 2019; De Jong, 2018 (76,151)	NCT 03662022	Comoros and Madagascar	Protocol; trial registration	Active, recruitment complete	Post-exposure prophylaxis provided to household members of index case, neighborhood contacts within 100m, or contacts within 100m who test positive for a serological marker	No post-exposure prophylaxis	Cluster RCT	Village	None	Leprosy	Hyper-endemic infection
Ram, 2015 (100)	NCT 00880659	Bangladesh	Article	Completed	Intensive handwashing (soap and daily handwashing) behavioral promotion and provision of handwashing station	Standard practices	Ring trial	Household compounds	None	Influenza-like illness	Seasonal epidemic
Sagliocca, 1999 (152)	N/A	Italy	Article	Completed	Hepatitis A vaccine	No vaccine	Ring trial	Household	None	Hepatitis A infection	Endemic infection
Salazar-Austin, 2019 (87)	NCT 03074799	South Africa	Article	Completed	Symptom-based tuberculosis screening of contacts	Skin-test based screening of tuberculosis contacts	Cluster RCT	Clinic	Case notification rate and distance to hospital	Tuberculosis	Endemic infection
Seddon, 2018 (153)	ISRCTN 92634082	South Africa	Study protocol	Ongoing	Daily levofloxacin for 24 weeks	Placebo	Ring trial	Household	Study site	Tuberculosis	Endemic infection
Smit, 2020; Calmy, 2020 (113,154)	NCT 04364022	Switzerland	Study protocol; trial registration	Recruitment complete	Prophylactic Lopinavir/ ritonavir treatment of households with an asymptomatic index case	No treatment of households with an asymptomatic index case (standard of care)	Ring-stratified cluster RCT	Household	Study site	COVID-19	Emergency / emerging infection
Suess, 2012 (102)	NCT 00833885	Germany	Article	Completed	1) Mask/hygiene: households provided with facemask and alcohol based hand-rub and information of proper usage; 2) Mask: households	No masks or hand-rub provided	Ring trial	Household	None	Influenza	Seasonal epidemic

					provided with surgical facemasks and information on correct usage						
Tan, 2021 (114)	NCT 04321174	Canada	Study protocol	Recruitment ongoing	Oral lopinavir/ ritonavir course for 2 weeks as prophylactic	No intervention	Ring trial	Ring (e.g., household members, healthcare workers)	Study site	COVID-19	Emergency / emerging infection
van der Sande, 2014 (98,155)	NCT 01053377; NL92738	Netherlands	Article; trial registration	Completed	Oseltamivir as prophylactic	Placebo	Cluster RCT	Nursing home unit	None	Influenza	Seasonal epidemic
Vasiliu, 2021 (156)	NCT 03832023	Cameroon and Uganda	Protocol	Recruiting	Community-based tuberculosis screening of household contacts	Facility-based standard of care	Cluster RCT	Health facility catchment area	Country	Tuberculosis	Endemic infection
Vilakati, 2021; Hsiang, 2014 (60,157)	NCT 02315690	Eswatini	Article; trial registration	Completed	Presumptive antimalarial treatment of population within 200m of index cases	Testing and treatment of positives within 500m of index cases	Cluster RCT	Locality	Malaria history; cluster size	Malaria	Elimination
Wamuti, 2015; Cherutich, 2017 (158,159)	NCT 01616420	Kenya	Protocol	Completed	Assisted partner notification services immediately after index case enrollment	6-week delayed assisted partner notification services	Cluster RCT	HIV testing site	Country and rurality	HIV	Epidemic infection
Wang, 2021 (160)	NCT 04536298	United States	Study protocol	Recruitment ongoing	High-dose vitamin D3 supplementation as 1) early treatment, and 2) prophylactic	Placebo capsule of identical appearance and taste	Ring trial	Dyads (index case plus closest household member)	None	COVID-19	Emergency / emerging infection
Welliver, 2001 (90)	N/A	Belgium, Canada, Denmark, Finland, Germany, Netherlands, Norway, Switzerland, United Kingdom, United States	Article	Completed	Oseltamivir as prophylactic	Placebo	Ring trial	Household	None	Influenza	Seasonal epidemic
Wingfield, 2017 (161)	N/A	Peru	Article	Completed	Standard of care plus socioeconomic support	Standard of care	Ring trial	Household	None	Tuberculosis	Endemic infection
Agrawal, 2020 (162)	NCT 04342156	Singapore	Trial registration	Withdrawn	Hydroxy-chloroquine sulfate	No treatment	Ring trial	Household	None	COVID-19	Emergency / emerging infection

Bardin, 2020 (163)	NCT 04343248	United States	Trial registration	Trial ongoing	Nitazoxanide as prophylactic, with Vitamin Super B-Complex as dietary supplement	Placebo, with Vitamin Super B-Complex as dietary supplement	RCT [c]	Not specified	None	COVID-19 and other viral respiratory illnesses	Emergency / emerging infection
Bennett, 2020 (164)	NCT 04416945	Lao People's Democratic Republic	Trial registration	Not yet recruiting	Testing and treatment of positives in five nearest households to index case	Standard of care and village-based RACD	Cluster RCT	Health facility catchment area	None	Malaria	Elimination setting
Borrie, 2020 (165)	NCT 04397328	Canada	Trial registration	Not yet recruiting	Hydroxy-chloroquine as prophylactic	Placebo	RCT [c]	Not specified	None	COVID-19	Emergency / emerging infection
Bracchi, 2021 (166)	NCT 04842331	United Kingdom	Trial registration	Recruitment ongoing	RESP301 (Nitric Oxide generating solution) as prophylactic, with standard of care	Standard of care	Ring trial	Household	None	COVID-19	Emergency / emerging infection
Elvira, 2021 (167)	NCT 04938596	Chile	Trial registration	Not yet recruiting	Combination of mask provision, prevention recommendations, and education about tuberculosis	Standard of care	Cluster RCT	Health care area and correspond-ing clinics	None	Tuberculosis	Endemic infection
Gadisa, 2020 (168)	NCT 04241705	Ethiopia	Trial registration	Recruitment ongoing	1) Presumptive antimalarial treatment of population within 100m of index cases; 2) Testing and treatment of positives within 100m of index cases	Standard of care	Cluster RCT	District	None	Malaria	Elimination
Giles, 2021 (169)	NCT 04318444	United States	Trial registration	Recruitment ongoing	Hydroxy-chloroquine as prophylactic	Placebo	RCT [c]	Not specified	None	COVID-19	Emergency / emerging infection
Malin, 2021 (170)	NCT 04894474	Not specified	Trial registration	Withdrawn	Antibody BI 767551 medication	Placebo	RCT [c]	Individual	None	COVID-19	Emergency / emerging infection
McGeer, 2020 (171)	NCT 04448119	Canada	Trial registration	Active, recruitment complete	Favipiravir	Placebo	Ring trial	Long term care home	None	COVID-19	Emergency / emerging infection
Sued, 2021 (172)	NCT 04788407	Argentina	Trial registration	Recruitment ongoing	Nitazoxanide as prophylactic	Placebo	RCT [c]	Not specified	None	COVID-19	Emergency / emerging infection

RCT: Randomized controlled trial. COVID-19: Coronavirus disease 2019. RACD: Reactive case detection.

[a] First author last name for published articles, pre-prints, or protocols. Principal Investigator last name for trial registrations with no publication.

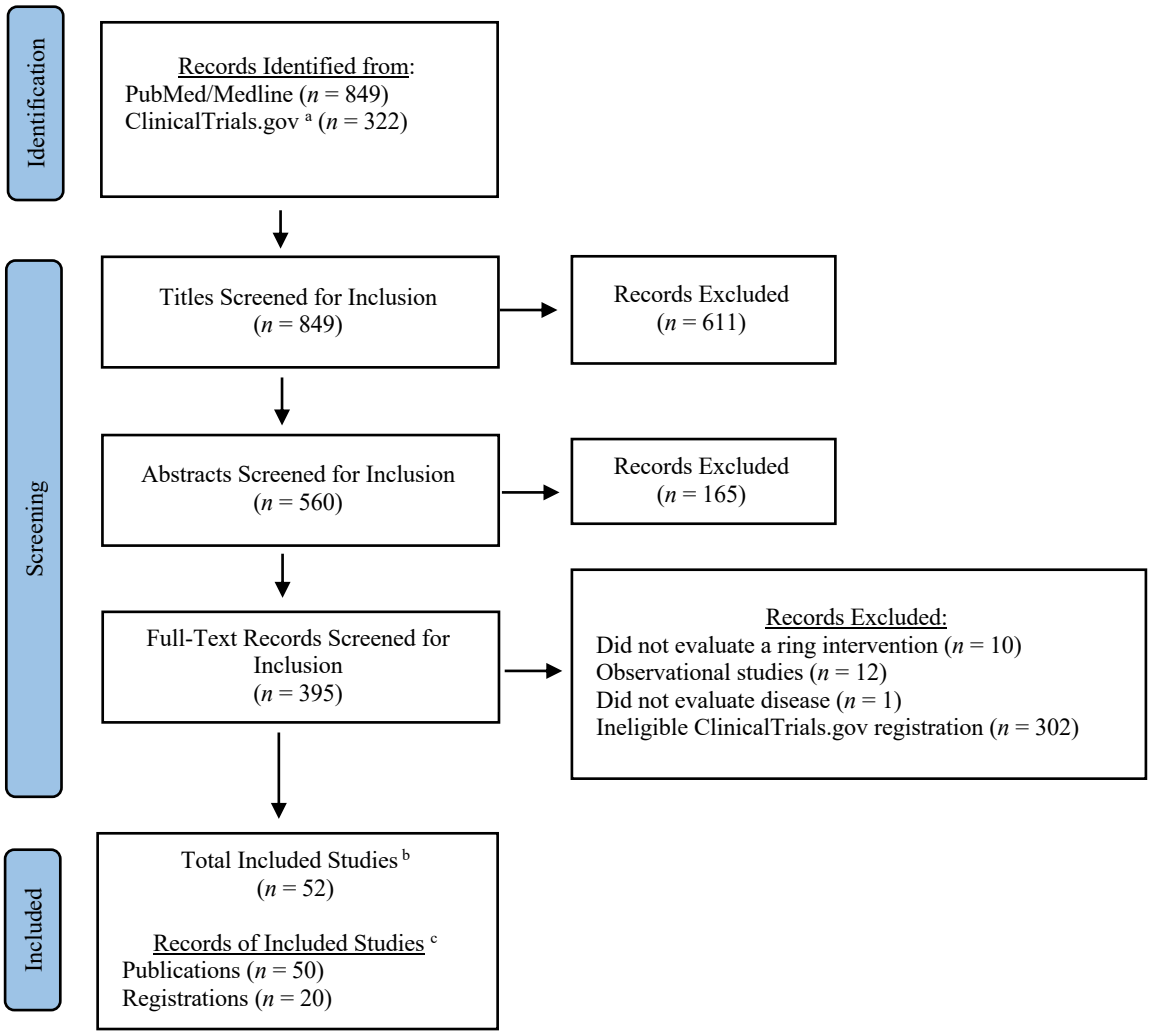
[b] Includes all types of articles retrieved in the systematic review.
[C] Insufficient information to determine type of trial.

Table 2. Number of included studies by study design, ring type, and study setting

Study characteristic	Study setting							
	Endemic setting (n = 12[a])		Epidemic setting (n = 15)		Emerging infection, emergency, and outbreak (n = 18[b])		Elimination setting (n = 7)	
	No.	%	No.	%	No.	%	No.	%
Trial design								
Cluster-randomized trial	5	45	2	13	1	8	7	100
Ring-stratified cluster-randomized trial	0	0	4	27	1	8	0	0
Ring trial	6	55	9	60	11	84	0	0
Ring type[c]								
Household including index case	8	73	14	93	10	67	1	14
Neighborhood around index case	2	18	0	0	1	7	7	100
Contacts of index case	1	9	2	13	3	20	0	0

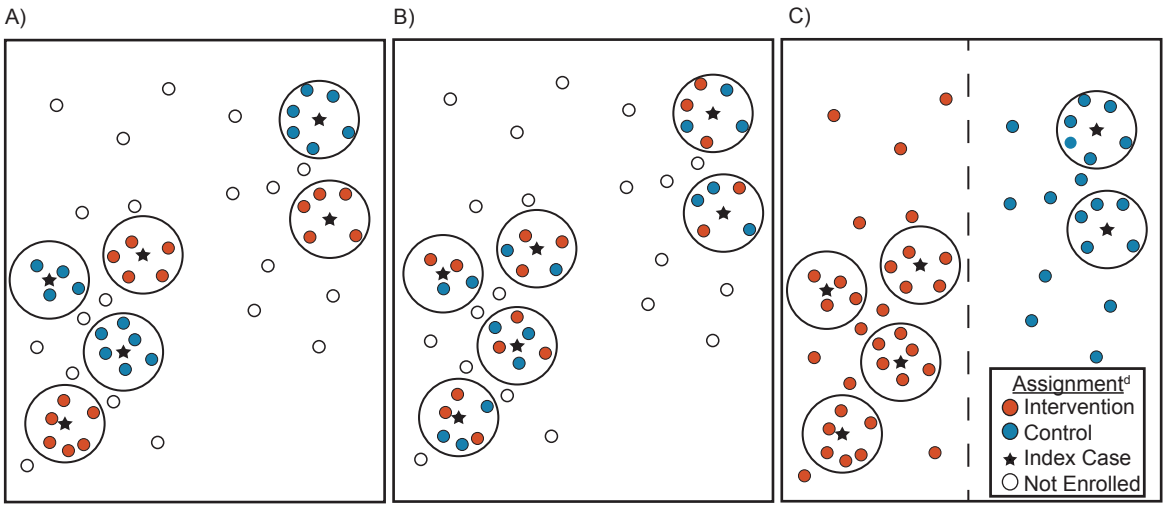
[a] Only 11 studies provided sufficient information to determine trial design and ring type.
[b] Only 13 studies provided sufficient information to determine trial design, and 15 provided sufficient information to determine ring type.
[c] If multiple types of rings were used, column percentages exceed 100

Figure 1. PRISMA flow diagram for systematic review screening and inclusion.



[a] All registrations from ClinicalTrials.gov were reviewed in a single stage of full-text review.
[b] Included studies refers to research projects for which one or more records were included.
[c] Records of studies include trial registrations, published trial protocols, and original research articles.

Figure 2. Types of ring intervention trial designs.



A) Ring trial design
B) Ring stratified randomized trial
C) Cluster-randomized trial of ring intervention. The dotted line separates Cluster 1 (left) from Cluster 2 (right). While all participants in Cluster 1 were assigned to the intervention group, only participants inside the four rings received the intervention.
[d] Red and blue dots indicate assignment to intervention or control, respectively. A star indicates a disease event that triggered a ring intervention, and a dot with a white center indicates an individual who was not enrolled or randomized.

1.7 Supplemental Material

Supplemental Material 1. Search strategy

In PubMed, we entered the following query to search title, abstract, and full text fields: ((“ring trial”) OR (“ring vaccination trial”) OR (“permuted locus trial”) OR (“ring vaccination”) OR (“ring treatment”) OR (“responsive target population”) OR (“case area targeted”) OR (“permuted locus”) OR (“reactive case detection”) OR (“reactive focal”) OR (“ring prophylaxis”) OR (“focal mass drug administration”) OR (“focal MDA”) OR (“targeted mass drug administration”) OR (“targeted MDA”) OR (“index case*” AND (contact* OR neighb* OR compound*)) OR (“patient households”) OR (“household contact*” AND “prophyla*”) OR (“post exposure prophyla*”) OR (“post exposition prophyla*”)) AND (“trial*” OR “randomize*” OR “randomis*”). We restricted search results to studies that were published in English.

In ClinicalTrials.gov, we used the Advanced Search feature and inputted the following search terms in the “Other terms” option while limiting “Study Type” to “Interventional Studies (Clinical Trials)”: “responsive target population”, “ring vaccine”, “ring intervention”, “ring vaccination”, “ring treatment”, “ring vaccine”, “responsive target population”, “case area targeted intervention”, “permuted locus”, “reactive case detection”, “reactive focal”, “ring prophylaxis”, “focal mass drug administration”, “index case” AND contacts, “index case AND neighbors”, “index case” AND compounds, “patient households”, “household contacts” AND “prophylaxis”, “household contacts” AND “prophylactic”, “post exposure prophylaxis”, “post exposure prophylactic”, “post exposition prophylaxis”, “post exposition prophylactic” independently and in combination with the terms “trial”, “randomized trial”, “randomized controlled trial”, “randomized control trial”, “randomised trial”, “randomised controlled trial”, “randomised control trial”, “controlled trial”, and “control trial”. We omitted search terms and that yielded no results in order to avoid the search tool’s expansion of inclusion criteria to include similar but terms or studies that matched a single component of the search term (eg. all “trials”).

Supplemental Material 2. Characteristics of rings in trials evaluating ring interventions.

Author [a], Year (Reference No.)	Ring Definition	Index Case Identification	Buffer Zone Between Clusters	Observation Period
Barnabas, 2020, 2021 (92,137)	Individuals in the household of an index case and medical staff who cared for index case without PPE	Passive surveillance	None	14 days
Bath, 2021 (57)	Households within 200m of index case	Passive surveillance	Natural boundaries or uninhabited space	Entire follow-up period (24 months)
Bridges, 2017 (58,138)	Radius of 140m from index case	Passive surveillance	3 km buffer between each health facility	2 years
Coldiron, 2017, 2018 (85,139)	Household members of index case or all members of index case's community	Passive surveillance	None	28 days
Cowling, 2008 (112)	Household members of an index case	Passive surveillance	None	9 days
Echevarria, 1995 (88)	Household members of index case	Passive surveillance	None	14 days
Egmose, 1965 (140)	Individuals regularly sleeping in the household of the index case	Passive surveillance	None	14 days
Eisele, 2015, 2016, 2020 (59,111,141)	All individuals in the household of an index case	Active surveillance [c]	1.5 km between health facility catchment areas	1 year
Fritz, 2012 (142)	Household members of index case	Passive surveillance	None	12 months
George, 2020; Masud, 2020 (143,144)	Household members of index case	Passive surveillance	None	12 months
Halperin, 1999 (89)	Household members of index case	Passive surveillance	None	28 days
Henao-Restrepo, 2015, 2017 (13,14,80)	Index cases' contacts and contacts of contacts	Passive surveillance [b]	None	31 days post-randomization with a delay of 10 days
Herzog, 1986 (91)	Families with 3-6 persons and a minimal age of 3 years, in contact with index case	Passive surveillance	None	10 days
Hsiang, 2020 (61,145,146)	Populations residing within 500m of index case	Passive surveillance	None	8 weeks from first intervention in each cluster
Ikematsu, 2020 (93)	Household contacts of an index case who had lived in the same household for at least 48 hours before enrollment	Passive surveillance	None	10 days
Iturriaga, 2021 (147)	Eligible households: index case with ≥ 1 household contacts	Passive surveillance	None	11 days
Kashiwagi, 2013 (94)	Household members who had been in contact with an index patient within 48 hours of symptom onset	Passive surveillance	None	10 days
Kashiwagi, 2016 (95)	Index case's cohabiting family members within 48 hours of symptom onset	Passive surveillance	None	11 days
Low, 2006 (148)	Index case's sexual contacts in the previous six months	Active and passive surveillance	None	6 weeks
Mitjá, 2021 (104)	Contacts who were epidemiologically linked to an index case: adults with recent history of close-contact exposure (> 15 minutes within 2 meters, up to 7 days before enrollment) with an index case, and adults with an increased risk of	Passive surveillance	None	14 days

	infection (e.g., healthcare worker, household contact, nursing home worker, nursing home resident)			
Murphy, 1983 (149)	Contact units of an index case: members of the index household, and nonresident contacts who spent ≥ 8 hours with index patient during the 7 days before hospitalization	Passive surveillance	None	2 weeks [e]
Nakano, 2016 (96)	Pediatric (< 10 years) household members of index case, in contact with the index case	Passive surveillance	None	11 days
Nanni, 2020 (150)	Household members and/or contacts of index case and healthcare professionals (Group 1), or patients with COVID asymptomatic or paucisymptomatic in home situations without anti-COVID meds (Group 2)	Passive surveillance	None	Group one: 1 month; Group two: 14 days
Okebe, 2021 (86)	Compound members of index case	Passive surveillance	None	4 days
Ortuno-Gutierrez, 2019 (76,151)	Household members of index case, neighborhood contacts within 100m of index case, or neighborhood contacts within 100m who test positive for a serological marker	Active surveillance	None	1 year
Ram, 2015 (100)	Members of an index case's household compound (single household or several households occupied by joint/extended family)	Passive surveillance	None	10 days
Sagliocca, 1999 (152)	Index case's family contacts older than 1 year and younger than 40 years	Passive surveillance	None	45 days
Salazar-Austin, 2019 (87)	Children under 5 years in same household as index case	Passive surveillance	None	6 months
Seddon, 2018 (153)	Children under 5 years in same household as index case	Passive surveillance	None	18 months
Smit, 2020 (113,154)	Asymptomatic contacts of individuals diagnosed with COVID-19 enrolled through contact tracing and social media	Passive surveillance	None	21 days
Suess, 2012 (102)	Household members of an index case	Passive surveillance	None	8 days
Tan, 2021 (114)	Index case's high risk close contacts within the past 1-7 days (e.g., direct caregiver, close physical contact, direct contact, indoor contact, cohabiting)	Passive surveillance	None	14 days
van der Sande, 2014 (98,155)	Residents of an index case's nursing home unit	Passive surveillance	None	10 days
Vasiliu, 2021 (156)	All children under 5 years and all symptomatic individuals in same household as index case or all household members of index case	Passive surveillance	None	6 months
Vilakati, 2021 (60,157)	Individuals living within 200m (intervention) or 500m (control) of case	Passive surveillance	None	35 days from day of first index case in each cluster
Wamuti, 2015 (158,159)	Index case's sexual partners in the previous 3 years	Passive surveillance	None	3 months
Wang, 2021 (160)	Closest cohabiting/household contact of an index case	Passive surveillance	None	4 weeks
Welliver, 2001 (90)	Households containing 2-8 contacts with index case	Passive surveillance	None	7 days
Wingfield, 2017 (161)	Household members of index case	Passive surveillance	None	24 weeks

Bennett & Vanisaveth, 2020 (164)	Five neighboring households and those who worked in the same forested area of index case	Passive surveillance	None	4 months
Gadisa, 2020 (168)	Residing within 100m of index case	Passive surveillance	None	2 years

[a] First author last name for published articles, pre-prints, or protocols. Principal Investigator last name(s) for trial registrations with no publication.

[b] Passive surveillance: Index cases present at health care facilities, where they are confirmed with laboratory tests and then reported to surveillance systems.

[c] Active surveillance: Health workers test entire communities or samples of communities using rapid diagnostic tests and regardless of symptom status.

[d] Hayden et al., 2000 measured outcomes at earlier time points within this period, including an in-person visit on days 11 and 28 and telephone screening on days 5 and 14.

[e] Murphy et al., 1983 measured the primary outcome as culture positives within 7 days and after 2 weeks of completing the intervention. However, additional measurements included telephone contact with households 6 months or longer after intervention.

Supplemental Material 3. Measures of compliance in published trials.

Author [a], Year	Median response time per arm [b]	Index case coverage [c]	Target population coverage [d]	Percentage receiving correct intervention
Barnabas, 2021	2 days (IQR: 1 – 3 days) [e]	Not applicable [f]	82% to 94% [g]	100%
Bath, 2021	Not reported	Not reported	Not reported	100%
Coldiron, 2017, 2018	2.4 days for village prophylaxis 24 hours for household prophylaxis	100%	77% for village prophylaxis 4% for household prophylaxis	100%
Cowling, 2008	1 day [h]	58% in control arm, 63% in face mask arm, 89% in hand hygiene arm	Not reported	100%
Echevarria, 1995	Within 24 hours	Not reported	52% of total	Not reported
Egmose, 1965	Within 2 weeks	Not reported	99% intervention 97% control	Not reported
Eisele 2015, 2016, 2020	Not reported	Not reported	71% mass drug administration 71% focal mass drug administration	100%
Fritz, 2012	Median 21 days	80% intervention 82% control	74% intervention group 90% control group	100%
George, 2020; Masud, 2020	Not reported	Not reported	Not reported	100%
Halperin, 1999	11.7 days intervention 12.5 days control	Not reported	85% intervention 87% control	100%
Hayden, 2000	Not reported	99% in intervention arm, 99% in placebo arm [i]	98% in intervention arm, 98% in placebo arm [i]	100%
Hayden, 2004	23 hours in the prophylactic arm, 23.2 hours in the expectant arm	Not reported	Not reported	96%
Henao-Restrepo, 2015, 2017	10 to 11 days	84%	66%	100%
Herzog, 1986	Not reported	99%	Not reported	Not reported
Hsiang, 2020	13 to 14 days	82% to 91%	86% to 93%	100%
Ikematsu, 2020	1 day	100%	>99%	100%
Kashiwagi, 2013	21.6 and 23 hours in the intervention arms, 22.5 hours in the placebo arm [j]	Not reported	Not reported [k]	100%
Kashiwagi, 2016	20.6 hours and 22.6 hours in the intervention arms, 21.9 hours in the placebo arm [j]	100% in the two intervention arms	100% in the two intervention arms	100%
Low, 2006	13.2 days intervention	100% intervention 69% control	45% across both arms	100%
Mitjá, 2020	4 days (IQR: 3 – 6 days)	Not applicable [f]	95% in intervention arm, 98% in the usual-care arm	99% in control arm, 100% in intervention arm
Murphy, 1983	Not reported	97%	Not reported	Not reported
Nakano, 2016	Not reported	100%	99% in intervention arm, 99% in placebo arm	100%
Okebe, 2021	Not reported	100% intervention	95% intervention	100%
Ram, 2015	Not reported	99%	Not reported	100%
Sagliocca, 1999	Within 4 days [h]	100%	87.8% in intervention arm, 86% in control arm	100%

Salazar-Austin, 2019	Not reported	100%	30% intervention 27% control	100%
Suess, 2012	Not reported	86% in control group, 70% in mask group, 72% in mask + hygiene group	Not reported	100%
van der Sande, 2014	Not reported	100% [l]	80% in intervention arm, 85% in placebo arm	100%
Vilakati, 2021	7 to 11 days	77% to 80%	76% to 81%	86% of intervention clusters, 95% of control clusters that received interventions
Welliver, 2001	Not reported	Not applicable [f]	99% in intervention arm, 99% in placebo arm	100%

[a] First author last name for published articles, pre-prints, or protocols. Principal Investigator last name(s) for trial registrations with no publication.

[b] Response time: the median time elapsed between index case ascertainment and intervention delivery.

[c] Index case coverage: the proportion of ascertained and eligible index cases that triggered a ring intervention.

[d] Target population coverage: the proportion of eligible ring members that received the intervention.

[e] Time between most recent exposure and first dose of study medication

[f] Index case did not receive any intervention by design

[g] Participants' self-reported adherence to the intervention

[h] Time between symptom onset in the index subject to application of the intervention

[i] Percentage receiving and completing the intervention, rather than the percentage solely receiving the intervention.

[j] Time between symptom onset in the index subject to first dose, reported as the mean rather than median

[k] Kashiwagi et al., 2013: percentages of total participants are reported, but not parsed by index cases and target population. 99% and 100% coverage in the two intervention arms, and 99% coverage in the placebo arm, for all participants.

[l] Based on the number of outbreaks that triggered interventions, which may include multiple index cases that were detected simultaneously

Supplemental Material 4. Risk of bias in included trials.

Author, Year	Study Design	Randomization Process	Recruitment of participants	Deviations from intended intervention	Missing outcome data	Outcome measurement	Reported result selection	Overall risk of bias	Predicted direction of bias
Barnabas, 2021	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Bath, 2021	Cluster RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Coldiron, 2018	Cluster RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Cowling, 2008	Ring trial	Some concerns	Low risk	Low risk	Some concerns	Low risk	Low risk	Some concerns	Unpredictable
Echevarria, 1995	RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Egsmose, 1965	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Eisele, 2020	Cluster RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Fritz, 2012	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
George, 2020	Ring trial	Low risk	Low risk	Low risk	Low risk	Some concerns	Low risk	Some concerns	Away from null
Halperin, 1999	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Hayden, 2000	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Hayden, 2004	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Henao-Restrepo, 2015	Ring Trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Henao-Restrepo, 2017	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Herzog, 1986	Ring trial	Some concerns	Low risk	Low risk	High risk	Low risk	Some concerns	High risk	Unpredictable
Hsiang, 2020	Cluster RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Ikematsu, 2020	Ring-stratified trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Kashiwagi, 2016	Ring-stratified RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Low, 2006	Ring trial	Low risk	Low risk	Low risk	Some concerns	Low risk	Low risk	Some concerns	Unpredictable
Mitja, 2020	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Murphy, 1983	Ring trial	Low risk	Low risk	Low risk	Low risk	Some concerns	Low risk	Some concerns	Away from null
Nakano, 2016	Ring-stratified RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Okebe, 2021	Cluster RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Ram, 2015	Ring trial	Some concerns	Low risk	Low risk	Low risk	Low risk	Low risk	Some concerns	Unpredictable

Sagliocca, 1999	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Salazar-Austin, 2019	Cluster RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Vilakati, 2021	Cluster RCT	Low risk	Low risk	Some concerns	Low risk	Low risk	Low risk	Some concerns	Towards null
Wingfield, 2017	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA
Suess, 2012	Ring trial	Low risk	Low risk	Some concerns	Low risk	Low risk	Some concerns	Some concerns	Unpredictable
van der Sande, 2014	Ring trial	Some concerns	Low risk	Low risk	Low risk	Low risk	Low risk	Some concerns	Unpredictable
Welliver, 2001	Ring trial	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	NA

RCT: randomized controlled trial.

Supplemental Material 5. Sample size assumptions, planned sample size, and enrolled sample size of published trials.

Author [a], Year	Baseline Risk	Minimum Detectable Effect[b]	Expected Correlation	Sample Size Required	Sample Size Included
Barnabas, 2020, 2021	6% attack rate of SARS-CoV-2; 10% attack rate	50% reduction in incident SARS-CoV-2	Not reported	2000 participants; 1240 participants (recalculated)	689 participants
Bath, 2021	2.2 cases of malaria per 1000 person-years	Noninferiority margin of 1 case per 1000 person-years	Coefficient of variation between clusters of 0.5	62 clusters	62 clusters 393,387 participants in study clusters
Bridges, 2017	5% malaria seropositivity	64.5% relative difference	Intra-class correlation coefficient 0.5	16 clusters	16 clusters 2,820 households in study clusters
Coldiron, 2017, 2018	Control attack rate between 0.005 and 0.05	Attack rate reduction between 50% and 90%	Intra-class correlation 0.025	Not established a priori	49 village clusters
Cowling, 2008	Not reported	24% secondary attack ratio to within +/- 7%	Not reported	51 households with an average of 3.8 members	128 households, 198 subjects
Echevarria, 1995	15% incidence	13.5% incidence difference	Not reported	200 household contacts	213 household contacts
Egsmose, 1965	2% prevalence	Not reported	Not reported	Not reported	775 household contacts
Fritz, 2012	Not reported	50% relative reduction	Not reported	183 households	183 households
George, 2020; Masud, 2020	2-week diarrhea prevalence of 8%	25% prevalence difference	Within household correlation 0.1	750 households	769 households
Halperin, 1999*	Not reported	Not reported	Not reported	Not reported	152 households
Hayden, 2000	Not reported	70% rate of efficacy of prophylactic	Not reported	270 families	321 families, 1158 participants
Hayden, 2004	50% of index cases with confirmed influenza infection; ≥ 20% influenza incidence in households	80% effectiveness of prophylactic in preventing further spread to household contacts	Not reported	200 households	277 households, 298 index cases, 812 household contacts
Henao-Restrepo, 2015, 2017	Contacts' Ebola virus disease illness risk 1% to 5%	Vaccine efficacy between 50% and 90%	Intra-class correlation coefficient 0.05	96 rings[c]	98 clusters 9,096 participants in rings
Herzog, 1986	Not reported	Not reported	Not reported	Not reported	191 families, 587 participants (189 index cases, 337 contacts)
Hsiang, 2020	24.4 cases of malaria per 1000 individuals	50% or more relative reduction in incidence for clusters receiving a single reactive intervention, and 75% or more relative reduction in incidence for clusters receiving combined interventions	Coefficient of variation 0.95	56 clusters	55 clusters 8,948 participants in rings
Ikematsu, 2020	10% clinical influenza in the placebo group	Risk ratio of 0.4	Not reported	748 participants	545 index patients, 752 household participants
Iturriaga, 2021	85% of untreated participants shedding virus at end of day 11	Odds ratio of 0.5 for a reduction in transmission to household contact	Intra-class correlation coefficient 0.15	310 households x 4 members = 1240 participants	251 households, as of protocol publication date
Kashiwagi, 2013	1.65% clinical influenza in treatment	70% protective efficacy	Not reported	470 participants in each group	FAS group [d]: {550, 550, 558}, FASII

	groups, 5.5% in the placebo group				group: {543, 539, 546}, FASIINAB group: {487, 486, 478}
Kashiwagi, 2016	3% clinical influenza in treatment groups, 10% in the placebo group	70% protective efficacy	Not reported	250 participants in each group	FAS group: {267, 269, 265}, FASII group: {267, 262, 261}, FASIINAB group: {248, 243, 241}
Low, 2006	40% of partners treated in control	20% difference in partners treated	Not reported	214 participants	140 participants
Mitjá, 2020	5% expected incidence in treatment group, 15% expected incidence in control group	10 percentage point between-group difference in incidence; Re-estimation: 3.5 percentage point difference	Intra-class correlation coefficient 1.0	190 clusters with 15 contacts per cluster; Re-estimation: 640 clusters with 3.5 contacts per cluster	672 clusters; 2314 contacts
Murphy, 1983	Not reported	Not reported	Not reported	Not reported	312 subjects
Nakano, 2016	15% clinical influenza in the placebo group, 4.5% in the treatment group	70% relative risk reduction	Not reported	300 subjects	341 subjects
Nanni, 2020	Not reported	Not reported	Not reported	Group one (hydroxychloroquine for a month): 1000-1300 index case clusters, 2000 contacts as participants; Group two (hydroxychloroquine for 5-7 days): 1000-1300 index case clusters * 25-30% = 300 participants	Not reported (protocol only)
Okebe, 2021	2.8% prevalence	60% prevalence difference	Coefficient of variation 0.7	34 villages	50 villages
Ortuno-Gutierrez, 2019[c]	Incidence 1.5 per 1000	50% incidence reduction	Coefficient of variation 0.29	124,0000	Not reported
Ram, 2015	Estimates based on 30%, 20%, and 10% secondary attack rate in the control group	50% relative risk reduction	Intra-cluster correlation: 0.37	200 household compounds	377 household compounds
Sagliocca, 1999	10% incidence of secondary hepatitis A virus infection in the untreated group	80% protective efficacy of the vaccine	Not reported	160 households per group	146 index case households, 351 household contacts
Salazar-Austin, 2019	80% uptake	40% increase in uptake	Coefficient of variation 0.3	1,152 participants per arm	550 and 467 participants in each arm, respectively
Seddon, 2018	7% incidence	50% incidence reduction	Intra-class correlation 0.1	1,556 participants	Not reported
Smit, 2020	20% SARS-CoV-2 incidence among close contacts	60% relative risk reduction	Intra-class correlation coefficient .05; design effect 1.1	300 participants	100 clusters 300 participants in rings
Suess, 2012	20% secondary attack rate in household contacts of control group	75% difference in secondary attack rates (20% in control group, 5% in intervention group)	Intra-cluster correlation coefficient: 0.3	114 household members	84 households, 302 participants (218 household contacts)

Tan, 2021	Secondary attack rate of 15%	40% decrease in relative risk of COVID-19	Intra-class correlation coefficient 0.05	244 rings x 5 contacts = 1220 participants	Not reported (protocol only)
van der Sande, 2014	40% of nursing home units experience influenza outbreaks in control	70% reduction further transmission in units with new symptomatic influenza-confirmed cases	Not reported	60 nursing home units	15 nursing home units
Vasiliu, 2021	Between 60% and 70% therapy completion	10% difference in therapy completion	Intra-cluster correlation 0.01	1,500 participants	Not reported
Vilakati, 2021	4 malaria cases per 1000 individuals	50% reduction in incidence if at least 63 out of 77 clusters have at least one index case	Coefficient of variation 0.9	63 clusters	77 clusters 3,628 participants in rings; 47 clusters with at least one index case
Wamuti, 2015	1 to 2 partners per site will seek testing in control arm	Twofold increase in partners seeking testing	Coefficient of variation 0.25	1,080 index cases	Not reported
Welliver, 2001	Not reported	Not reported	Not reported	Not reported	371 households, 955 contacts
Wingfield, 2017	Not reported	50% increase in treatment initiation	Not reported	400 participants	410 participants

[a] First author last name for published articles, pre-prints, or protocols. Principal Investigator last name(s) for trial registrations with no publication.

[b] The Ebola ring vaccine trial (Henao-Restrepo 2017; Henao-Restrpo 2015) used a power of 90%.

[c] Ortuna-Gutierrez, 2019 used a significance level of 0.019 in order to account for multiple comparisons.

[c] The Ebola ring vaccine trial reported a range of possible required sample sizes given the range of baseline risk and minimum detectable effect, and this value reflects the required sample size with 2% risk of disease and 90% vaccine efficacy.

SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2.

[d] FAS is the full analysis set based on the intention-to-treat principle, FASII is the FAS with index-infected participants, and FASIINAB is the FAS with index-infected and virus-negative at baseline participants.

2. Chapter 2. Treatment Heterogeneity of Water, Sanitation, Hygiene, and Nutrition Interventions on Child Growth by Environmental Enteric Dysfunction and Pathogen Status for Young Children in Bangladesh

2.1 Abstract

Background: Despite widespread use of water, sanitation, hygiene (WSH), nutrition (N), and combined (N+WSH) interventions by global health and international development organizations, investigators have found mixed evidence regarding these interventions' impact on child growth. WSH interventions may insufficiently reduce pathogen exposure, and nutrition interventions may be modified by environmental enteric dysfunction (EED), a condition of increased intestinal permeability and inflammation. This study sought to investigate the relationship between these factors and intervention effectiveness by evaluating N, WSH, and N+WSH treatment heterogeneity based on individual pathogen and EED biomarker status with respect to child linear growth.

Methods: We used targeted learning methods, which allow for doubly-robust estimation with minimal parametric assumptions, to estimate drivers of treatment heterogeneity. We applied cross-validated targeted maximum likelihood estimation and super learner ensemble machine learning to assess the conditional treatment effects based on biomarker and pathogen status. We analyzed treatment (N+WSH, WSH, N, or control) randomly assigned in-utero, child pathogen and EED data at 14 months of age, and child LAZ at 28 months of age. We controlled for baseline covariates related to household socioeconomic status, living conditions, and maternal health. The primary measures of association included the change in mean child length for age Z-score (LAZ) given the application of the individualized optimal treatment rule, assessment of individual biomarkers' and pathogens' association with treatment heterogeneity, and the difference in stratified conditional average treatment effect (treatment effect difference) comparing high (above zero for pathogens, above median for EED biomarkers) versus low (zero for pathogens, below median for EED biomarkers) pathogen and EED biomarker status.

Results: We analyzed data from 1,522 children, who had median LAZ of -1.56. We found that EED biomarkers and pathogens were associated with treatment heterogeneity for all interventions, and myeloperoxidase (N+WSH treatment effect difference 0.0007 LAZ, WSH treatment effect difference 0.1032 LAZ, N treatment effect difference 0.0037 LAZ) and *Campylobacter* infection (N+WSH treatment effect difference 0.0011 LAZ, WSH difference 0.0119 LAZ, N difference 0.0255 LAZ) were associated with greater effect of all interventions on growth. Children with both detected *Campylobacter* and above median myeloperoxidase (N+WSH difference 0.039 LAZ, WSH difference 0.106 LAZ, N difference 0.022 LAZ) had a greater N+WSH and WSH treatment effect than children with either factor alone. We found that a treatment rule that assigned the N+WSH (LAZ difference 0.23, 95% CI (0.05, 0.41)) and WSH (LAZ difference 0.17, 95% CI (0.04, 0.30)) interventions based on baseline covariates, EED biomarkers and pathogens, rather than randomly, increased predicted child growth.

Conclusions: These findings indicate that EED biomarker and pathogen status, particularly *Campylobacter* and myeloperoxidase (a measure of gut inflammation), may be related to impact of N+WSH, WSH, and N interventions on child linear growth, although the magnitude of these relationships was small to modest. Children with greater *Campylobacter* or myeloperoxidase burden at 14 months experienced a greater impact of the interventions on growth at 28 months, which is consistent with the findings of a multi-site birth cohort study (MAL-ED). This

contributes to the body of literature suggesting that EED may be a key factor associated with N+WSH, WSH, and N intervention effectiveness.

2.2 Introduction

Approximately 156 million children globally experience linear growth faltering, which can have lifelong consequences (29). Studies have consistently found a positive relationship between child growth and child development, leading investigators to use child linear growth as a proxy for overall development (30,173). In adulthood, children who experienced early-life growth faltering are more likely to experience low educational attainment and low income (30). Children of mothers who are stunted have an increased risk of experiencing stunting themselves, which can perpetuate the cycle of poverty (31).

Water, sanitation, hygiene, and nutrition

Experts in public health and international development have identified water, sanitation, hygiene (WSH), nutrition (N), and combined (N+WSH) programs as potentially effective methods to improve child growth. WSH interventions aim to reduce children's exposure to pathogens, which can improve nutrient utilization by reducing malabsorption, redirection of nutrients for immune response, and other symptoms associated with infection, while nutrition interventions aim to directly provide nutrient supplementation (174,31). The United Nations has established universal access to WSH by the year 2030 as a Sustainable Development Goal (38). Despite the widespread implementation of N and WSH interventions, based on the assumption that these interventions improve child growth, few randomized controlled trials have evaluated the impact of N+WSH, WSH, and N interventions on child growth.

Several observational studies indicated a positive relationship between household WSH interventions and child growth (39). In contrast to these observational findings, the WASH Benefits study, which enrolled pregnant mothers and young children in rural Bangladesh and Kenya, and the SHINE (Zimbabwe) randomized controlled trial found that household WSH interventions did not improve child linear growth in a randomized context (29,31,40,41). These findings suggested that positive associations between WSH and child growth in observational settings may be due to residual confounding. The null effect of these environmental interventions on growth indicated the possibility that additional sources of growth impairment might exist for children facing extreme poverty. Alternatively, the lack of impact of these interventions may reflect an inability of these household interventions to sufficiently reduce pathogen exposure and environmental enteric dysfunction (40).

The WASH Benefits study found that nutritional supplementation led to modest improvements in child linear growth compared to control (40). This is consistent with other randomized controlled trials in low and middle-income countries, which have also found that early nutritional supplementation can improve child growth (31,42,43). The combined N+WSH intervention did not provide any additional benefit to child linear growth compared to the nutrition intervention alone (40). The authors indicated that this small and variable impact of nutrition interventions on child linear growth may be due to contextual underlying factors that influence participants' amenability to nutrition interventions (40).

Effect Measure Modification by EED and Pathogens

In addition to finding a null main effect of WSH interventions on growth and modest effects of nutrition (and N+WSH) on growth, the WASH Benefits study did not detect significant effect

modification of interventions by child age, child sex, maternal education, maternal age, child parity, economic factors, or child hunger (40). Despite this lack of evidence of interaction, pathogen and environmental enteric dysfunction (EED) biomarker data may provide additional information on which subgroups of children, defined by pathogen or biomarker levels, are amenable or resistant to intervention.

EED is a condition characterized by increased gut permeability, gut barrier disruption, increased gut and systemic inflammation, and is hypothesized to be caused by chronic exposure to pathogens (45,46). Although clear diagnostic criteria for EED have not been established, several studies have speculated that it could be a key intermediate between poverty and growth impairment for children in low and middle-income countries (45,46). Observational data and animal models have indicated that *Campylobacter* infection may contribute to EED (175). Among young children in Bangladesh, small intestine bacterial overgrowth is associated with both intestinal inflammation, a key component of EED, and child growth impairment (176,177). The WASH Benefits Bangladesh study found that the nutrition intervention was associated with reduction of neopterin at 3 and 14 months of age, and all interventions reduced lactulose and mannitol at 3 and 14 months (178). At 28 months, contrary to a-priori hypotheses, WSH and nutrition interventions were associated with increased myeloperoxidase, and WSH was associated with increased mannitol (178). Although these findings at age 3 and 14 months support N+WSH interventions' ability to reduce some EED biomarkers, the counterintuitive results at 28 months highlight uncertainty regarding the relationship between N+WSH interventions and presumed biomarkers for EED.

Investigators of the WASH Benefits study suggested that insufficient reduction of pathogen exposure could explain the null effects of WSH interventions on child linear growth (40). Investigation of the relationships between N+WSH interventions and enteropathogens at Year 1 (age 14 months) in Bangladesh found that children who received WSH interventions had a lower prevalence and quantity of some individual viruses (norovirus, sapovirus, and adenovirus 40/41) compared to children in the control group, although investigators did not find a significant difference in bacteria, parasites, or stunting-related pathogens between these groups (44). Furthermore, this study found that 99% of children at Year 1 had at least one enteropathogen (44). At Year 2 (age 31 months), investigators found that individual sanitation and hygiene interventions were associated with decreased *Giardia* infections and that drinking water and nutrition interventions were not associated with a change in *Giardia* infections (179). Regarding soil-transmitted helminths, investigators found that the drinking water intervention was associated with reduced hookworm (180). Lastly, analysis of interventions and fecal contamination found that drinking water and handwashing interventions reduced contamination of water and food, but did not reduce contamination of indirect pathways such as child hands and objects, and that combined WSH interventions provided no additional benefit compared to individual interventions (181). These cumulative findings indicate that household WSH interventions can reduce child exposure to certain pathogens, although these results highlight complex relationships between interventions and individual pathogens.

Methodological Utility of Optimal Treatment Regime Analysis

Public health research typically seeks to identify population-level drivers of incidence rates, rather than individual causes of cases (15). But, methodological advances have enabled the creation of dynamic treatment rules, where susceptible individuals can be targeted for interventions based on individual characteristics or treatment history (16,17). Despite this focus

on optimizing interventions based on individual covariate information, we retain the public health goal of maximizing population-level health outcomes (16). Even if there is a true effect of the intervention on the outcome of interest among certain individuals, a study may fail to detect this relationship if the effect is heterogeneous in the study sample or the subgroup of amenable individuals is small. We can assess the variance of the stratum-specific treatment effect to evaluate treatment heterogeneity (18).

Targeted maximum likelihood estimation is a doubly-robust method that optimizes the bias-variance tradeoff in estimating a specific parameter of interest (26). This method of estimation is optimally efficient when the data generating distribution (DGD) is correctly specified, and its doubly-robust properties ensure consistent results as long as either part of the DGD, the outcome regression or treatment mechanism (propensity score), is estimated consistently (26,27). We can gain additional insight through analysis of optimal individualized treatment effect, where we seek to maximize population outcomes by assigning treatment based on individual characteristics that are associated with the most beneficial treatment effect (25). Estimation of this optimal individualized treatment effect has gained popularity with the rise in precision health, but much of these efforts have relied on unrealistic parametric assumptions (19–24). If the parametric model is incorrect (which is inevitable), the resulting estimates will be biased for the parameters of interest (e.g., average treatment effect) (25). Using targeted learning methods, we can assess the mean outcome, where the candidate treatment rules can be estimated on the same data for which the impact of the rule is also estimated, using a robust cross-validated estimation procedure (25). One can gain efficiency by constraining the statistical model when the constraints are true, so the only restrictions that we will place on the data distribution relate to the probability of a participant receiving treatment (randomized assignment) (25). The use of cross-validated targeted maximum likelihood estimator (CV-TMLE) for the mean outcome under optimal individualized treatment uses a data-adaptive estimation of both the DGD and the rule, while still providing valid inference on the impact of the optimal treatment without making parametric assumptions.

Using data from the WASH Benefits Bangladesh study, analysis of treatment heterogeneity through estimating conditional average treatment effects and optimal treatment regimens can improve our understanding of child growth in low- and middle-income countries. Despite widespread use of N+WSH interventions, investigators have found mixed evidence regarding these interventions' impact on child growth (29,31,40). This study will apply targeted machine learning methods to assess the conditional treatment effect of N+WSH, WSH, and N interventions on child linear growth (child length for age Z score (LAZ)) by pathogen and EED biomarker status and explore rules for the optimal allocation of N+WSH, WSH, and N interventions in resource-constrained settings.

2.3 Methods

Study design, participants, and interventions

This analysis involves data from a substudy of the WASH Benefits Bangladesh randomized controlled trial. The trial randomized pregnant mothers and their children to receive one of six interventions – water treatment, sanitation, handwashing, nutrition (N), combined water treatment, sanitation and handwashing (WSH) and combined nutrition plus WSH (N+WSH), or control (40). In a substudy focused on evaluation of EED, investigators assessed additional

biomarker data in a subset of children in four of the study arms – N, WSH, N+WSH, and control (with an allocation ratio of 1:1:1:1) (178).

Intervention promoters, who were residents of the study area, visited participants to promote intervention behaviors at the level of the compound (cluster of nearby houses). Each promoter received at least five days of training prior to visiting compounds, and received periodic refresher courses throughout the intervention period. The behavioral components of these interventions included treating drinking water for children less than 3 years of age (water), using latrines and removing animal feces from the compound (sanitation), washing hands with soap before preparing food and after defecating or contacting feces (hygiene), and practicing age-appropriate nutrition practices from pregnancy up until two years of age and using lipid nutrient supplements for children six months to two years of age (nutrition) (182). These promoters used various strategies to promote intervention behaviors. For example, promoters promoted the hygiene intervention (handwashing) by framing it as a nurturing intervention that was facilitated by the handwashing station and soap provided by the intervention (182,183). Promoters were instructed to visit study compounds at least once per week for the first six months, and then once every two weeks for the following 1.5 years. The intervention hardware and consumables were provided free of charge and replenished by promoters as needed throughout the study period (additional details on interventions can be found in Supplemental Material 1).

Investigators followed the cohort of children for approximately 2.5 years after birth. It was not feasible to retain the geographic matching of the parent trial in this subset due to logistical challenges regarding specimen collection and transportation. The trial was conducted in contiguous rural subdistricts in Gazipur, Mymensingh, Tangail and Kishoreganj districts of Bangladesh. The trial enrolled mothers in their second trimester of pregnancy (additional information on recruitment and eligibility can be found in Supplemental Material 2) (182).

Covariates:

Although randomization of participants led to a balanced distribution of covariates between study arms, this analysis conditioned on post-randomization biomarker values, leading to the possibility of collider stratification bias. Therefore the primary outcome was adjusted for additional covariates related to socioeconomic status, maternal health, living conditions, child age at data collection, and month of data collection. We considered and tested potential confounders using super learner and cross-validated targeted maximum likelihood estimation. The full list of baseline and time-varying covariates can be found in Supplemental Material 3.

Biomarkers:

EED Biomarkers

The EED measures included in this study were fecal alpha-1-antitrypsin, myeloperoxidase, and regenerating gene 1β (REG1B). These measures are markers of intestinal permeability (alpha-1-antitrypsin), inflammation (myeloperoxidase), and intestinal repair (REG1B) (184). We excluded EED biomarkers (neopterin, lactulose, and mannitol) that were associated with the interventions of interest in a previous analysis of this sample and therefore were potential mediators of the exposure-outcome relationship (178).

To reduce inter-laboratory variation, all feces samples were assayed by the same research team member at the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b)

laboratory. All biomarkers were assessed at ages 3 and 14 months. Laboratory methods are included in Supplemental Material 4 and were published previously (44,178).

Pathogens

The pathogens of interest were relative concentrations of 34 viral, bacterial, and parasitic enteropathogens assessed at 14 months in feces using quantitative polymerase chain reaction (PCR) via TaqMan array card (44,185,186). We excluded three pathogens (norovirus, sapovirus, and adenovirus 40/41) that were associated with the interventions of interest in a previous analysis of this sample and therefore were potential mediators of the exposure-outcome relationship (44). We quantified pathogens via quantification cycle, where one unit represented twice the pathogen quantity, and the analytical limit of detection was at quantification cycle 35 (187). We standardized these measures using the efficiency of per-sample extraction/amplification. The full list of pathogens is included in Supplemental Material 5.

A single infection event is unlikely to elicit growth impairment in itself, but repeated exposure to pathogens and chronic disruptions such as EED are associated with delayed growth (188–190). This analysis assumes that the detection of pathogens and EED biomarkers at 14 months indicates an increased likelihood of chronic exposure to these factors throughout early childhood.

Outcomes:

The growth outcome was length for age Z-score (LAZ) assessed at Year 2 (median age 28 months). Following standard protocols for anthropometric outcomes measurement (191,192), pairs of trained anthropometrists measured recumbent length (accurate to 0.1 cm) in triplicate to calculate median growth using 2006 WHO child growth standards (40).

Analyses:

These analyses assessed the conditional average treatment effect and optimal individualized treatment regime using a targeted learning approach (25). A static treatment approach, as used in the WASH Benefits primary analysis, aims to assess the counterfactual impact of uniform interventions (regardless of individual covariate information) (25,40). In contrast, an optimal treatment regime analysis assesses the impact of the intervention given individual covariate status (25,40). In these analyses, the individual covariate information was child pathogen and EED biomarker status. We used cross-validated targeted maximum likelihood estimation (CV-TMLE) and super learner (SL) ensemble machine learning for estimation to define the optimal individualized treatment regime (18). In our learner list for the treatment mechanism, we included the least absolute shrinkage and selection operator (LASSO) learner, random forest learner, mean learner, generalized linear model (fast) learner, and non-negative least squares to construct the final ensemble (the meta-learner) (25). In our learner list for both the outcome regression and the individualized treatment effect, we included LASSO learner, random forest learner, generalized linear model (fast) learner, and non-negative least squares as the meta-learner (25).

First, we estimated the outcome regression function and propensity score (treatment mechanism) using SL. Next, we used doubly-robust augmented inverse probability weighting to transform the outcome to a random variable that has as the conditional average treatment effect (CATE) as its mean and regressed this transformed outcome to assess treatment heterogeneity. Specifically, we estimated the function of the individualized outcome by regressing this contrast on biomarker status using SL with a non-negative least squares loss function based on the Lawson-Hanson

algorithm. As these analyses assess the impact of the randomized intervention (the treatment mechanism), the doubly-robust nature of this estimator will ensure asymptotically consistent estimation of the CATE even if the outcome regression is not consistently estimated (25). Finally, we use the estimate of the CATE function to derive an optimal individualized treatment rule where we would treat individuals with the greatest CATE.

As providing optimal treatment to all children may be desirable, in this circumstance, one might also be interested to limit the intervention to the children most likely to benefit from the intervention (i.e., have the greatest CATE). In order to assess the impact of the individualized treatment regime in resource-constrained settings (i.e., preventing all children from being allocated to intervention), we restricted the maximum allocation to treatment in each binary (treatment to control) contrast to be no more than 50%, which is approximately equivalent to the original trial's allocation ratio (1:1:1:1). If less than 50% of individuals in a single binary (treatment to control) contrast have a positive CATE (beneficial effect of treatment), then the optimal treatment rule will assign all individuals with a positive CATE to intervention. If more than 50% of individuals in a single contrast have a positive CATE, the optimal treatment rule will only assign the 50% of individuals with the greatest CATE to intervention.

In order to assess the role of each biomarker or pathogen in the optimal treatment rule, we evaluated Pearson's correlation between each of these covariates and the CATE. In order to contextualize the magnitude of these relationships, we estimated the stratified CATE among children with high (above zero for pathogens, above median for EED biomarkers) versus low (zero for pathogens, below median for EED biomarkers) pathogen and EED biomarker status, hereafter referred to as "treatment effect difference."

Covariate screening

We screened all covariates for missingness, excluding all covariates with missingness greater than 30% and median-imputing all other missing covariate data. We only included observations for which the primary outcome, length for age z-score at 28 months, was observed. We also excluded variables with near zero variance, which we defined as covariates with a frequency ratio (ratio of most frequent value to second most frequent value) greater than 2 and a percent of unique values less than 20%, using the R package "caret" (version 6.0-92) (193). The analysis plan was publicly pre-registered on Open Science Framework (osf.io/qtdm8/).

2.4 Results

We analyzed data from 1,522 children, and our analytic sample had a median LAZ of -1.56 at Year 2 (median age 28 months; Table 1).

Our treatment rule included the pathogens *Campylobacter jejuni/coli*, enteroaggregative *Escherichia coli*, any enterotoxigenic *Escherichia coli*, atypical enteropathogenic *Escherichia coli*, enteropathogenic *Escherichia coli* spp., and *Campylobacter* spp., as well as EED biomarkers REG1B, myeloperoxidase, and alpha-1-antitrypsin. EED markers assessed at 3 months were excluded due to high missingness (>30%). The full list of excluded covariates and reasons for exclusion are defined in Supplemental Material 6.

In addition to the biomarkers and pathogens included in the treatment rule, our analyses adjusted for child sex, birth order, number of children under 18 years of age in the household, number of individuals in the compound (group of nearby houses), household wall material, household wealth (first principal component of a principal components analysis), maternal age and height,

age in days at urine and stool assessments, month of urine and stool assessments, and age at anthropometry assessment.

Correlation of pathogens and biomarkers with conditional average treatment effect

We found that the following covariates were associated with a greater impact of N+WSH intervention on growth under the optimal treatment rule: enterotoxigenic *Escherichia coli* spp. (correlation 0.45, treatment effect difference 0.0019 LAZ)*, Campylobacter jejuni/coli* (correlation 0.37, treatment effect difference 0.0016 LAZ), *Campylobacter* spp. (correlation 0.33, treatment effect difference 0.0011 LAZ), REG1B (correlation 0.20, treatment effect difference 0.0005 LAZ), and myeloperoxidase (correlation 0.15, treatment effect difference 0.0007 LAZ) (Table 2). The following covariates were associated with a lower impact of N+WSH intervention: atypical enteropathogenic *Escherichia coli* (correlation -0.41, treatment effect difference -0.0018 LAZ), enteroaggregative *Escherichia coli* (correlation -0.39, treatment effect difference -0.0015 LAZ), alpha-1-antitrypsin (correlation -0.38, treatment effect difference -0.0013 LAZ), and enteropathogenic *Escherichia coli* spp. (correlation -0.22, treatment effect difference -0.0009 LAZ).

The following EED biomarkers and pathogens were associated with greater WSH impact on growth under the optimal treatment rule: myeloperoxidase (correlation 1.00, treatment effect difference 0.1032 LAZ), alpha-1-antitrypsin (correlation 0.26, treatment effect difference 0.0259 LAZ), REG1B (correlation 0.17, treatment effect difference 0.0105 LAZ), *Campylobacter jejuni/coli* (correlation 0.15, treatment effect difference 0.0143), *Campylobacter* spp. (correlation 0.13, treatment effect difference 0.0119 LAZ), enteropathogenic *Escherichia coli* spp. (correlation 0.11, treatment effect difference 0.014 LAZ), and atypical enteropathogenic *Escherichia coli* (correlation 0.08, treatment effect difference 0.0099 LAZ) (Table 3). No EED biomarkers or pathogens were associated with lower WSH treatment effect.

The following EED biomarkers and pathogens were associated with greater impact of N on growth under the optimal treatment rule: *Campylobacter* spp. (correlation 0.17, treatment effect difference 0.0255 LAZ), *Campylobacter jejuni/coli* (correlation 0.15, treatment effect difference 0.0269 LAZ), myeloperoxidase (correlation 0.06, treatment effect difference 0.0037 LAZ), and enterotoxigenic *Escherichia coli* (correlation 0.05, treatment effect difference 0.0098 LAZ) (Table 4). Enteroaggregative *Escherichia coli* (correlation -0.07, treatment effect difference -0.0181 LAZ) was associated with a lower impact of N intervention.

Treatment allocation and predicted child growth

When comparing the combined N+WSH (mean LAZ -1.62) and control (mean LAZ -1.54) arms (n = 756), an optimal treatment allocation assigned 331 children to N+WSH and 425 children to control (Table 5). The optimal treatment rule predicted greater child growth than the observed randomized intervention (growth difference 0.23 LAZ, 95% CI (0.05, 0.41)).

In the contrast of WSH (mean LAZ -1.69) and control (mean LAZ -1.54) arms (n = 752), the optimal treatment rule assigned 9 children to receive WSH interventions and 743 children to receive control. The optimal treatment rule had greater predicted child growth than the observed randomized, static intervention (growth difference 0.17 LAZ, 95% CI (0.04, 0.3)).

After comparing the nutrition (mean LAZ -1.53) and control (mean LAZ -1.54) arms (n = 726), the optimal treatment rule assigned 317 children to receive the intervention and 409 children to be in the control group. The optimal treatment rule did not have significantly greater child

growth compared to the observed randomized intervention (growth difference 0.07 LAZ, 95% CI (-0.09, 0.22).

Post-hoc analysis

Campylobacter spp. and myeloperoxidase were associated with a greater treatment effect across all three interventions (Supplemental Materials 7-12). We conducted an exploratory evaluation of the combined impact of *Campylobacter* infection (any detection) and high myeloperoxidase (above median concentration) on the conditional treatment effect under the optimal treatment rule (Table 6). The difference in treatment effect, comparing those with both *Campylobacter* spp. infection and high myeloperoxidase to those with no *Campylobacter* spp. detection and below median myeloperoxidase, was 0.039 LAZ for N+WSH, 0.106 LAZ for WSH, and 0.022 LAZ for N.

2.5 Discussion

These findings highlight the potential for targeted learning methods to identify and explore treatment heterogeneity within a study and for optimal treatment regime analysis to estimate the effects of targeting treatments to children who would benefit the most when resource constraints prevent intervening on all children. These findings provide empirical support for the notion that pathogen exposure contributes to EED, EED contributes to growth faltering, and that the interruption of these processes is protective. Within rural Bangladesh, these effects were small, but they provide support for a biological mechanism.

Across all three interventions, myeloperoxidase, an EED biomarker of gut inflammation, and *Campylobacter*, one of the leading causes of bacterial diarrhea, were associated with a greater treatment effect (184,194). In other words, children with the greatest burden of *Campylobacter* infection and myeloperoxidase experienced the greatest benefit from the interventions, although the magnitude of these differences in treatment effects was typically small. There was a greater N+WSH and WSH treatment effect among those with both *Campylobacter* infection and high myeloperoxidase than those with either factor alone. The correlation of both *Campylobacter* and myeloperoxidase biomarkers with the treatment effect indicates that these factors, implicated as a cause (*Campylobacter*) and a marker (myeloperoxidase) of EED, supports that EED may play a role in the interventions' impact on growth (188).

We found that children with *Campylobacter* infection, compared to children with no *Campylobacter*, experienced a greater impact of the N+WSH, WSH, and N interventions. This is consistent with previous findings that young children with *Campylobacter* infection may face increased risk of growth impairment and are therefore a high-need group for intervention. A multi-site birth cohort study (MAL-ED), which found that *Campylobacter* infection was highly prevalent and was associated with child growth shortfall in the first two years of life (195,189). While *Campylobacter* infection is typically associated with food contamination in high-income settings, it is endemic in low income settings, and even asymptomatic infection is associated with growth shortfall (189). *Campylobacter* alters the gut microbiota composition, disrupts the intestinal barrier, and elicits chronic intestinal inflammation (196–201). Across eight study sites in low-resource settings, MAL-ED found that breastfeeding, access to WSH, and targeted antibiotic treatment were associated with *Campylobacter* infection (189). In addition, these investigators found that *Campylobacter* infection was associated with increased intestinal permeability, intestinal inflammation, and systemic inflammation, which are key components of EED (189).

We found that myeloperoxidase, an EED marker of gut inflammation, was associated with a greater impact of N+WSH, WSH, and N interventions on child growth. Regarding the nutritional supplementation interventions, this is consistent with MAL-ED's findings of a heterogeneous relationship between nutrient intake and micronutrient deficiency based on EED status (202). Regarding WSH, our findings were also consistent with MAL-ED's findings that EED and inflammation likely mediated the relationship between infection and growth faltering (188). In addition, MAL-ED investigators found that myeloperoxidase was associated with pathogen infection, and more specifically, that *Campylobacter* and myeloperoxidase were positively associated across all eight study sites (203).

After comparing both WSH and combined N+WSH interventions to control, we found that an optimal treatment rule selected via cross-validation and based on EED and pathogen status led to greater expected mean child growth than the observed, randomized intervention. This indicates that pathogen and EED biomarker status may define, in part, which children are responsive to WSH and N+WSH interventions.

Strengths

This manuscript provides a roadmap for researchers who hope to use targeted learning and optimal treatment regime analyses to assess the drivers of treatment heterogeneity within a study. This applied example demonstrates the utility of these methods to flexibly model these relationships despite high dimensionality. These methods may be applicable for a range of other research areas, including targeted medicine, adaptive trials, and analyses of secondary data in both randomized and observational contexts.

The rich data source of the WASH Benefits Bangladesh EED substudy is a major strength of this analysis. This data source included in-utero randomized interventions that were continued for two years after birth and robust collection of enrollment covariates, EED biomarkers, pathogens, and growth outcomes across multiple timepoints.

The analysis methods are a second major strength of this study. We used targeted maximum likelihood estimation, which is maximally efficient in finite samples and doubly-robust (26,27). Assessment of optimal individualized treatment effects allows us to evaluate the relationships between pathogen exposure, EED, and intervention effects without making parametric assumptions (19–25). Given the high-dimensionality of these biomarker and pathogen data, these targeted learning methods allow flexible modeling of complex relationships without requiring parametric assumptions regarding relationships between interventions, biomarkers, pathogens, and child growth that would inevitably be violated.

Future directions and limitations

These findings highlight the utility of applying targeted learning methods to explore treatment heterogeneity in a population. In future evaluations, this may support the use of co-interventions, where children who show little responsiveness to N, WSH, or N+WSH interventions could receive a co-intervention in order to increase amenability to treatment. While we identified biomarkers and pathogens that indicated greater treatment effect, consistent identification of biomarkers associated with lower treatment effect (i.e., resistance to treatment) could indicate the need for co-interventions. For example, certain types of persistent bacterial infection (e.g., *Mycobacterium tuberculosis* or *Salmonella typhi*) may not be responsive to WSH interventions, and may require additional medical intervention (204–206). In these cases, co-interventions, such

as antibiotic treatment, may supplement interventions in order to ameliorate these conditions and improve N+WSH, WSH, or N intervention effectiveness (204).

We focused our interpretation on *Campylobacter* and myeloperoxidase, which demonstrated consistent correlations (in terms of direction) with the CATE across interventions. Our analysis of individual biomarkers' and pathogens' correlations with the conditional treatment effect provided some evidence of effect heterogeneity being associated with factors beyond *Campylobacter* and myeloperoxidase, although the lack of consistency of these observations across similar covariates (e.g., N+WSH and any enterotoxigenic *Escherichia coli* versus enteroaggregative *Escherichia coli*) and similar interventions (e.g., N+WSH versus WSH) led us to believe that these relationships may be spurious. On the other hand, it is plausible that these unique correlations across similar biomarkers point to unique actions of related covariates or unique mechanisms of combined versus individual interventions, respectively. Future studies could incorporate cluster analysis methods to assess the combined role of related biomarkers and pathogens on treatment effectiveness.

One limitation of this study arises from using post-intervention biomarkers, as no baseline EED biomarkers or pathogens were measured. Conditioning on these post-intervention nodes potentially introduces confounding and bias. We accounted for this possible confounding by adjusting for additional baseline covariate information related to family health and socioeconomic status and by excluding pathogens and EED biomarkers that were associated with the interventions in previous analyses of this sample (i.e., potential mediators or colliders), although residual confounding or bias may be present. However, the identification of these relationships remains useful for generating hypotheses about the causes of N+WSH, WSH, and N treatment heterogeneity. In the future, we hope to analyze biological samples that were collected from these children at a younger age (4-8 months) in order further evaluate these relationships.

The small or null overall effects of the study interventions is another limitation. In the presence of a null overall effect, in order to detect subpopulations that have a significant effect, there must be equivalent populations with a deleterious effect (implausible for N+WSH interventions) or much larger populations with a null effect. In contrast, optimal treatment regime analysis in a population with a greater treatment effect will have much greater power to detect subpopulations of interest. Furthermore, the subsample analyzed here did not retain the same growth characteristics as the total trial population. While the trial reported that N and N+WSH interventions led to a modest improvement in growth (40), these effects were not seen for this subsample. This may be the reason that more than half of the children were assigned to control rather than the interventions in the optimal treatment regime. These findings should be taken as a finite sample limitation of a trial with null effects on children within the small substudy, not as an indication that N+WSH, WSH, or N interventions could be harmful. Follow up evaluation of these relationships in a separate population may provide insight on the replicability of these findings.

2.6 Conclusion

This analysis provides an example of how targeted learning methods can explore treatment effect heterogeneity, and the cumulative results here indicate that EED and pathogens may be related to N+WSH, WSH, and N interventions' impact on child growth. In particular, we found that *Campylobacter* infection and high myeloperoxidase were associated with a greater effect of

N+WSH (treatment effect difference 0.039 LAZ), WSH (treatment effect difference 0.106 LAZ), and N (treatment effect difference 0.022 LAZ) interventions on child LAZ at 28 months. These findings are consistent with the MAL-ED study (188,189,202). These results may help distinguish what defines a responsive versus nonresponsive individual to N+WSH, WSH, and N interventions and should motivate future etiological research that seeks to estimate the causal impact of EED and pathogen burden on intervention effectiveness.

2.7 Tables and Figures

Table 1. Descriptive statistics of sample population.

			n (%) or median (IQR)
Child		Female	748 (49%)
	Anthropometry (14 months, Year 1)	Length-for-age z-score	-1.41 (-2.06, -0.74)
		Weight-for-age z-score	-1.31 (-2.01, -0.63)
		Weight-for-length z-score	-0.89 (-1.55, -0.21)
		Head circumference-for-age z-score	-1.78 (-2.34, -1.12)
	Anthropometry (28 months, Year 2)	Length-for-age z-score	-1.56 (-2.27, -0.94)
		Weight-for-age z-score	-1.58 (-2.2, -0.93)
		Weight-for-length z-score	-1.03 (-1.62, -0.38)
		Head circumference-for-age z-score	-1.81 (-2.39, -1.2)
	Diarrhea (14 months, Year 1)	Caregiver-reported 7-day recall	192 (13%)
	Diarrhea (28 months, Year 2)	Caregiver-reported 7-day recall	114 (7%)
Mother		Age (years)	23 (20, 27)
	Anthropometry at enrollment	Height (cm)	150.28 (146.81, 154.15)
	Education	Schooling completed (years)	7 (4, 9)
	Depression at Year 1	CES-D score	9 (6, 16)
	Depression at Year 2	CES-D score	10 (5, 17)
	Perceived stress at Year 2	Perceived Stress Scale score	14 (10, 18)
	Intimate partner violence	Any lifetime exposure	835 (57%)

IQR, interquartile range; CES-D, Center for Epidemiologic Studies Depression Scale

Table 2. Biomarker and Pathogen Correlation with NWSH Conditional Average Treatment effect

Biomarker or pathogen	Correlation	Treatment effect (LAZ difference) at non-detection (pathogen) or below median (EED biomarker)	Treatment effect (LAZ difference) at detection (pathogen) or above median (EED biomarker)	Difference in Treatment effect (LAZ difference)
Enterotoxigenic Escherichia coli spp.	0.45	-0.0006	0.0013	0.0019
Campylobacter jejuni/coli	0.37	-0.0004	0.0013	0.0016
Campylobacter spp.	0.33	-0.0004	0.0008	0.0011
REG 1B	0.20	-0.0003	0.0003	0.0005
Myeloperoxidase	0.15	-0.0003	0.0004	0.0007
Enteropathogenic Escherichia coli spp.	-0.22	0.0006	-0.0003	-0.0009
Alpha-1-antitrypsin	-0.38	0.0007	-0.0006	-0.0013
Enteroaggregative Escherichia coli	-0.39	0.0012	-0.0002	-0.0015
Atypical enteropathogenic Escherichia coli	-0.41	0.0006	-0.0011	-0.0018

N+WSH, Combined nutrition, water, sanitation, and hygiene intervention; REG1B, Regenerating gene 1β; LAZ, length-for-age z score; EED, environmental enteric dysfunction

Table 3. Biomarker and Pathogen Correlation with Conditional Average WSH Treatment effect

Biomarker or pathogen	Correlation	Treatment effect (LAZ difference) at non-detection (pathogen) or below median (EED biomarker)	Treatment effect (LAZ difference) at detection (pathogen) or above median (EED biomarker)	Difference in Treatment effect (LAZ difference)
Myeloperoxidase	1.00	-0.1973	-0.0941	0.1032
Alpha-1-antitrypsin	0.26	-0.1586	-0.1328	0.0259
REG 1B	0.17	-0.151	-0.1405	0.0105
Campylobacter jejuni/coli	0.15	-0.1486	-0.1343	0.0143
Campylobacter spp.	0.13	-0.1494	-0.1375	0.0119
Enteropathogenic Escherichia coli spp.	0.11	-0.1535	-0.1395	0.014
Atypical enteropathogenic Escherichia coli	0.08	-0.1483	-0.1384	0.0099
Enteroaggregative Escherichia coli	0.04	-0.1522	-0.1434	0.0088
Enterotoxigenic Escherichia coli spp.	0.03	-0.1452	-0.1445	0.0007

WSH: water, sanitation, and hygiene intervention; REG1B, Regenerating gene 1β; LAZ, length-for-age z score; EED, environmental enteric dysfunction

Table 4. Biomarker and Pathogen Correlation with Nutrition Conditional Average Treatment Effect

Biomarker or pathogen	Correlation	Treatment effect (LAZ difference) at non-detection (pathogen) or below median (EED biomarker)	Treatment effect (LAZ difference) at detection (pathogen) or above median (EED biomarker)	Difference in Treatment effect (LAZ difference)
Campylobacter spp.	0.17	-0.008	0.0175	0.0255
Campylobacter jejuni/coli	0.15	-0.0049	0.0221	0.0269
Myeloperoxidase	0.06	0.0038	0.0075	0.0037
Enterotoxigenic Escherichia coli spp.	0.05	-0.0015	0.0083	0.0098
REG 1B	0.04	-0.0046	0.016	0.0207
Atypical enteropathogenic Escherichia coli	-0.01	0.001	0.0026	0.0017
Enteropathogenic Escherichia coli spp.	-0.01	0.0053	-0.0008	-0.006
Alpha-1-antitrypsin	-0.05	0.0102	0.0007	-0.0095
Enteroaggregative Escherichia coli	-0.07	0.0157	-0.0024	-0.0181

REG1B, Regenerating gene 1β; LAZ, length-for-age z score; EED, environmental enteric dysfunction

Table 5. Average child growth given optimized vs randomized treatment.

Study arms	*n*	Observed growth in treatment arm	Optimal allocation ratio (treatment: control)	Overall observed child growth	Optimized child growth	Predicted growth difference
N+WSH vs. control	756	-1.62	(331:425)	-1.58	-1.35 (-1.53 , -1.17)	0.23 (0.05 , 0.41)
WSH vs. control	752	-1.69	(9:743)	-1.62	-1.45 (-1.58 , -1.32)	0.17 (0.04 , 0.3)
Nutrition vs. control	726	-1.53	(317:409)	-1.53	-1.47 (-1.62 , -1.31)	0.07 (-0.09 , 0.22)

N+WSH, combined nutrition, water, sanitation, and hygiene intervention; WSH: water, sanitation, and hygiene intervention; LAZ, length for age Z-score

Table 6. Stratified conditional average treatment effect given levels of both Campylobacter and myeloperoxidase at 14 months

Treatment arm	Treatment effect (LAZ difference) given Campylobacter nondetection and below median myeloperoxidase	Treatment effect (LAZ difference) given Campylobacter detection and above median myeloperoxidase	Difference in treatment effect (LAZ difference)
N+WSH	-0.014	0.026	0.039
WSH	-0.198	-0.092	0.106
Nutrition	-0.003	0.019	0.022

N+WSH, combined nutrition, water, sanitation, and hygiene intervention; WSH: water, sanitation, and hygiene intervention; LAZ, length for age Z-score

2.8 Supplemental Materials

Supplemental Material 1. Study interventions.

We included four study arms: control, combined water treatment, sanitation, and handwashing (WSH), nutrition, and nutrition plus WSH (N+WSH). The control arm was passive, including no visit by a health promoter. The water treatment involved provision of chlorine tablets (Aquatabs; NaDCC) and a safe storage vessel to treat and store drinking water. Sanitation involved upgrading latrines to double pit latrines for all households in study compounds, providing child potties, and sani-scoops to remove feces from households and compounds. The handwashing intervention involved providing handwashing stations near latrines and kitchens, which included soapy water bottles and detergent soap. The nutrition intervention involved the provision of a lipid nutrient supplement and age-appropriate recommendations on maternal nutrition and child feeding.

Supplemental Material 2. Inclusion criteria.

We included households if a resident was a pregnant mother in her first or second trimester, the household was located in rural area that was not fully submerged during monsoon season and did not have water, sanitation, hygiene, or nutrition programs ongoing or planned in the next two years.

We excluded households whose residents had plans to move during the following year, did not own their home, or drew water from a source with high iron content or high arsenic content.

We included child participants who were born to enrolled mothers meeting household inclusion criteria within six months of the baseline survey. We excluded children if their growth score fell outside of the WHO plausible range (192).

Supplemental Material 3. **Adjustment covariates.**

Enrollment characteristics: child sex (male or female), child birth order (first born, second born or greater), maternal age (years), maternal height (cm.) maternal education level (no education, primary, or secondary/greater), household food insecurity (4-level HFIAS categories), number of children in the household less than 18 years old, total number of individuals living in the compound (group of nearby houses), household's distance to primary drinking water source (in minutes), household construction materials (floor, walls, and roof), and an asset-based household wealth index, calculated from the first principal component of a principal components analysis of household assets (electricity, wardrobe, table, chair or bench, khat, chouki, working radio, working black/white or color television, refrigerator, bicycle, motorcycle, sewing machine, mobile phone, land phone, number of cows, number of goats, number of chickens).

Time-varying characteristics: child age (days) at assessment and month of data collection.

Supplemental Material 4. Laboratory methods.

EED Biomarkers

Study staff followed enzyme-linked immunosorbent assay kit protocols for fecal alpha-1-antitrypsin, myeloperoxidase, and REG1B.

Pathogens

The child's primary caregiver collected the fecal sample, it was placed on cold chain within three hours, and then transported on dry ice to the laboratory where it was stored at -80 degrees Celsius (44). We extracted DNA and RNA using QIAamp Fast DNA Stool Mini kit (Qiagen, Venlo, The Netherlands) as well as spike-ins of two extrinsic controls which aimed to assess efficiency of extraction and amplification (185). We assessed enteropathogens at icddr,b using quantitative polymerase chain reaction (PCR) via TaqMan array card (185,186). We quantified pathogens using quantification cycle, where one unit corresponded to twice the pathogen quantity and there was an analytical limit of detection at quantification cycle 35 (187). These quantities were normalized based on the efficiency of per-sample extraction/amplification.

Supplemental Material 5. Pathogens

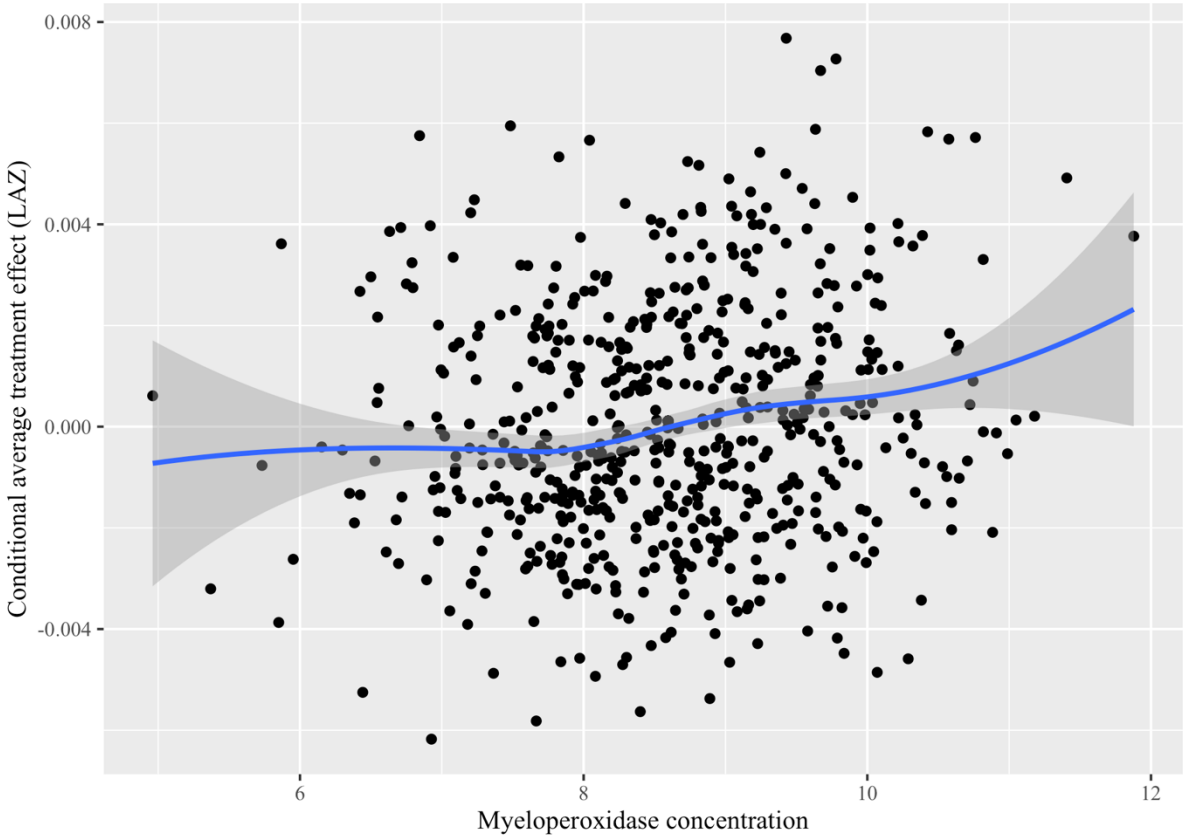
We assessed the relative concentration of the following pathogens: *Campylobacter jejuni/coli*, enteroaggregative *Escherichia coli*, enterotoxigenic *Escherichia coli* spp., atypical enteropathogenic *Escherichia coli*, enteropathogenic *Escherichia coli* spp., *Campylobacter* spp., *E. coli* with heat-stable toxin, typical enteropathogenic *E. coli*, Shiga toxin–producing *E. coli*, *Shigella*/enteroinvasive *E. coli*, *Ancyclostoma, Necator, E. bieneusi, E.intestinalis, E.histolytica*, *Entamoeba* spp*., Giardia, Cryptosporidium, Salmonella, H. nana, Schistosoma, B. fragilis, H. pylori,* rotavirus*, Ascaris, Trichurism, Cyclospora, Isospora, Cryptosporidium hominis, Cryptosporidium parvum, Strongyloides, Blastocystis, V. cholerae, M. tuberculosis, C. difficile, Plesiomonas, Aeromonas*, and astrovirus.

Supplemental Material 6. Covariate, EED biomarker, and pathogen exclusion

We excluded the following variables due to missingness greater than 30%, all of which were assessed at median 3 months of age: lactulose concentration, mannitol concentration, myeloperoxidase, alpha-1-antitrypsin, and age and month of assessment for stool and urine tests.

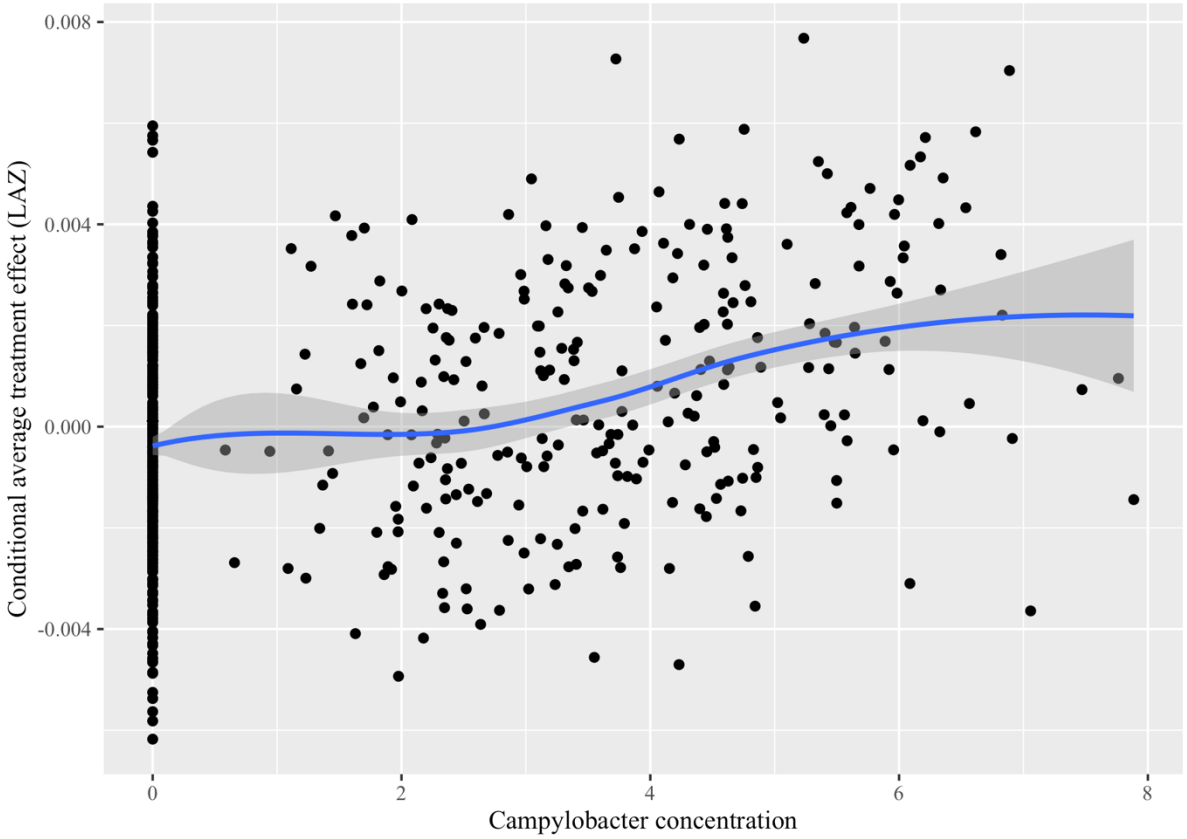
We excluded the following covariates due to near zero variance: maternal education, household food security, distance to water source, household floor material, enterotoxigenic *E. coli* with heat-labile toxin, enterotoxigenic *E. coli* with heat-stable toxin, typical enteropathogenic *E. coli*, Shiga toxin–producing *E. coli*, *Shigella*/enteroinvasive *E. coli*, *Ancyclostoma, Necator, E. bieneusi, E.intestinalis, E.histolytica*, *Entamoeba* spp.*, Giardia, Cryptosporidium, Salmonella, H. nana, Schistosoma, B. fragilis, H. pylori,* rotavirus, *Ascaris, Trichurism, Cyclospora, Isospora, Cryptosporidium hominis, Cryptosporidium parvum, Strongyloides, Blastocystis, V. cholerae, M. tuberculosis, C.difficile, Plesiomonas, Aeromonas,* and astrovirus.

Supplemental Material 7. Correlation of individual N+WSH conditional treatment effect and myeloperoxidase concentration



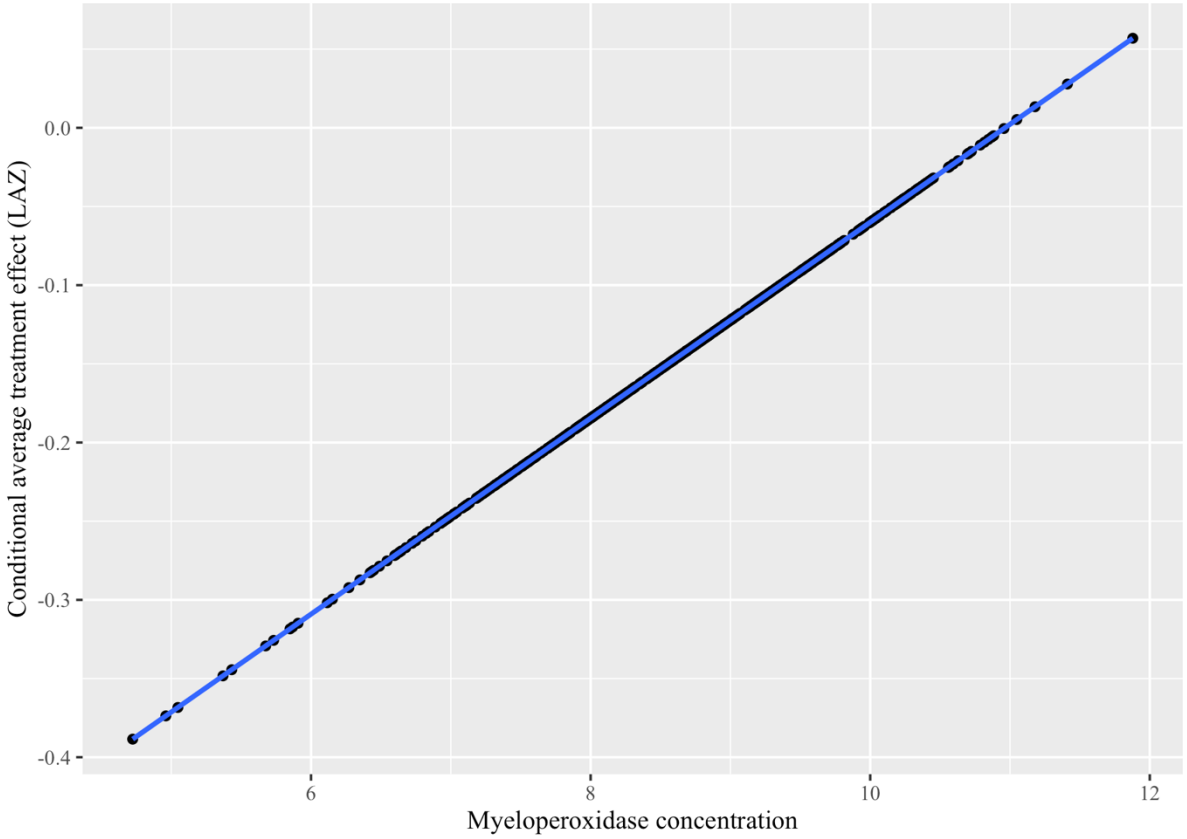
N+WSH, combined nutrition, water, sanitation, and hygiene intervention; LAZ, length-for-age z-score

Supplemental Material 8. Correlation of individual N+WSH conditional treatment effect and *Campylobacter* concentration



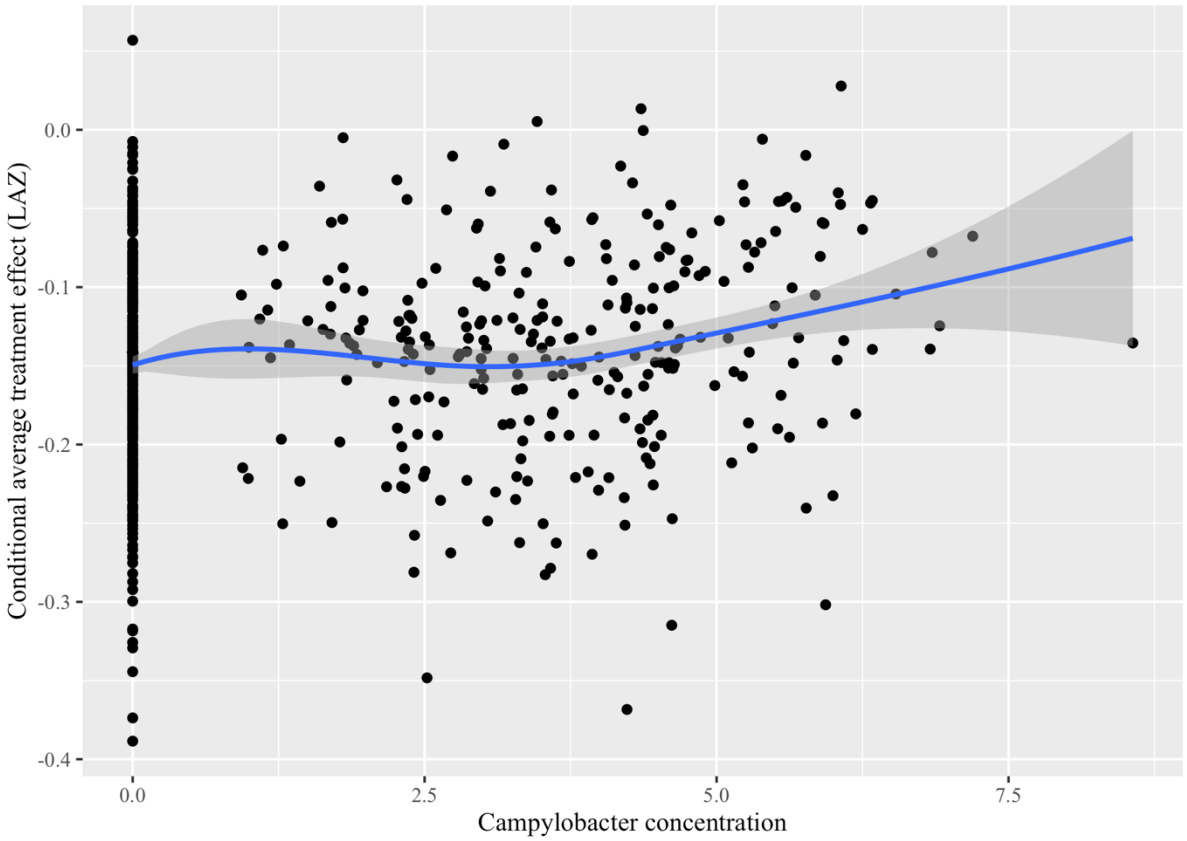
N+WSH, combined nutrition, water, sanitation, and hygiene intervention; LAZ, length-for-age z-score

Supplemental Material 9. Correlation of individual WSH conditional treatment effect and myeloperoxidase concentration



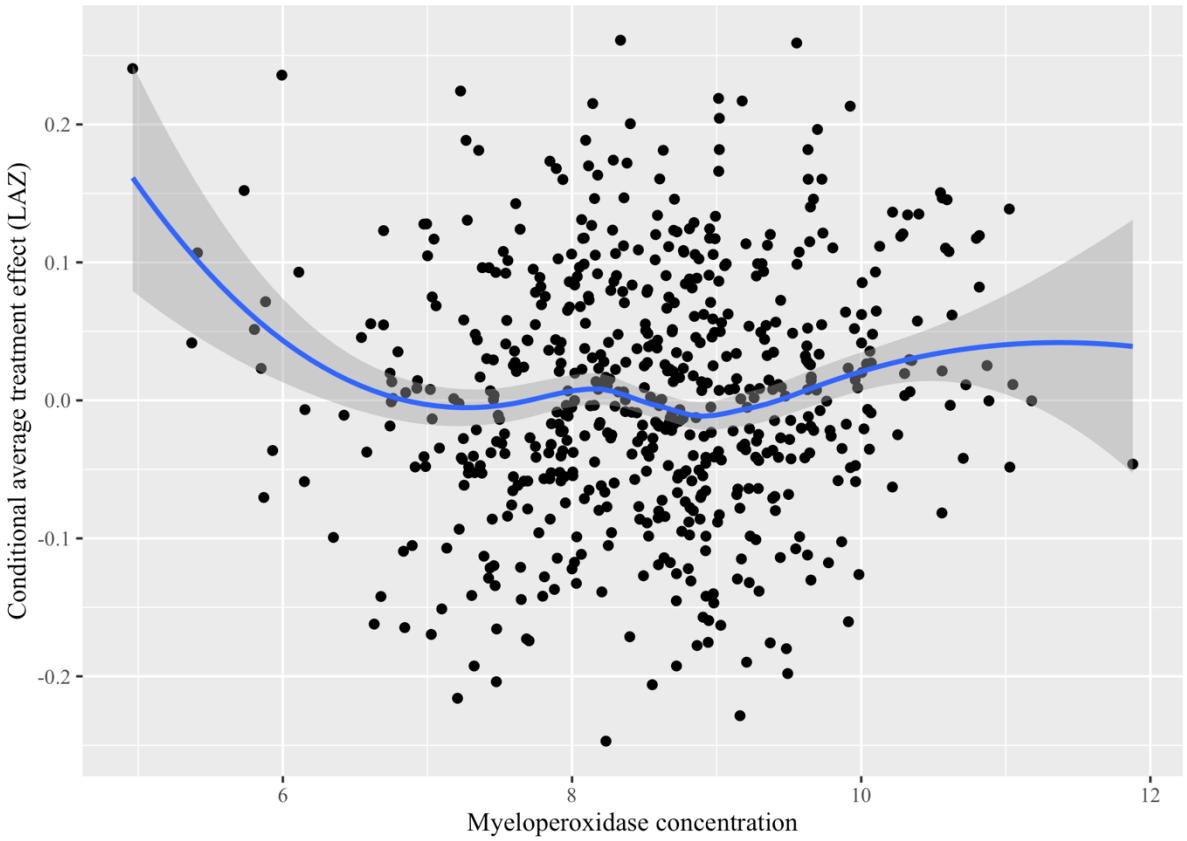
WSH, water, sanitation, and hygiene intervention; LAZ, length-for-age z-score

Supplemental Material 10. Correlation of individual WSH conditional treatment effect and *Campylobacter* concentration



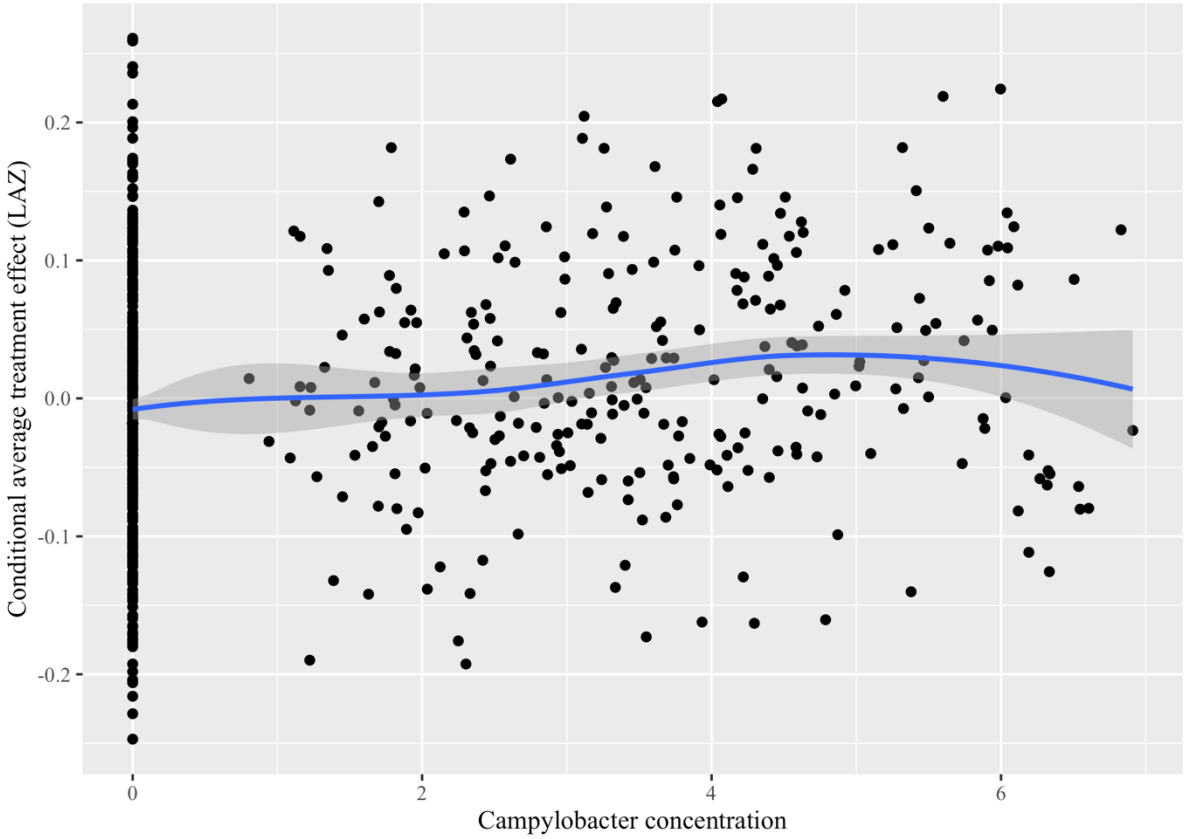
WSH, water, sanitation, and hygiene intervention; LAZ, length-for-age z-score

Supplemental Material 11. Correlation of individual nutrition conditional treatment effect and myeloperoxidase concentration



LAZ, length-for-age z-score

Supplemental Material 12. Correlation of individual nutrition conditional treatment effect and *Campylobacter* concentration



LAZ, length-for-age z-score

3. Chapter 3. Stress Biomarkers and Child Development in Young Children in Bangladesh

3.1 Abstract

Background: Hundreds of millions of children in low- and middle-income countries are exposed to chronic stressors, such as poverty, poor sanitation and hygiene, and sub-optimal nutrition. These stressors can have physiological consequences for children and may ultimately have detrimental effects on child development. This study explores associations between biological measures of chronic stress in early life and developmental outcomes in a large cohort of young children living in rural Bangladesh.

Methods: We assessed physiologic measures of stress using measures of the hypothalamic-pituitary-adrenal (HPA) axis (salivary cortisol and glucocorticoid receptor gene methylation, assessed at Year 2), the sympathetic-adrenal-medullary (SAM) system (salivary alpha-amylase, heart rate, and blood pressure, assessed at Year 2) and oxidative status (F2-isoprostanes, assessed at Year 1). Salivary cortisol and alpha-amylase were collected before and after a standard challenge (venipuncture), while heart rate, blood pressure, urinary F2-isoprostanes, and glucocorticoid receptor methylation (salivary) were assessed in children at rest. Child development was assessed with the MacArthur-Bates Communicative Development Inventories (CDI, assessed at Year 1 and Year 2), the WHO gross motor milestones (assessed at Year 1), and the Extended Ages and Stages Questionnaire (EASQ, assessed at Year 2). Our goal was to compare development outcomes of children at the 75th and 25th percentiles of stress biomarker distributions while adjusting for potential confounders. To do this, we constructed generalized additive models, which are statistical models where the outcome is predicted by a potentially non-linear function of predictor variables. Confounders included baseline measures of household socioeconomic status, living conditions, and maternal factors, as well as timing and age at assessment. Investigators evaluated the consistency of the direction of observed associations across related biomarkers to avoid false discovery.

Results: We analyzed data from 684 children at 14 and 28 months of age; an additional 765 children were included at 28 months of age (49% female). We observed 135 primary contrasts of the differences in child development outcomes at the 75^{th} and 25^{th} percentiles of stress exposures, where we detected significant relationships in 5 out of 30 contrasts (17%) of HPA axis activity, 1 out of 30 contrasts (3%) of SAM activity, and 3 out of 75 contrasts (4%) of oxidative status. Results from generalized additive models revealed that measures of HPA axis activity were associated with poor development outcomes. Although increased pre-stressor salivary alpha-amylase was associated with worse developmental outcomes, we did not find associations between alpha-amylase reactivity or post stressor alpha-amylase and development. We found some evidence that moderate oxidative status was associated with better concurrent child development, but we did not observe a consistent association between oxidative status and subsequent child development. We did not find evidence of an association between heart rate or blood pressure and child development.

Conclusions: Our observations reveal associations between the physiological evidence of stress exposure in the HPA axis with developmental status in early childhood and support the use of HPA axis biomarkers as possible markers of young children at risk of poor development outcomes. These findings add to the existing evidence exploring the developmental consequences of early life stress.

3.2 Introduction

There are more than 250 million children in low and middle-income countries who are at risk of failing to reach their developmental potential (28). Children may not reach their potential in a range of domains, including cognitively, linguistically, or psychologically (30). Biological consequences of chronic stress experienced in early childhood can increase risk for poor developmental outcomes later in life (30,32,33). Chronic early-life stress has the potential to negatively affect multiple biological systems and interfere with learning and memory, dysregulate metabolism and sleep, and potentiate extreme behavior including mental illness (35,36). Given the evidence of associations between stress and development, there is an increased demand for improved understanding of how specific biomarkers can affect development, which may have implications on pediatrics as well as policy (207). Evaluating these associations in early childhood is particularly important, as this is can be a very effective time to intervene on developmental outcomes (34).

The quantification of neural disruptions may be particularly impactful for rural communities where poverty is prevalent, as children in these communities face increased biological, environmental, and psychosocial stressors (37). Social and economic factors can lead to developmental differences in children, and developmental neuroscience demonstrates how early experiences can influence development (34). Poverty and other sociocultural factors can alter neurological functioning, brain structure, and child behavior, which can in turn affect developmental status (37). Inadequate cognitive stimulation is a major risk factor for poor developmental outcomes, and children from low-income homes are less likely to have high quality stimulation at home (34,37). Although cumulative exposure to stressors leads to an increased risk of poor developmental outcomes, the impact of these exposures depends on their timing, co-occurrence, and an individual's reactivity (physiologic response) to these stressors (208). These cumulative stressors are often measured through stress hormones (209).

Stress can be an adaptive or maladaptive response to challenging stimuli. Primary stress biochemical pathways include the hypothalamic-pituitary-adrenal (HPA) and sympathetic adreno-medullary (SAM) axes (210,211). The HPA axis is controlled by a negative feedback loop in which pro-inflammatory cytokines stimulate HPA activation, triggering the release of anti-inflammatory cortisol, a glucocorticoid, which in turn dampens HPA axis activity (211). The cortisol response enables individuals to respond to challenging circumstances, and cortisol is involved in mobilizing biological resources for metabolic, sensory, and learning processes (212). The *NR3C1* gene encodes glucocorticoid receptors, and early life stress is associated with *NR3C1* methylation (213). Prolonged activation of the HPA axis and excess cortisol can lead to oxidative stress, which is an excess of reactive oxygen species relative to antioxidants (214). Urinary F2-isoprostanes are the biomarkers most frequently used to measure oxidative status (215–217). While oxidative stress can be harmful, reactive oxygen species play a critical role in the human body and immune function (216).

The second major component of the psychobiology of the stress response is activation of the SAM axis, which increases blood pressure and heart rate through the release of epinephrine and norepinephrine (218,219). Whereas the HPA response to stress is linked with negative affect, distress, withdrawal, and being overwhelmed, the SAM response to stress is associated with increased engagement, cognitive effort, attentional focus, work, and arousal (218,219). This axis also triggers the secretion of salivary alpha-amylase, a carbohydrate digestion enzyme that has been recently used as a salivary stress biomarker (210).

Early studies addressed the association between individual differences in the psychobiology of the stress response and the consequences of these differences on early child development (220–224). More recent studies have extended these research questions to investigate the impact of rural poverty (225), intimate partner violence (226,227), extreme neglect (223,228), divorce (229), nutrition (230,231), and maternal substance use (232) on developmental outcomes in early childhood. Correlational studies have indicated that poverty is associated with increased child cortisol, and a 2009 quasi-experimental study found that children from families who participated in a cash transfer program had lower cortisol compared to children from families who did not participate (225,233). A 2003 study in Nepal as well as a 1998 study in Jamaica found that stress reactivity (salivary cortisol and heart rate) was associated with growth impairment (230,231).

Throughout decades of research on stress and development, several themes have emerged. First, the biobehavioral manifestations of chronic stress are heterogenous (234). Second, differences in biological responses to environmental exposures (i.e. reactivity) are largely responsible for translating experiences into differential outcomes (233,235–237). Third, the social context of the family and quality of family care moderate the effects of exposures on development (212,238–240). Further evaluation of these associations in the context of low- and middle-income countries, where there is a large burden of both early-life chronic stress and poor developmental outcomes, as well as exposure to inflammation and infection, may provide important insights on a high-risk population (28).

This study aims to evaluate the associations between markers of HPA axis activity, SAM axis activity, oxidative stress, and child development outcomes in a cohort of young children in rural Bangladesh. We hypothesized that decreased oxidative status and decreased salivary alpha-amylase would be associated with higher child development scores, while higher salivary cortisol, higher glucocorticoid receptor methylation, and higher heart rate and blood pressure would be associated with higher child development scores.

3.3 Methods

These analyses utilize data from the WASH Benefits study, described in detail previously (40,241). The trial enrolled pregnant mothers in Bangladesh in their first or second trimester of pregnancy in rural subdistricts in Gazipur, Mymensingh, Tangail and Kishoreganj and followed the cohort of children from birth until 2.5 years of age. Here, we describe observational analyses of the associations between child stress biomarkers and concurrent and subsequent child development in a subsample of children from the trial. This sample included 684 children aged 14 months (median age, Year 1) and 1,449 children aged 28 months (median age, Year 2).

Correlates of Stress Biomarkers

Child stress biomarkers included markers of the HPA axis, the SAM axis, and oxidative status. HPA axis biomarkers were salivary assessments of cortisol and glucocorticoid receptor (*NR3C1*) methylation. SAM axis measures were salivary alpha-amylase, resting heart rate, and mean arterial pressure. We measured oxidative status via urinary F2-isoprostanes at Year 1 (median age 14 months), while all other stress biomarkers were assessed at Year 2 (median age 28 months; Table 1). Resting heart rate and mean arterial pressure were measured in four study arms, while other stress biomarker exposures were only assessed in two arms, which led to a greater sample size at Year 2 compared to Year 1.

We analyzed four urinary isoprostane isomers separately (iPF(2a)III, 2,3-dinor-iPF(2a)III, iPF(2a)-IV, and 8,12-iso-iPF(2a)-VI), and we used the first component of a principal components analysis of the four measures of urinary F2-isoprostanes (as these measures were correlated; P-value <0.2) to assess overall oxidative status (242). We assessed stress reactivity at Year 2 as the change in salivary cortisol and salivary alpha-amylase following venipuncture (see Supplemental Material 1 for additional details). In this setting, venipuncture serves as both a physical and psychological stressor, as it involves physical discomfort as well as separation from the mother. We collected saliva pre-stressor, 5 minutes post-stressor, and 20 minutes post-stressor. We measured cortisol pre- and 20 minutes post-stressor, and measured alpha-amylase pre- and 5 minutes post-stressor (243,244). We calculated cortisol and alpha-amylase reactivity as the post-stressor value minus the pre-stressor value, divided by the time elapsed between samples. We recorded time of salivary biomarker assessment to account for circadian patterns of hormone production. We assessed percent methylation across the entire glucocorticoid receptor (*NR3C1*) exon 1F promoter (39 assayed CpG sites) as well as the nerve-growth factor inducing protein A (NGFI-A) transcription factor binding site, which is a specific site within the *NR3C1* exon that is associated with hippocampal glucocorticoid receptor expression (245,246). We measured resting heart rate and blood pressure in triplicate to ensure reliability at Year 2, where we included the median of the three measurements, and we assessed the mean arterial pressure as two times the diastolic blood pressure, plus the systolic blood pressure, divided by three (247–249). We log-transformed F2-isoprostane, cortisol, salivary alpha-amylase, and glucocorticoid receptor methylation distributions to account for skewness. We assessed child stimulation in the home through family care indicator (FCI) score, which is based on Home Observations for Measurement of the Environment (250). The FCI includes subscales for play activities, variety of play materials, sources of play materials, books, and magazines and newspapers, and each of these subscales have demonstrated reliability for children in Bangladesh (250). Additional details of laboratory methods can be found in Appendix 1.

Assessments of Child Development

Primary outcomes included child development data measured via the MacArthur-Bates Communicative Development Inventories (CDI) at Years 1 and 2, the WHO gross motor milestones module at Year 1, and the Extended Ages and Stages Questionnaire (EASQ) at Year 2. Details of child development measures have previously been published and can be found in Appendix 2 (241).

The CDI assessment provides scores for language expression and comprehension. WHO motor milestones include six developmental markers – sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone. Motor milestone attainment was analyzed as a sum score of the 2nd, 4th, 5th, and 6th milestones (1st and 3rd milestones excluded from sum score due to missingness) to assess risk as well as through a time-to-event analysis to assess hazard, which is consistent with previous analyses of this measure of development (241). The EASQ has five domains, and only three were used in this study due to field-work constraints: child communication, gross motor development, and personal-social development. We also generated a combined EASQ score (241). We age-standardized both CDI and EASQ scores using the control group as the standard population in 2-month age bins using standard techniques (241).

Analysis

We used R (version 4.1.1) to conduct observational analyses nested within a randomized controlled trial in accordance with a pre-registered analysis plan (https://osf.io/hzb6m/) (251). We evaluated the association between each exposure of interest (e.g., stress reactivity at Year 2) and each outcome of interest (e.g., CDI comprehension score at Year 2) independently (Table 2), as each association potentially required its own, unique set of adjustment covariates to reduce confounding.

We used natural smoothing splines to accommodate potential nonlinearity and summarized mean developmental outcomes across stress biomarker distributions after controlling for potential confounders and covariates of interest in accordance with our pre-registered statistical analysis plan (252–254). All adjusted analyses included child age and sex, and we screened the following covariates for potential inclusion: birth order, maternal age and education, food insecurity, household crowding, access to drinking water, household assets, prior growth, treatment arm, month of assessment, assessment time, and maternal depression, stress, and exposure to intimate partner violence. Additional information regarding covariate screening and inclusion can be found in Appendix 3. We then plotted these general additive model curves along with simultaneous confidence intervals (255). The primary contrast was the difference in the mean outcome at the 75th and 25th percentile of each exposure level after adjusting for relevant covariates, which we describe as "adjusted difference" hereafter (253). We assessed potential modification of the association between stress exposure and development outcome by FCI score at Year 1 for outcomes assessed at Years 1 and 2 and Year 2 for outcomes assessed at Year 2.

As these observational analyses were exploratory in nature, interpretations included both the strength of associations between individual biomarkers as well as the consistency of the direction of these associations across related biomarker groups. While typical corrections for false discovery rate aim to determine the probability of an individual result being due to random variation, adjusting for the number of repeated tests, we aimed to evaluate whether multiple measures of a similar exposure-outcome domain (e.g., salivary cortisol and child development) indicated a underlying association. For example, if we found that a domain of exposure-outcome associations (e.g. oxidative status and subsequent development) was clustered around (both above and below) the null, but individual measures indicated statistically significant associations (e.g. iPF(2a)-VI at Year 1 and EASQ personal-social score at Year 2), we concluded that these individual results may be spurious associations due to repeated testing. On the other hand, if we found a consistent direction of associations (e.g., point estimates consistently indicating positive correlations) we concluded that these observed estimates might reflect a true association between the domain of exposures and outcomes. In addition, we corrected for repeated testing to evaluate the robustness of individual associations using the Benjamini-Hochberg procedure (256,257).

Ethics

Primary caregivers of all children provided written informed consent prior to enrollment. Human subjects protection committees at International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), the University of California, Berkeley, and Stanford University approved the study protocols. Investigators registered the parent trial at ClinicalTrials.gov (NCT01590095) and a safety monitoring committee convened by icddr,b oversaw the study.

3.4 Results

We analyzed data from 684 children at Year 1 (median age 14 months) and 1,449 children at Year 2 (median age 28 months) for this study (Figure 1). The children had a median cortisol

reactivity of 0 (IQR 0, .01) ug/dl/min and a median salivary alpha-amylase reactivity of 1.41 (IQR -0.08, 4.41) U/ml/min (Table 1). The children had a median length-for-age z-score (LAZ) of -1.42 and a diarrhea prevalence of 13% at Year 1 and a median LAZ of -1.56 and a diarrhea prevalence of 7% at Year 2. Among women in this sample, there was a median educational attainment of 7 years and a 57% prevalence of having experienced intimate partner violence.

We observed 135 contrasts (excluding subgroup analyses) of the differences in child development outcomes at the 75th and 25th percentiles of stress exposures across three domains of stress (HPA axis, SAM axis, and oxidative status). We found that markers of HPA axis activity (cortisol and glucocorticoid receptor methylation) were associated with child development in five out of 30 contrasts (17%), markers of SAM activity (salivary alpha-amylase, heart rate, and blood pressure) were associated with child development in 1 out of 30 contrasts (3%), and markers of oxidative status (F2 isoprostanes) were associated with child development in 3 out of 75 contrasts (4%). The proportion of significant results for HPA axis biomarkers was greater than we could expect due to random variation alone (5%; $\alpha = 0.05$), but the proportion of significant results for SAM axis and oxidative status biomarkers were less than we would expect due to random variation, leading us to believe that significant associations with SAM axis and oxidative status biomarkers may be spurious due to repeated testing. No observed individual associations were statistically significant following false discovery rate correction for multiple testing.

Salivary stress biomarkers

We found that increased salivary cortisol production was associated with worse child development outcomes. These analyses indicated that increased concurrent cortisol reactivity was associated with a lower CDI comprehension score (adjusted difference -0.15 standard deviations (SD), 95% CI (-0.29, -0.01)) as well as a near-significant inverse correlation between cortisol reactivity and CDI expression score (adjusted difference -0.23 SD, 95% CI (-0.45, 0)) at Year 2 (Table 3). In addition, we found that higher post-stressor cortisol was associated with lower combined EASQ score (adjusted difference -0.22 SD, 95% CI (-0.41, -0.04)) and lower CDI comprehension Z-score (adjusted difference -0.12 SD, 95% CI (-0.24, -0.01)).

We observed that greater pre-stressor salivary alpha-amylase was associated with worse child development outcomes. There was a significant association between pre-stressor salivary alpha-amylase and EASQ gross motor score (adjusted difference -0.18 SD, 95% CI (-0.35, 0.00)), and the direction of this association was consistent across communication, personal social, and combined EASQ scores. We did not detect significant associations between post-stressor salivary alpha-amylase or salivary alpha-amylase reactivity and measures of development (Figure 2).

Glucocorticoid receptor methylation

We observed that higher glucocorticoid receptor methylation was associated with worse child development outcomes. Greater mean overall glucocorticoid receptor methylation was correlated with lower concurrent CDI expressive language score (adjusted difference -0.09 SD, 95% CI (-0.17, -0.01)), and there was a consistently negative association between overall glucocorticoid receptor methylation and concurrent measures of child development (Table 4). A greater percent methylation of transcriptor NGFI-A binding site was associated with higher EASQ gross motor score (adjusted difference 0.18 SD, 95% CI (0, 0.37)), but this association was not consistent across other measures of child development (Figure 3).

Oxidative status

We found evidence of an association between concurrent oxidative status and child development at Year 1 (Table 5, Figure 4). Increased concurrent 2,3-dinor-iPF(2a)-III (ng/mg creatinine) was associated with greater WHO sum score (adjusted difference 0.27, 95% CI (0.04, 0.51)) as well as 8,12-iso-iPF(2a)-VI (ng/mg creatinine) and greater CDI comprehension Z-score (adjusted difference 0.15, 95% CI (0.04 SD, 0.27)). We assessed the possibility of a curvilinear association between concurrent oxidative status and child development by plotting the spline curves of these associations (Figure 5). The association between concurrent oxidative stress and CDI comprehension score largely indicated a positive correlation, while correlations of concurrent oxidative status and CDI expression and WHO sum score often depicted nonlinear associations, in which the second and third quartiles (moderate levels) of F2-isoprostanes were associated with better development outcomes relative to the first quartile.

We did not find evidence of a consistent association between measures of oxidative status at Year 1 and subsequent child development at Year 2 (Table 6). Higher levels of iPF(2a)-VI (ng/mg creatinine) were associated with lower EASQ personal-social score (adjusted difference -0.14 , 95% CI (-0.25, -0.03)), but this inverse correlation was not consistent across other urinary F2-isoprostanes or measures of child development (Figure 6). As we would expect to find consistent associations across F2-isoprostane isomers, this observed correlation may be spurious. We did not detect a significant association between any measure of oxidative status and time to attainment of any WHO motor milestone (Table 7) or an association with mean arterial pressure or mean resting heart rate and any measures of concurrent development (Table 8; Figure 7).

Quality of family care

We analyzed potential modification of the association between stress exposure and development outcome by family care indicators (FCI) score at Year 1 and Year 2. The contrasted FCI scores (75th and 25th percentiles) were 9 and 5 at Year 1, and 11 and 6 at Year 2. Although we found some evidence of effect measure modification in specific exposure-outcome associations at specific timepoints, the lack of consistency of these associations over time (e.g., FCI at Year 1 and Year 2) and across related exposure-outcome domains led us to conclude that these associations may be spurious (Supplemental Tables 1-5).

3.5 Discussion

Our observations reveal some consistent associations between multi-level and -system biological signatures of exposure to chronic stress and early child development (Supplemental Materials 1-5). Results suggest that individual differences in the biology of the stress response play an important role in the translation of experience and exposure into developmental consequences in early life. Yet, the nature of these associations is not the same for all children. That is, in this high risk (low-income, rural) developmental context in Bangladesh, using a large sample size and analytical approach that selectively compared those in the lower to upper quartile of biomarker distributions, the magnitude of the effect was small to modest. For some children, chronic stress exposure had measurable negative cognitive and social developmental consequences early in life, but most children appeared to be resilient.

These findings support the use of HPA axis biomarkers, namely cortisol reactivity and overall glucocorticoid receptor methylation, as biomarkers for young children who are at risk of poor developmental outcomes, although the sensitivity of these measures is limited. While pre-

stressor salivary alpha-amylase as well as moderate oxidative status showed some evidence of associations with developmental status, associations were most consistent and strongest for markers of HPA axis activity. These findings suggest that higher cortisol reactivity is associated with worse concurrent child development. This is consistent with previous studies' findings that HPA axis hyperactivity may be related to delays in learning, memory, and neurological development (30,32,33,35,36,212). The age of the sample population has considerable implications. In older children, low or blunted (little change throughout the day) cortisol production is associated with poor developmental outcomes, and investigators have hypothesized HPA axis hypoactivity be the result of early-life HPA axis hyperactivity (258). This hypothesis is partially motivated by the social buffering theory, which posits that children with secure attachment associations may show little HPA axis reactivity in early childhood as their attachment figure serves as a buffer from these stressors (259). Our findings provide indirect support for this hypothesis by indicating a link between early childhood HPA axis hyperactivity and poor developmental status. Follow-up evaluation of HPA axis activity in this cohort once they have reached school age may provide insights on the developmental origins of HPA axis activity.

We found that greater glucocorticoid receptor methylation was associated with worse child development outcomes. This indicates that glucocorticoid receptor hypermethylation, which is an indicator of early-life stress, may be associated with poor developmental outcomes for high-risk children in rural Bangladesh (213). This is consistent with previous findings that glucocorticoid receptor methylation is positively associated with externalizing behavior and depressive symptoms in school-aged children (260). In a previous analysis of this sample evaluating the impact of randomized assignment to interventions on child stress, we found that the control group had greater glucocorticoid receptor methylation relative to the combined N+WSH intervention group (261). These cumulative findings indicate that glucocorticoid receptor methylation may be a pathway or marker of environmental stressors' contribution to developmental status.

Although we detected a positive correlation between pre-stressor salivary alpha-amylase and child development, the lack of consistency of relationships (in terms of direction and significance) across measures of SAM axis activity and child development (3% significant contrasts) prevents us from concluding that this relationship is not spurious. Similarly, although we found limited evidence that moderate oxidative status was associated with child development, the inconsistency of these correlations across measures of oxidative status and child development (4% significant contrasts) prevents us from concluding that these correlations were not due to random variation.

Strengths and limitations

Our study evaluated the association between stress and child development using a comprehensive set of biomarkers representing the HPA axis, SAM system, and oxidative stress. As each of these biomarkers reflects a unique stress response, analysis of these individual correlations between each stress biomarker and measure of child development highlights these associations at multiple levels and multiple biological systems.

The observational and often concurrent nature of these analyses does not readily enable causal inference regarding the impact of child stress on child development. As it would not be ethical or feasible to randomize children to experience varying levels of stress, we conducted observational

analyses with multivariate adjustment to control for potential confounders and covariates of interest, although residual confounding may still be present. Future analyses should include a greater number of time points for observations with additional temporal separation. While investigators assessed some measures prior to assessing development outcomes, the majority of stress and development measures were assessed concurrently at Year 2. Our interpretations are based on the assumption that stress biomarker exposures may cause a change in child development, as the stress response is a shorter-term outcome than child development, although it is also possible that child development outcomes lead to changes in child stress neurobiology (i.e., reverse causation). In addition, assessment of child development measures for children greater than 3 years of age, such as school attendance, executive functioning, and intelligence, would enable inference of the impact of stress on longer-term development.

The inclusion of multiple measures of both stress and development is both a strength of this study and a limitation, as multiple comparisons lead to an increased risk of Type 1 error. We aimed to account for this risk by assessing the consistency of the direction (positive vs. negative) of point estimates in each domain of exposure-outcome assessments, in addition to evaluation of each contrast's statistical significance, although the possibility of Type 1 errors remains plausible. Furthermore, the use of corrections for false discovery, such as Benjamini-Hochberg, may be overly conservative (i.e. low power) for correlational studies of stress biomarkers and child development, where we would expect to see small to modest effect sizes. Therefore, we recommend that these inferences inform futures studies that can deliberately target and evaluate potential associations of interest.

3.6 Conclusions

Our findings support the use of HPA axis biomarkers (cortisol regulation and glucocorticoid receptor methylation) as markers of young children in Bangladesh who are at risk of poor developmental outcomes. These associations contribute to the body of evidence that supports interventions that aim to improve child development by intervening on early-life stress. Given the context of this study in rural Bangladesh, where poverty, pathogen exposure, and malnutrition are relatively common, the possibility of intervening on stress may provide a low-cost intervention to bolster child development.

3.7 Tables and Figures

Table 1. Descriptive statistics of sample population

			n (%) or median (IQR)
Child		Female	761 (49%)
	Urinary F2-isoprostanes (ng/mg creatinine; Year 1)	iPF(2a)-III	-0.42 (-0.72, -0.09)
		2,3-dinor-iPF(2a)-III	1.76 (1.55, 1.97)
		iPF(2a-VI	2.57 (2.31, 2.87)
		8,12-iso-iPF(2a)-VI	2.58 (2.17, 2.91)
	Salivary cortisol reactivity (ug/dl; Year 2)	Cortisol reactivity	0 (0, 0.01)
		Cortisol residualized gain score	-0.09 (-0.21, 0.14)
	Salivary alpha-amylase reactivity (U/ml; Year 2)	Salivary alpha-amylase reactivity	1.41 (-0.08, 4.41)
		Salivary alpha-amylase residualized gain score	-25.83 (-51.4, 28.68)
	Sympathetic-adreno-medullar biomarkers (Year 2)	Mean arterial pressure (mmHg)	64.44 (60.78, 68.78)
		Resting heart rate (bpm)	109 (99.33, 118.67)
	Glucocorticoid receptor percent methylation	NR3C1 exon 1F promoter	-5.66 (-6.06, -5.32)
		NGFI-A transcription factor binding site	-4.48 (-4.65, -4.14)
	Child development (Year 1)	WHO gross motor milestone sum score	2 (1, 4)
		CDI expressive language z-score	0.02 (-0.54, 0.74)
		CDI language understanding z-score	0.09 (-0.56, 0.78)
	Child development (Year 2)	EASQ communication z-score	0.37 (-0.39, 0.75)
		EASQ motor development z-score	-0.15 (-0.59, 0.87)
		EASQ personal-social development z-score	0.14 (-0.49, 1)
		EASQ combined z-score	0.3 (-0.37, 0.86)
		CDI expressive language z-score	0.27 (-0.57, 0.8)
		CDI language understanding z-score	0.12 (-0.48, 0.74)
	Anthropometry (14 months, Year 1)	Length-for-age z-score	-1.42 (-2.07, -0.76)

			n (%) or median (IQR)
		Weight-for-age z-score	-1.31 (-2.01, -0.64)
		Weight-for-length z-score	-0.9 (-1.56, -0.23)
		Head circumference-for-age z-score	-1.79 (-2.35, -1.13)
	Anthropometry (28 months, Year 2)	Length-for-age z-score	-1.56 (-2.28, -0.95)
		Weight-for-age z-score	-1.58 (-2.2, -0.94)
		Weight-for-length z-score	-1.03 (-1.62, -0.38)
		Head circumference-for-age z-score	-1.81 (-2.38, -1.21)
	Diarrhea (14 months, Year 1)	Caregiver-reported 7-day recall	195 (13%)
	Diarrhea (28 months, Year 2)	Caregiver-reported 7-day recall	110 (7%)
Mother		Age (years)	24 (20, 27)
	Anthropometry at enrollment	Height (cm)	150.2 (146.8, 154.05)
	Education	Schooling completed (years)	7 (4, 9)
	Depression at Year 1	CES-D score	10 (6, 16)
	Depression at Year 2	CES-D score	10 (5, 17)
	Perceived stress at Year 2	Perceived Stress Scale score	14 (10, 17.25)
	Intimate partner violence	Any lifetime exposure	810 (57%)

CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Table 2. Study hypotheses, exposures, and outcomes.

Hypothesis	*Exposures*	*Outcomes*
Salivary stress biomarkers are associated with child development	Cortisol (hypothesis: positively correlated) and salivary alpha-amylase (hypothesis: negatively correlated) concentrations pre-stressor, post-stressor, and reactivity at Year 2	CDI and EASQ at Year 2
Glucocorticoid receptor methylation is inversely associated with child development	Percentage methylation at NGFI-A transcription factor binding site and mean overall glucocorticoid receptor methylation at Year 2	CDI and EASQ at Year 2
Oxidative status is inversely associated with child development	Individual urinary F2 isoprostane isomers and combined score at Year 1	WHO motor milestones and CDI at Year 1; CDI and EASQ at Year 2
Blood pressure and heart rate are negatively associated with child development	Mean arterial pressure and resting heart rate at Year 2	CDI and EASQ at Year 2

CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Table 3. Salivary stress biomarkers and child development at Year 2

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
Cortisol reactivity (ug/dl/min)	EASQ communication score	545	0	0.01	0.23	0.25	0.02 (-0.07, 0.11)	0.67
	EASQ gross motor score	538	0	0.01	-0.09	-0.11	-0.01 (-0.13, 0.1)	0.81
	EASQ personal social score	544	0	0.01	0.57	0.43	-0.14 (-0.4, 0.12)	0.3
	Combined EASQ score	543	0	0.01	0.13	0.13	-0.01 (-0.1, 0.09)	0.9
	CDI expressive language score	552	0	0.01	-0.01	-0.24	-0.23 (-0.45, 0)	0.05
	CDI comprehension score	547	0	0.01	-0.01	-0.16	-0.15 (-0.29, -0.01)	0.04
Pre-stressor cortisol (ug/dl)	EASQ communication score	588	-2.51	-1.66	0.32	0.24	-0.07 (-0.32, 0.17)	0.57
	EASQ gross motor score	582	-2.51	-1.66	0.24	0.17	-0.07 (-0.21, 0.07)	0.31
	EASQ personal social score	588	-2.51	-1.66	0.57	0.47	-0.1 (-0.33, 0.12)	0.38
	Combined EASQ score	587	-2.51	-1.66	0.27	0.09	-0.17 (-0.38, 0.03)	0.1
	CDI expressive language score	595	-2.51	-1.66	-0.06	-0.11	-0.04 (-0.13, 0.05)	0.35
	CDI comprehension score	591	-2.51	-1.66	0.02	-0.03	-0.05 (-0.14, 0.04)	0.28
Post-stressor cortisol (ug/dl)	EASQ communication score	550	-2.08	-0.61	0.34	0.19	-0.16 (-0.32, 0.01)	0.06

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
	EASQ gross motor score	543	-2.06	-0.61	0.12	0.07	-0.05 (-0.17, 0.07)	0.42
	EASQ personal social score	549	-2.08	-0.61	0.46	0.43	-0.04 (-0.16, 0.09)	0.59
	Combined EASQ score	548	-2.07	-0.61	0.29	0.07	-0.22 (-0.41, -0.04)	0.02
	CDI expressive language score	557	-2.06	-0.61	-0.09	-0.16	-0.07 (-0.18, 0.04)	0.24
	CDI comprehension score	552	-2.07	-0.61	0.01	-0.12	-0.12 (-0.24, -0.01)	0.04
Salivary alpha-amylase reactivity (U/ml/min)	EASQ communication score	561	-0.08	4.61	0.11	0.13	0.02 (-0.04, 0.07)	0.57
	EASQ gross motor score	555	-0.08	4.54	-0.08	-0.05	0.04 (-0.03, 0.1)	0.26
	EASQ personal social score	560	-0.08	4.61	0.32	0.37	0.06 (-0.18, 0.3)	0.65
	Combined EASQ score	559	-0.08	4.6	0.03	0.18	0.15 (-0.06, 0.36)	0.16
	CDI expressive language score	568	-0.08	4.52	-0.2	-0.18	0.02 (-0.03, 0.08)	0.45
	CDI comprehension score	564	-0.08	4.49	-0.05	-0.02	0.03 (-0.03, 0.09)	0.29
Pre-stressor salivary alpha-amylase (U/ml)	EASQ communication score	584	3.5	4.64	0.25	0.17	-0.09 (-0.18, 0.01)	0.07
	EASQ gross motor score	578	3.5	4.65	0.04	-0.14	-0.18 (-0.35, 0)	0.04
	EASQ personal social score	584	3.5	4.65	0.45	0.43	-0.02 (-0.23, 0.19)	0.84

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
	Combined EASQ score	583	3.5	4.65	0.23	0.04	-0.19 (-0.4, 0.02)	0.07
	CDI expressive language score	591	3.51	4.64	-0.1	-0.16	-0.07 (-0.22, 0.08)	0.39
	CDI comprehension score	587	3.5	4.65	0.05	0.01	-0.04 (-0.15, 0.06)	0.42
Post-stressor salivary alpha-amylase (U/ml)	EASQ communication score	571	3.94	5.17	0.22	0.22	0 (-0.1, 0.09)	0.98
	EASQ gross motor score	565	3.94	5.16	0.21	0.19	-0.02 (-0.12, 0.09)	0.78
	EASQ personal social score	570	3.94	5.17	0.45	0.47	0.02 (-0.08, 0.13)	0.66
	Combined EASQ score	569	3.94	5.17	0.18	0.19	0.01 (-0.09, 0.11)	0.9
	CDI expressive language score	578	3.94	5.16	-0.19	-0.16	0.03 (-0.06, 0.13)	0.53
	CDI comprehension score	574	3.94	5.16	-0.08	-0.04	0.04 (-0.05, 0.14)	0.37

Analyses adjusted for child age and child sex, and screened the following covariates for potential inclusion (see Appendix 3 for details) -- child birth order, maternal age, maternal height, maternal education, household food insecurity, number of children in the household, number of individuals living in the compound, distance to primary drinking water source, household assets, prior anthropometry, month of assessment, treatment arm, pre-stressor sample collection time, maternal Center for Epidemiologic Studies Depression Scale score, maternal Perceived Stress Scale score, and maternal lifetime cumulative exposure to intimate partner violence.

* P-value < 0.2 after adjusting for multiple comparisons using the Benjamini-Hochberg procedure

CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Table 4. Urinary isoprostanes and child development at Year 1

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
Mean overall percentage glucocorticoid receptor methylation	EASQ communication score	557	-6.05	-5.31	0.25	0.22	-0.02 (-0.1, 0.06)	0.58
	EASQ gross motor score	551	-6.05	-5.3	-0.06	-0.05	0.01 (-0.08, 0.1)	0.82
	EASQ personal social score	557	-6.05	-5.31	0.44	0.36	-0.08 (-0.17, 0.01)	0.07
	Combined EASQ score	556	-6.05	-5.3	0.12	0.13	0.01 (-0.21, 0.23)	0.95
	CDI expressive language score	563	-6.05	-5.3	0.18	0.09	-0.09 (-0.17, -0.01)	0.03
	CDI comprehension score	559	-6.05	-5.3	0.1	0.01	-0.09 (-0.19, 0.01)	0.07
Percentage methylation at NGFI-A transcription factor binding site (CpG site #12)	EASQ communication score	339	-4.65	-4.16	0.04	0.02	-0.02 (-0.13, 0.09)	0.71
	EASQ gross motor score	333	-4.65	-4.16	0.02	0.2	0.18 (0, 0.37)	0.05
	EASQ personal social score	339	-4.65	-4.16	0.35	0.36	0.01 (-0.11, 0.13)	0.9
	Combined EASQ score	338	-4.65	-4.16	0	0.03	0.03 (-0.08, 0.14)	0.62
	CDI expressive language score	342	-4.66	-4.16	0.32	0.29	-0.03 (-0.14, 0.08)	0.63
	CDI comprehension score	339	-4.66	-4.16	-0.13	-0.2	-0.07 (-0.18, 0.05)	0.26

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
						Adjusted		
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value

Analyses adjusted for child age and child sex, and screened the following covariates for potential inclusion (see Appendix 3 for details) -- child birth order, maternal age, maternal height, maternal education, household food insecurity, number of children in the household, number of individuals living in the compound, distance to primary drinking water source, household assets, prior anthropometry, month of assessment, treatment arm, pre-stressor sample collection time, maternal Center for Epidemiologic Studies Depression Scale score, maternal Perceived Stress Scale score, and maternal lifetime cumulative exposure to intimate partner violence.

* P-value < 0.2 after adjusting for multiple comparisons using the Benjamini-Hochberg procedure

CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Table 5. Urinary isoprostanes and child development at Year 1

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
IPF(2a)-III (ng/mg creatinine)	Sum of 2nd, 4th, 5th, and 6th WHO motor milestones	571	-0.73	-0.09	1.84	1.79	-0.05 (-0.37, 0.27)	0.77
	CDI expressive language Z-score	672	-0.72	-0.09	0.26	0.31	0.05 (-0.1, 0.2)	0.52
	CDI comprehension Z-score	576	-0.73	-0.09	0.37	0.43	0.06 (-0.04, 0.16)	0.24
2,3-dinor-iPF(2a)-III (ng/mg creatinine)	Sum of 2nd, 4th, 5th, and 6th WHO motor milestones	571	1.54	1.97	1.81	2.09	0.27 (0.04, 0.51)	0.02
	CDI expressive language Z-score	672	1.55	1.97	0.25	0.27	0.01 (-0.09, 0.11)	0.8
	CDI comprehension Z-score	576	1.55	1.97	0.35	0.36	0.01 (-0.11, 0.12)	0.89
iPF(2a)-VI (ng/mg creatinine)	Sum of 2nd, 4th, 5th, and 6th WHO motor milestones	571	2.3	2.87	1.92	1.92	0 (-0.15, 0.15)	0.98
	CDI expressive language Z-score	672	2.31	2.87	0.2	0.22	0.02 (-0.2, 0.24)	0.86
	CDI comprehension Z-score	576	2.3	2.87	0.38	0.43	0.05 (-0.08, 0.18)	0.43
8,12-iso-iPF(2a)-VI (ng/mg creatinine)	Sum of 2nd, 4th, 5th, and 6th WHO motor milestones	571	2.17	2.91	1.9	1.91	0.01 (-0.17, 0.19)	0.94
	CDI expressive language Z-score	672	2.17	2.92	0.19	0.36	0.17 (-0.03, 0.38)	0.09
	CDI comprehension Z-score	576	2.18	2.91	0.29	0.44	0.15 (0.04, 0.27)	0.01

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
Combined urinary oxidative status score	Sum of 2nd, 4th, 5th, and 6th WHO motor milestones	571	2.48	3.54	1.88	1.96	0.09 (-0.11, 0.29)	0.41
	CDI expressive language Z-score	672	2.5	3.55	0.1	0.31	0.21 (-0.03, 0.44)	0.08
	CDI comprehension Z-score	576	2.48	3.54	0.22	0.4	0.18 (-0.08, 0.44)	0.18

Analyses adjusted for child age and child sex, and screened the following covariates for potential inclusion (see Appendix 3 for details) -- child birth order, maternal age, maternal height, maternal education, household food insecurity, number of children in the household, number of individuals living in the compound, distance to primary drinking water source, household assets, prior anthropometry, month of assessment, treatment arm, pre-stressor sample collection time, maternal Center for Epidemiologic Studies Depression Scale score, maternal Perceived Stress Scale score, and maternal lifetime cumulative exposure to intimate partner violence.

* P-value < 0.2 after adjusting for multiple comparisons using the Benjamini-Hochberg procedure

CDI: MacArthur-Bates Communicative Development Inventories

Table 6. Urinary isoprostanes at Year 1 and child development at Year 2

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
IPF(2a)-III (ng/mg creatinine)	EASQ communication score	551	-0.74	-0.09	0.19	0.21	0.01 (-0.07, 0.1)	0.79
	EASQ gross motor score	546	-0.73	-0.09	0.32	0.24	-0.08 (-0.18, 0.01)	0.08
	EASQ personal social score	551	-0.73	-0.09	0.42	0.41	0 (-0.12, 0.11)	0.97
	Combined EASQ score	550	-0.73	-0.09	0.23	0.21	-0.01 (-0.1, 0.08)	0.8
	CDI expressive language score	558	-0.74	-0.09	0	0.05	0.04 (-0.04, 0.13)	0.34
	CDI comprehension score	554	-0.73	-0.09	0.07	0.13	0.06 (-0.03, 0.15)	0.21
2,3-dinor-iPF(2a)-III (ng/mg creatinine)	EASQ communication score	551	1.54	1.97	0.21	0.17	-0.04 (-0.15, 0.07)	0.49
	EASQ gross motor score	546	1.54	1.97	0.37	0.29	-0.08 (-0.34, 0.18)	0.55
	EASQ personal social score	551	1.54	1.97	0.42	0.35	-0.07 (-0.22, 0.08)	0.38
	Combined EASQ score	550	1.54	1.97	0.21	0.12	-0.09 (-0.22, 0.05)	0.21
	CDI expressive language score	558	1.54	1.97	0.01	-0.03	-0.04 (-0.18, 0.1)	0.6
	CDI comprehension score	554	1.54	1.97	0.08	0.05	-0.03 (-0.15, 0.09)	0.64
iPF(2a)-VI (ng/mg creatinine)	EASQ communication score	551	2.3	2.87	0.18	0.1	-0.08 (-0.21, 0.04)	0.19

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
	EASQ gross motor score	546	2.31	2.87	0.33	0.27	-0.07 (-0.17, 0.04)	0.22
	EASQ personal social score	551	2.31	2.87	0.42	0.28	-0.14 (-0.25, -0.03)	0.01
	Combined EASQ score	550	2.3	2.87	0.22	0.12	-0.1 (-0.23, 0.04)	0.15
	CDI expressive language score	558	2.3	2.87	0.01	-0.07	-0.08 (-0.17, 0.02)	0.12
	CDI comprehension score	554	2.31	2.87	0.06	0.05	-0.01 (-0.15, 0.13)	0.91
8,12-iso-iPF(2a)-VI (ng/mg creatinine)	EASQ communication score	551	2.16	2.91	0.2	0.26	0.06 (-0.07, 0.19)	0.36
	EASQ gross motor score	546	2.17	2.91	0.29	0.31	0.02 (-0.09, 0.13)	0.76
	EASQ personal social score	551	2.17	2.91	0.41	0.42	0 (-0.12, 0.12)	0.97
	Combined EASQ score	550	2.17	2.91	0.21	0.25	0.04 (-0.1, 0.17)	0.6
	CDI expressive language score	558	2.17	2.91	0.01	-0.02	-0.03 (-0.22, 0.16)	0.76
	CDI comprehension score	554	2.17	2.91	0.03	0.09	0.05 (-0.13, 0.24)	0.57
Combined urinary oxidative status score	EASQ communication score	551	2.48	3.56	0.19	0.19	-0.01 (-0.11, 0.1)	0.9
	EASQ gross motor score	546	2.49	3.55	0.32	0.27	-0.06 (-0.17, 0.05)	0.31

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
	EASQ personal social score	551	2.48	3.56	0.42	0.38	-0.04 (-0.16, 0.08)	0.49
	Combined EASQ score	550	2.48	3.55	0.23	0.19	-0.04 (-0.15, 0.08)	0.54
	CDI expressive language score	558	2.47	3.55	0	0.02	0.02 (-0.08, 0.12)	0.72
	CDI comprehension score	554	2.49	3.54	0.05	0.11	0.06 (-0.05, 0.17)	0.28

Analyses adjusted for child age and child sex, and screened the following covariates for potential inclusion (see Appendix 3 for details) -- child birth order, maternal age, maternal height, maternal education, household food insecurity, number of children in the household, number of individuals living in the compound, distance to primary drinking water source, household assets, prior anthropometry, month of assessment, treatment arm, pre-stressor sample collection time, maternal Center for Epidemiologic Studies Depression Scale score, maternal Perceived Stress Scale score, and maternal lifetime cumulative exposure to intimate partner violence.

* P-value < 0.2 after adjusting for multiple comparisons using the Benjamini-Hochberg procedure

CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Table 7. Urinary isoprostanes and time to WHO motor milestone at Year 1

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile	
					Adjusted	
					Hazard Ratio (95% CI)	P-value
IPF(2a)-III (ng/mg creatinine)	Time to sitting unsupported	577	-0.73	-0.09	0 (0, 727975445160237568)	0.61
	Time to crawling	682	-0.72	-0.09	1.02 (0.91, 1.14)	0.76
	Time to standing with support	671	-0.72	-0.1	1.01 (0.88, 1.17)	0.86
	Time to walking with support	670	-0.72	-0.1	1.03 (0.9, 1.17)	0.72
	Time to standing unsupported	672	-0.72	-0.09	0.95 (0.65, 1.39)	0.79
	Time to walking unsupported	679	-0.72	-0.09	1.05 (0.88, 1.25)	0.6
2,3-dinor-iPF(2a)-III (ng/mg creatinine)	Time to sitting unsupported	577	1.55	1.97	0 (0, 1.18552134902425e+59)	0.76
	Time to crawling	682	1.55	1.97	1 (0.72, 1.37)	0.98
	Time to standing with support	671	1.55	1.97	0.95 (0.81, 1.11)	0.55
	Time to walking with support	670	1.55	1.97	0.97 (0.82, 1.15)	0.72
	Time to standing unsupported	672	1.55	1.97	1.12 (0.93, 1.33)	0.23
	Time to walking unsupported	679	1.55	1.97	1 (0.81, 1.23)	1
iPF(2a)-VI (ng/mg creatinine)	Time to sitting unsupported	577	2.3	2.87	624959704.4 (0, 8.78286913170087e+65)	0.78
	Time to crawling	682	2.31	2.87	1 (0.89, 1.14)	0.97
	Time to standing with support	671	2.31	2.87	0.91 (0.78, 1.06)	0.22
	Time to walking with support	670	2.31	2.87	1.12 (0.88, 1.42)	0.35

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile	
					Adjusted	
					Hazard Ratio (95% CI)	P-value
8,12-iso-iPF(2a)-VI (ng/mg creatinine)	Time to standing unsupported	672	2.31	2.87	0.8 (0.51, 1.26)	0.35
	Time to walking unsupported	679	2.31	2.87	1.02 (0.76, 1.37)	0.88
	Time to sitting unsupported	577	2.19	2.91	0.03 (0, 2645605214851898)	0.87
	Time to crawling	682	2.17	2.92	0.98 (0.86, 1.12)	0.79
	Time to standing with support	671	2.16	2.91	0.88 (0.74, 1.04)	0.12
	Time to walking with support	670	2.16	2.92	0.99 (0.8, 1.23)	0.95
	Time to standing unsupported	672	2.16	2.92	1.07 (0.79, 1.45)	0.69
	Time to walking unsupported	679	2.17	2.91	1.12 (0.81, 1.53)	0.51
Combined urinary oxidative status score	Time to sitting unsupported	577	2.5	3.55	0 (0, 520427180630980)	0.52
	Time to crawling	682	2.49	3.54	0.99 (0.87, 1.14)	0.93
	Time to standing with support	671	2.48	3.54	1.12 (0.75, 1.69)	0.59
	Time to walking with support	670	2.48	3.54	1 (0.84, 1.2)	0.98
	Time to standing unsupported	672	2.48	3.54	1.2 (0.8, 1.82)	0.38
	Time to walking unsupported	679	2.49	3.54	0.98 (0.79, 1.22)	0.87

Analyses adjusted for child age and child sex, and screened the following covariates for potential inclusion (see Appendix 3 for details) -- child birth order, maternal age, maternal height, maternal education, household food insecurity, number of children in the household, number of individuals living in the compound, distance to primary drinking water source, household assets, prior anthropometry, month of assessment, treatment arm, pre-stressor sample collection time, maternal Center for Epidemiologic Studies Depression Scale score, maternal Perceived Stress Scale score, and maternal lifetime cumulative exposure to intimate partner violence.

*P-value < 0.2 after adjusting for multiple comparisons using the Benjamini-Hochberg procedure

Hazard ratio could not be estimated for sitting without support since nearly all children had achieved this milestone before time of measurement

Table 8. Mean arterial pressure and heart rate and child development at Year 2

Exposure	Outcome	N	25th Percentile	75th Percentile	Outcome, 75th Percentile v. 25th Percentile			
					Adjusted			
					Predicted Outcome at 25th Percentile	Predicted Outcome at 75th Percentile	Coefficient (95% CI)	P-value
Mean arterial pressure (mmHg)	EASQ communication score	1196	60.89	69	-0.18	-0.13	0.04 (-0.04, 0.12)	0.29
	EASQ gross motor score	1176	60.89	69	0.38	0.4	0.02 (-0.04, 0.09)	0.49
	EASQ personal social score	1194	60.89	69	0.04	-0.04	-0.08 (-0.26, 0.09)	0.35
	Combined EASQ score	1191	60.89	69	0.07	0.09	0.02 (-0.14, 0.17)	0.85
	CDI expressive language score	1209	60.89	69	-0.02	0.07	0.09 (-0.08, 0.25)	0.31
	CDI comprehension score	1195	60.94	69	0.02	0.01	-0.01 (-0.16, 0.15)	0.92
Mean resting heart rate (bpm)	EASQ communication score	1198	99.33	118.33	-0.15	-0.15	-0.01 (-0.07, 0.06)	0.89
	EASQ gross motor score	1178	99.33	118.33	0.39	0.37	-0.02 (-0.08, 0.05)	0.58
	EASQ personal social score	1196	99.33	118.33	0.03	-0.07	-0.1 (-0.25, 0.06)	0.23
	Combined EASQ score	1193	99.33	118.33	-0.09	-0.11	-0.02 (-0.1, 0.06)	0.68
	CDI expressive language score	1211	99.33	118.5	0.01	-0.03	-0.04 (-0.11, 0.02)	0.18
	CDI comprehension score	1196	99.33	118.33	-0.01	-0.04	-0.03 (-0.09, 0.03)	0.27

Analyses adjusted for child age and child sex, and screened the following covariates for potential inclusion (see Appendix 3 for details) -- child birth order, maternal age, maternal height, maternal education, household food insecurity, number of children in the household, number of individuals living in the compound, distance to primary drinking water source, household assets, prior anthropometry, month of assessment, treatment arm, pre-stressor sample collection time, maternal Center for Epidemiologic Studies Depression Scale score, maternal Perceived Stress Scale score, and maternal lifetime cumulative exposure to intimate partner violence.

* P-value < 0.2 after adjusting for multiple comparisons using the Benjamini-Hochberg procedure

CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Figure 1. Participant enrollment, follow-up, and analysis

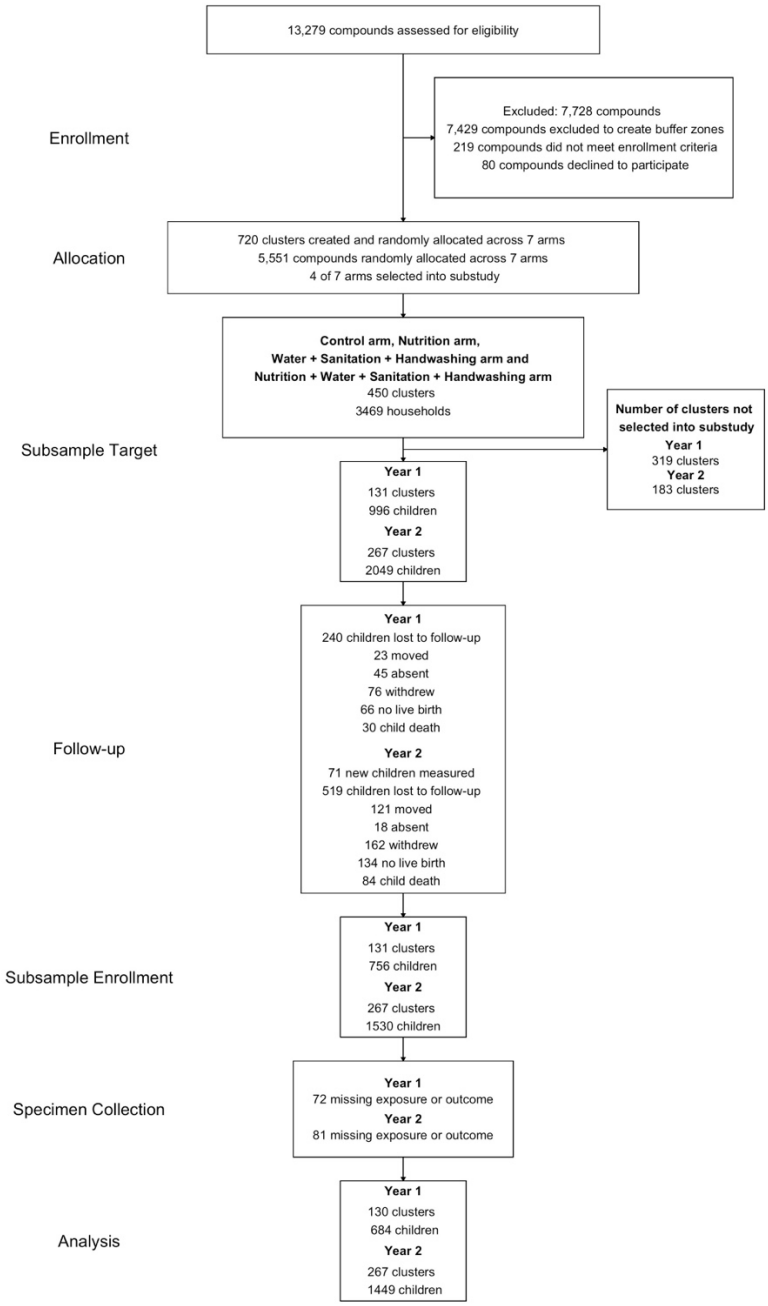
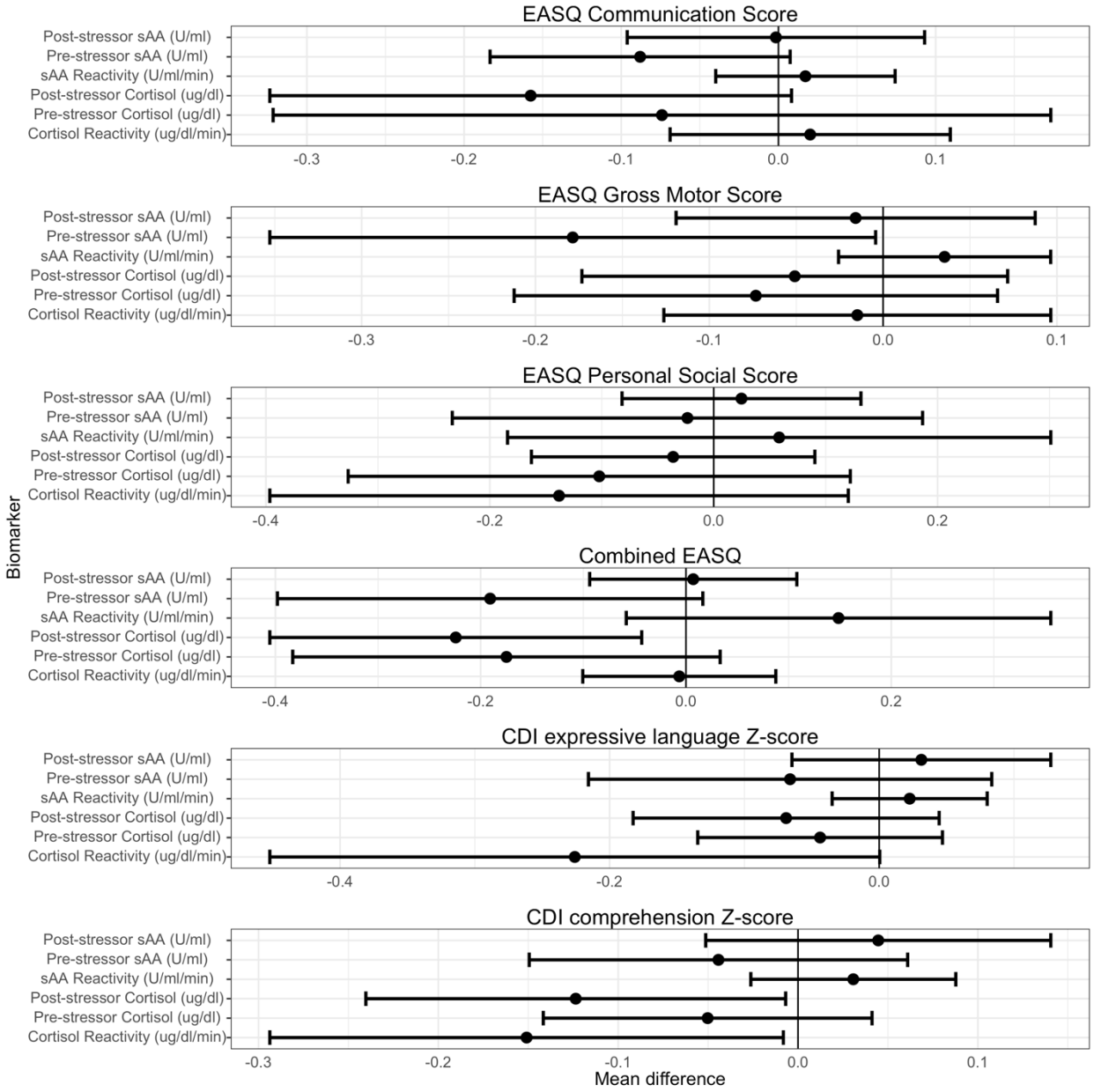
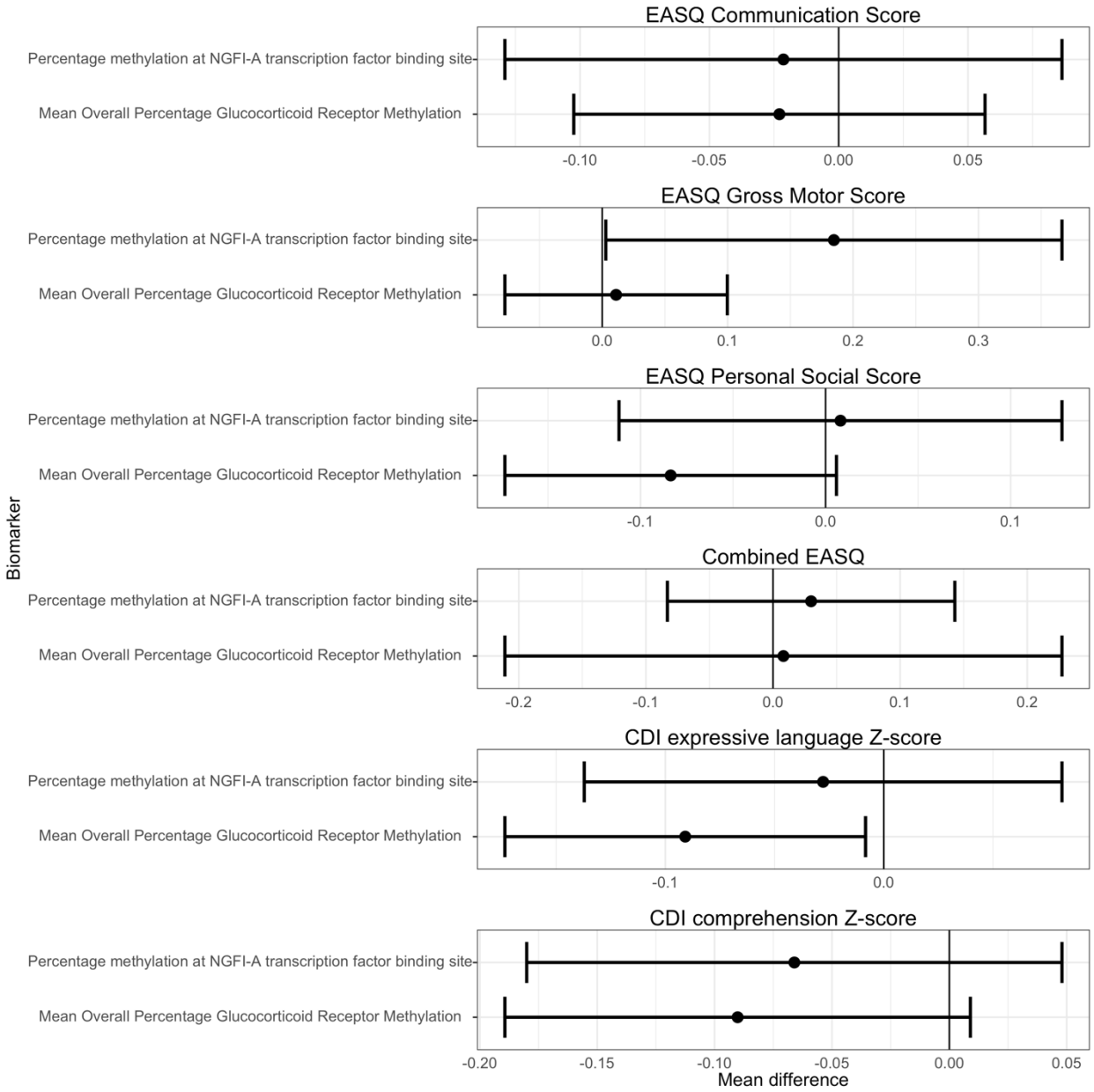


Figure 2. Salivary stress biomarkers and child development at Year 2



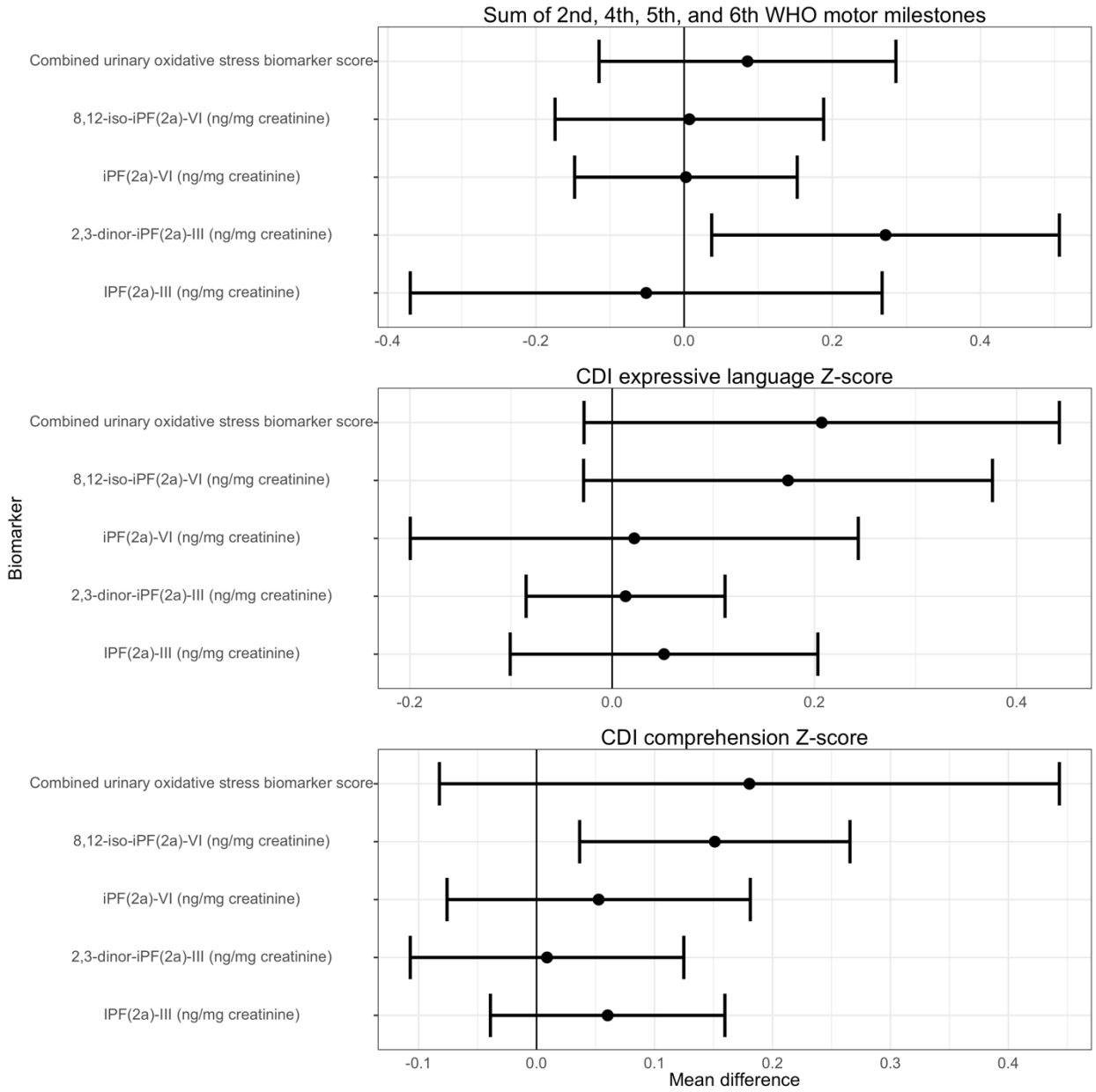
sAA: salivary alpha-amylase; CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Figure 3. Glucocorticoid receptor methylation and child development at Year 2



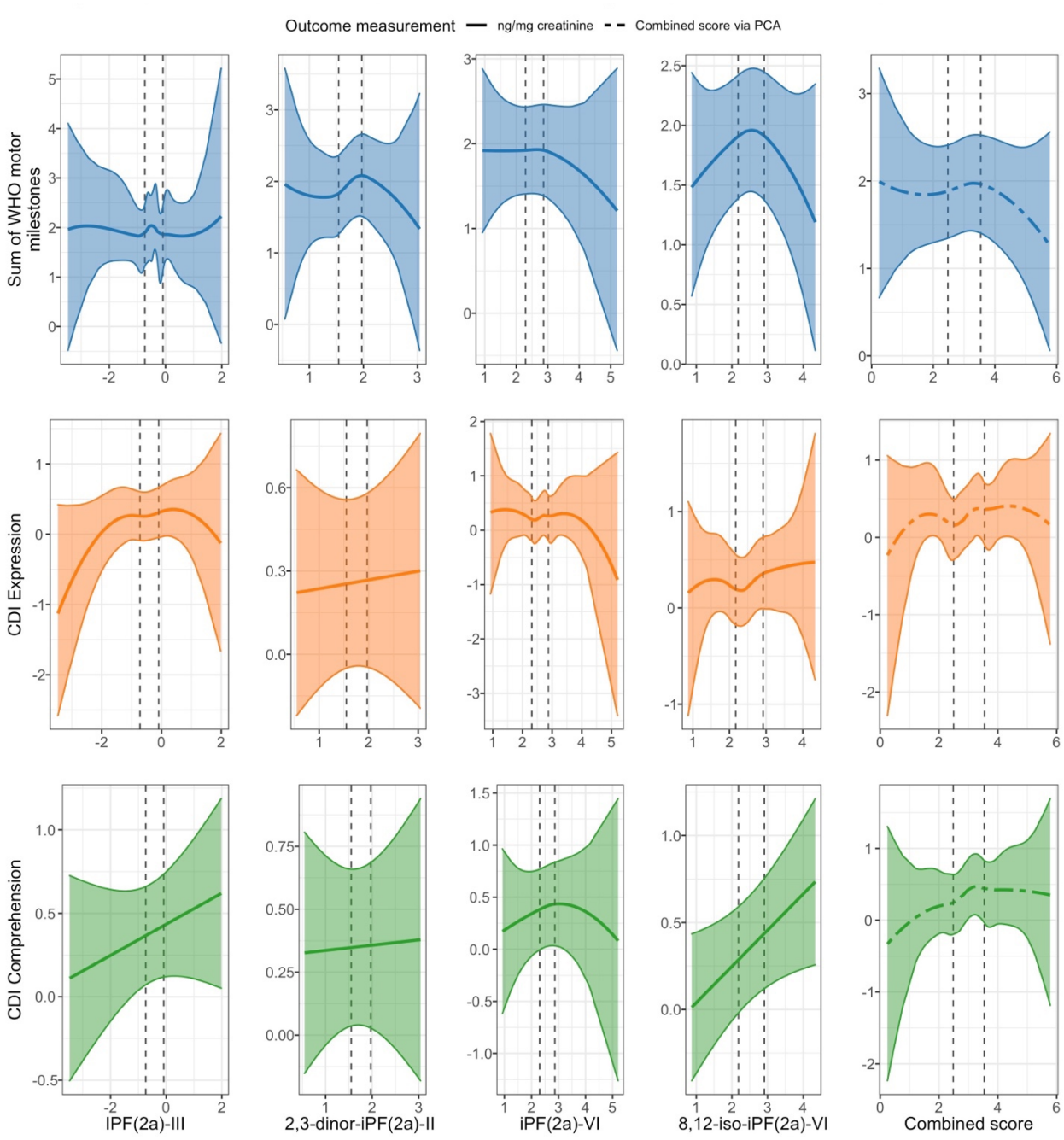
CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Figure 4. Urinary isoprostanes and child development at Year 1



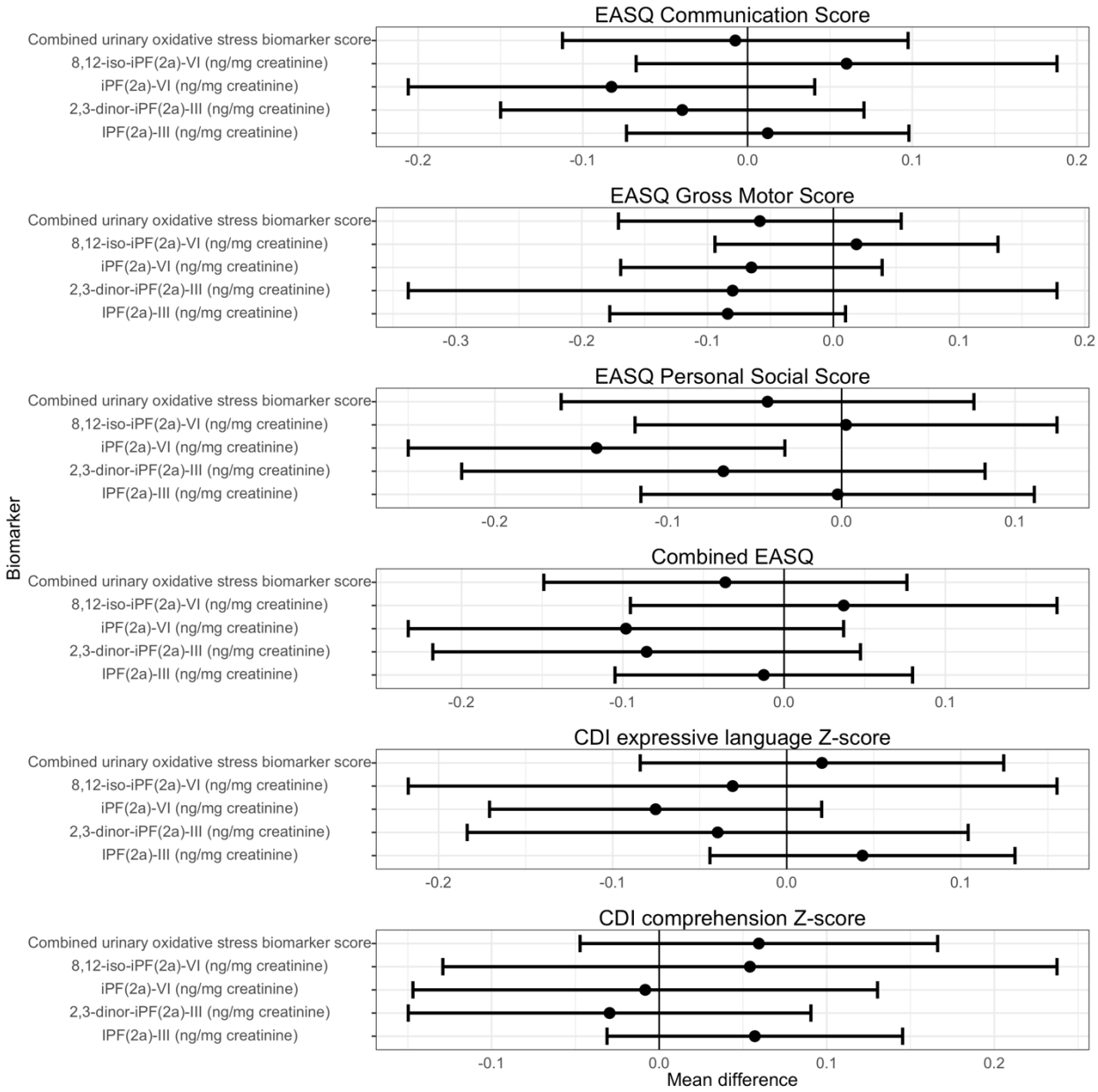
CDI: MacArthur-Bates Communicative Development Inventories

Figure 5. Spline curves of the relationships between concurrent urinary isoprostanes and child development at Year 1



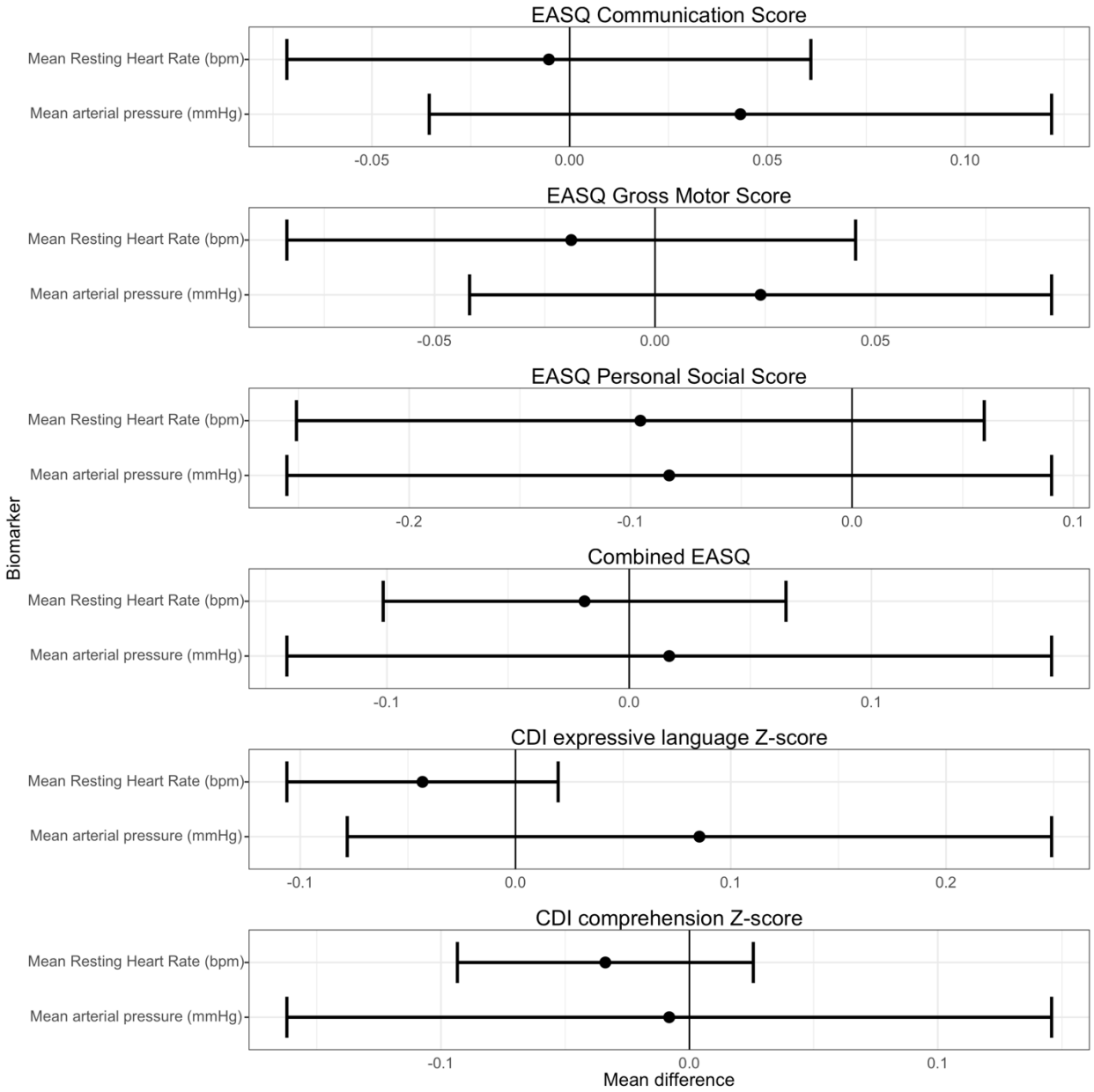
CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Figure 6. Urinary isoprostanes at Year 1 and child development at Year 2



CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

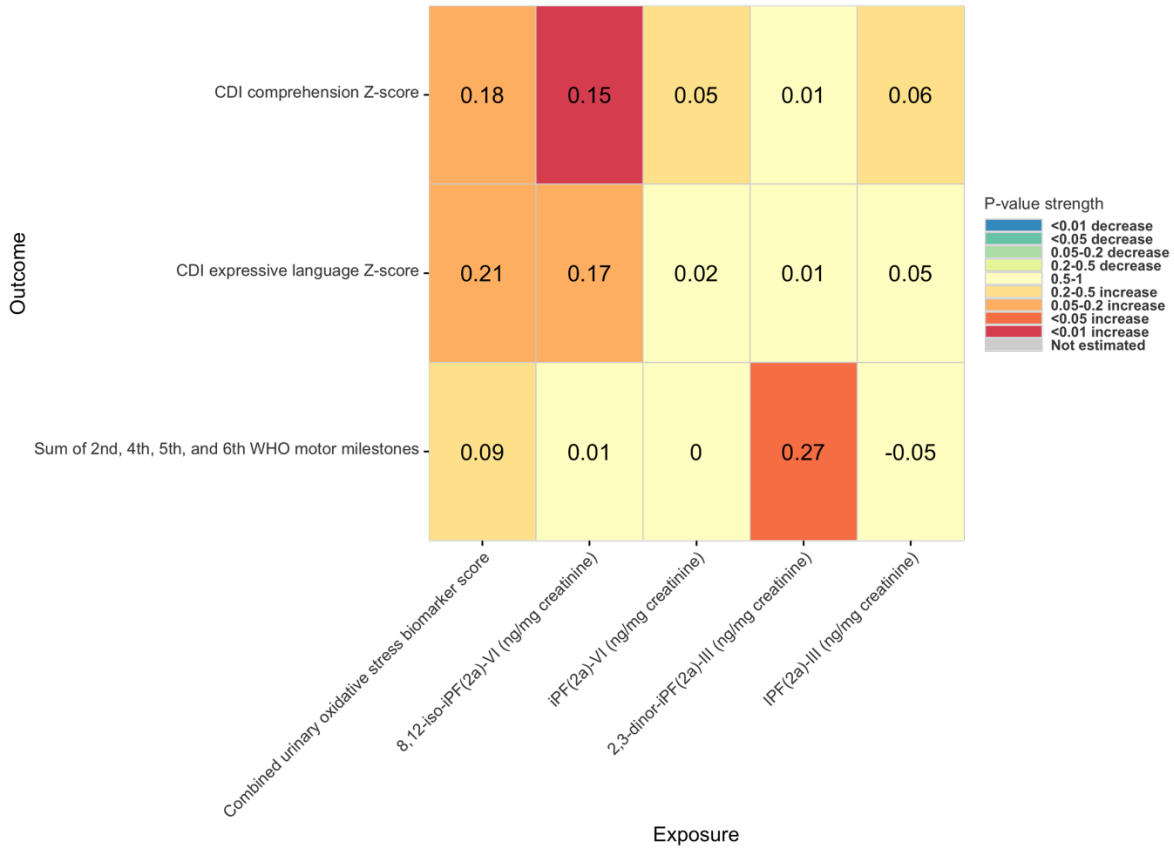
Figure 7. Mean arterial pressure, heart rate, and child development at Year 2



CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

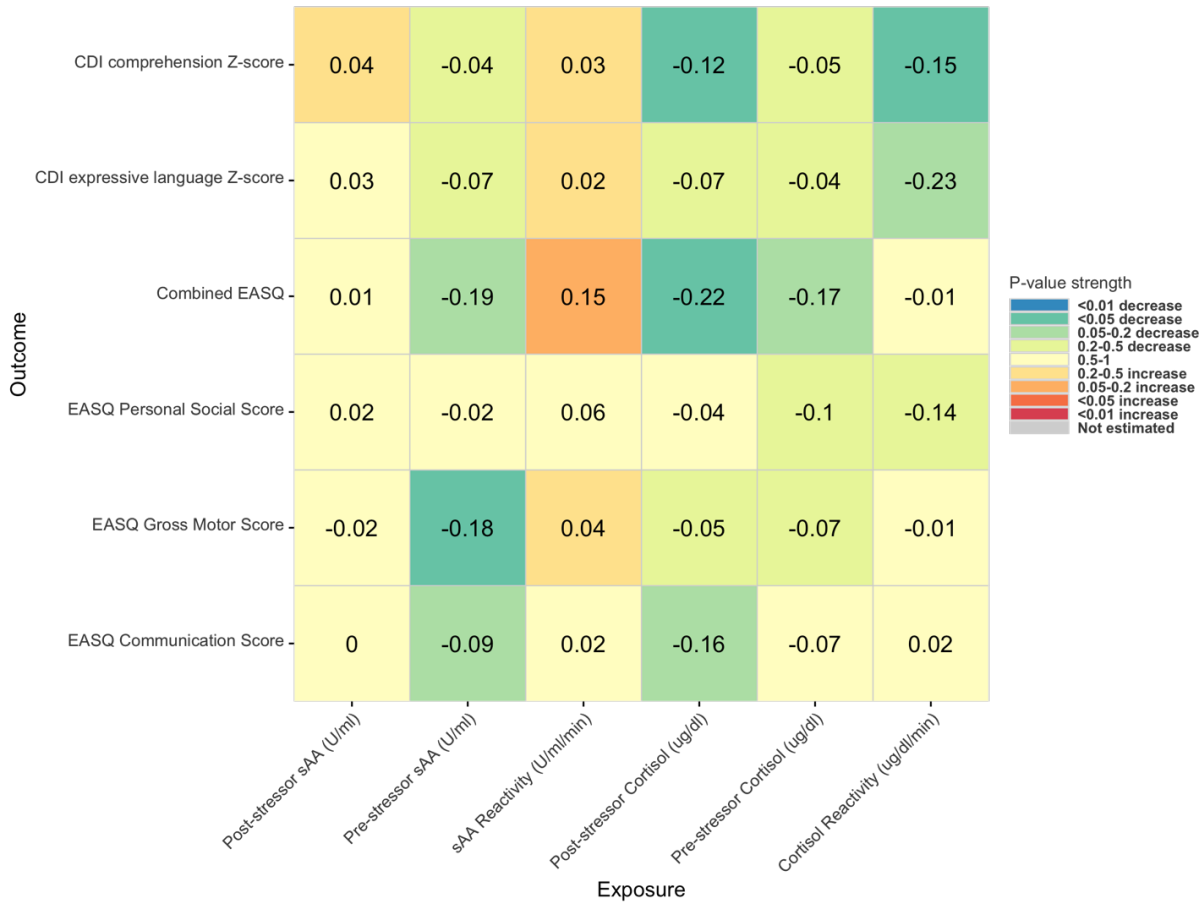
3.8 Supplemental Material

Supplemental Material 1. Heatmap of urinary isoprostanes and child development at Year 1



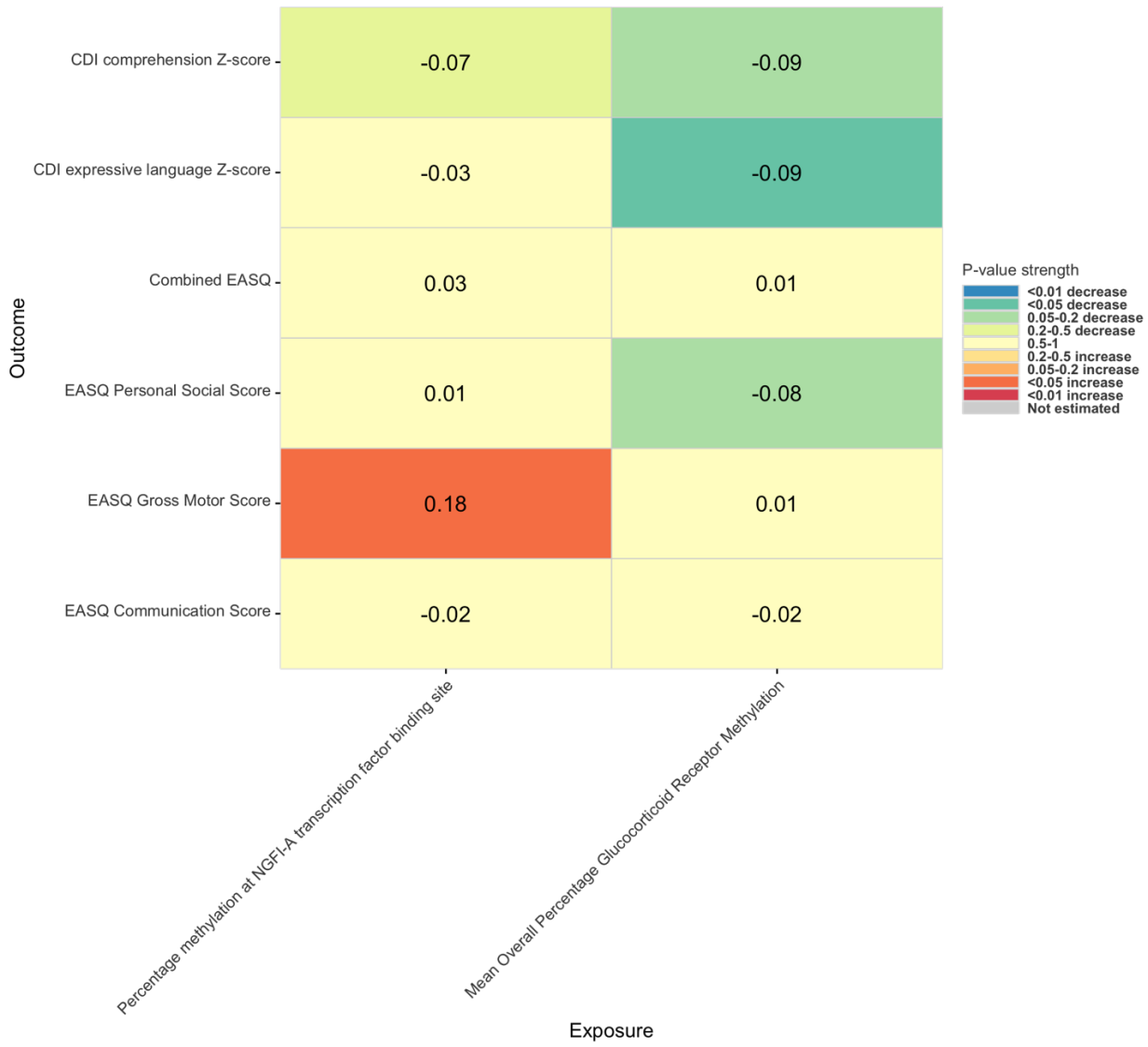
CDI: MacArthur-Bates Communicative Development Inventories

Supplemental Material 2. Heatmap of salivary stress biomarkers and child development at Year 2



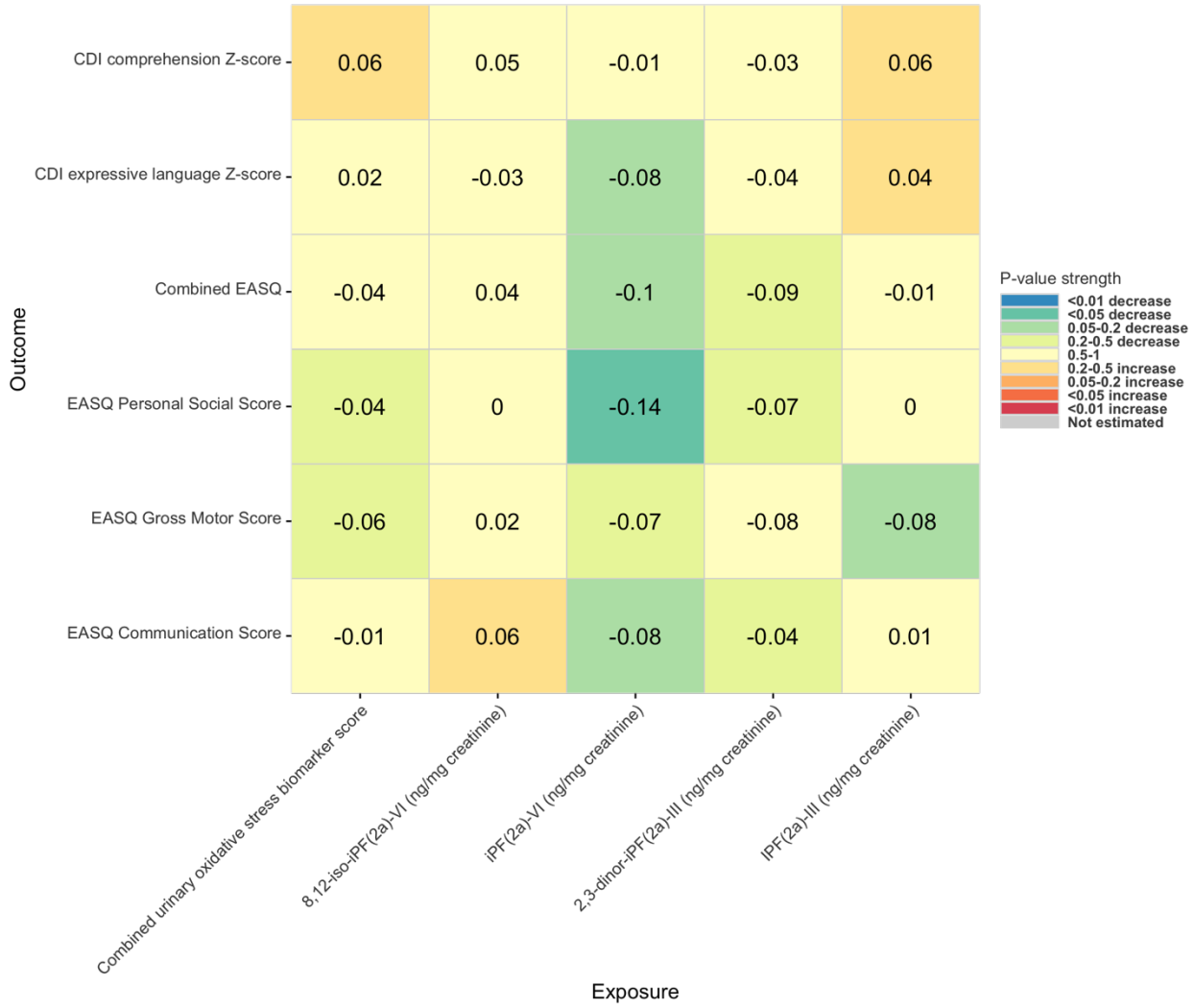
sAA: salivary alpha-amylase; CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Supplemental Material 3. Heatmap of glucocorticoid receptor methylation and child development at Year 2



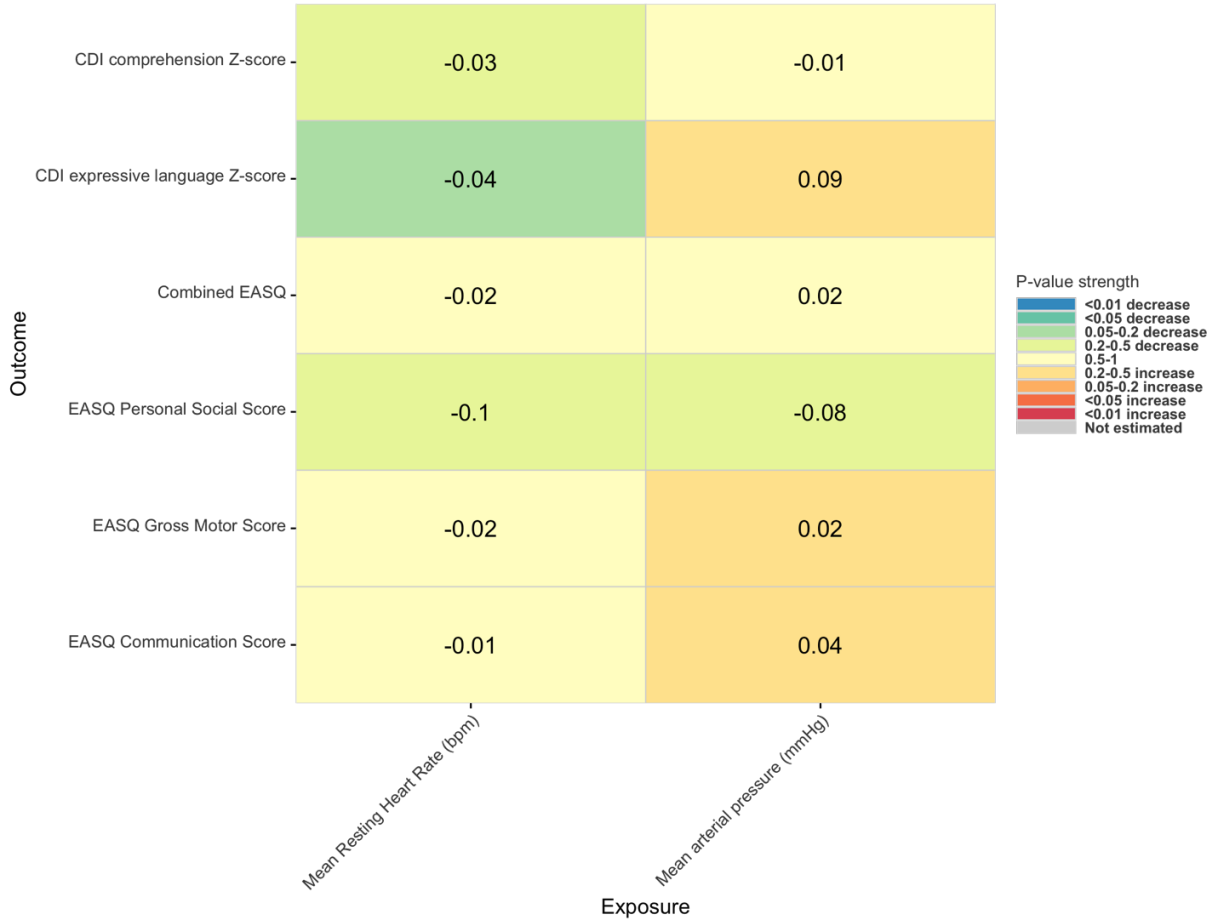
CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

Supplemental Material 4. Heatmap of urinary isoprostanes at Year 1 and child development at Year 2



EASQ: Extended Ages and Stages Questionnaire

Supplemental Material 5. Heatmap of mean arterial pressure, heart rate, and child development at Year 2



CDI: MacArthur-Bates Communicative Development Inventories; EASQ: Extended Ages and Stages Questionnaire

IV. CONCLUSION

The analyses and systematic review included here provide examples of methods to design and analyze trials in order to maximize inference. The first chapter reviews the applications of an underutilized trial design, while the second and third chapters demonstrate how alternative trial analyses can extend inferences beyond the primary analysis of the intervention effect. Although trials are the gold standard, solely relying on traditional methods in trial design and analysis can limit inference. The major limitations of trials include high cost and inability to evaluate exposures that cannot be feasibly or ethically assigned. These methods can ensure that trial resources are used effectively and that inferences are maximized using trial data.

Ring trials

Given the high cost of trials, it is crucial to ensure that trials will be adequately powered. In disease transmission settings with high spatiotemporal clustering, the use of a ring trial design can ensure that trials will have sufficient power despite heterogeneous incidence.

After reviewing 849 articles and 322 study protocols, we identified 26 ring trials, 15 cluster-randomized trials that used ring interventions, five trials that used ring recruitment and randomized within rings, and one individually-randomized trial that used a ring intervention. Ring trials require robust disease surveillance, accurate contact tracing, rapid intervention delivery, and an intervention with a strong post-exposure prophylactic effect. In settings with these characteristics, ring trials can retain power and randomization despite unpredictable clustering of the outcome of interest.

Future directions: applications of ring trials

Ring trial designs have already been used to evaluate prevention and treatment for SARS-CoV-2 infection, and these methods will continue to be essential to the evaluation of interventions to reduce the spread and severity of new variants of this virus (63). As ring trials are most effective for diseases in emergent or elimination transmission intensity settings, this study design will be most useful for waning variants of the virus or newly emerging variants. Beyond COVID-19, ring trial designs will be an effective tool for the evaluation of interventions related to novel infectious disease outbreaks.

In addition to infectious diseases, a ring trial design may be effective for non-infectious outcomes that spread through social networks. For example, interventions aiming to prevent gun violence have intervened on the social networks of gun violence victims in order to discourage retaliation and further violence (132,133). Similarly, social-behavioral outcomes such as extremist ideology, disinformation, and racial discrimination may spread through similar social network structures. A ring trial design may serve as an effective trial design to evaluate interventions that address these constructs, given the spatiotemporal clustering of these outcomes.

Treatment heterogeneity of N+WSH, WSH, and nutrition interventions on child growth by child pathogen and EED biomarker status

Investigators have hypothesized that EED and insufficient pathogen reduction may be to blame for WSH interventions' limited impact on child growth in a randomized context and have speculated that EED and infection may be associated with the impact of nutritional interventions (40). Given the high dimensionality of these biomarkers and pathogens, evaluation of these relationships using traditional regression methods is severely limited. In order to evaluate these

research questions, we applied cross-validated targeted maximum likelihood estimation and super learner to assess the treatment heterogeneity based on biomarker and pathogen status.

We analyzed data from 1,522 children to assess treatment heterogeneity based on individual pathogen and EED biomarker status. The included sample had a median LAZ of -1.56. Biomarkers of EED and pathogens were associated with the conditional average treatment effect for all interventions, and myeloperoxidase and *Campylobacter* at 14 months were associated with greater effect of all interventions on growth. An optimal treatment rule that allocated participants to intervention or control based on individual pathogen and EED biomarker status led to increased predicted child growth for N+WSH and WSH interventions, relative to the observed randomized treatment.

These findings highlight the potential of these targeted learning methods to assess treatment heterogeneity in a study population. These results indicate that pathogens and EED biomarkers are indicative of WSH and N+WSH interventions' impact on child growth. More specifically, the findings presented here point to *Campylobacter*, a common cause of diarrhea and potential contributor to EED, and myeloperoxidase, an EED marker of gut inflammation, as being related to nutrition, WSH, and N+WSH interventions' impact on child growth (188).

Future directions

These analyses evaluated biomarker and EED status after intervention had begun, leading to uncertainty regarding the causal impact of EED and pathogen status on intervention effectiveness. Future studies should evaluate child EED and pathogen status prior to randomization to assess the impact of these factors on treatment effectiveness.

Targeted learning and optimal treatment regime analysis have applications for a wide range of research questions. For both randomized and observational data, these analyses can assess treatment heterogeneity and identify subgroups of interest. For example, these methods could be used to identify patients with likely bacterial versus viral infections in order to optimize antibiotic allocation.

Child stress and child development

Previous studies have indicated relationships between child stress and child development (30). Adverse child experiences (ACES) are associated with poor developmental outcomes, although the impact of similar experiences varies widely between individuals (37). In order to quantify the risk of developmental impairment, investigators have sought to better understand which biomarkers best predict developmental status. We assessed various measures of HPA axis (cortisol and glucocorticoid receptor methylation), SAM axis (salivary alpha-amylase, heart rate, and blood pressure), oxidative status (F2-isoprostanes), and early childhood development (WHO gross motor milestones, EASQ, and CDI) in 684 children at 14 months of age and 1,449 children at 28 months of age.

We found that measures of HPA axis activity were correlated with child development, while we did not find a consistent relationship between SAM axis activity and child development. We found limited evidence that moderate oxidative status was positively associated with child development. This evidence supports the use of HPA axis biomarkers as measures of children who are at risk of poor developmental outcomes. These findings contribute to the literature base of the consequences of early life stress.

Future directions

These analyses were limited by primarily concurrent evaluation of stress exposures and developmental outcomes, which limits potential causal inference. Follow-up evaluation should include temporal separation of the exposures and outcomes of interest in order to bolster the possibility of causal inference and limit the possibility of reverse causation.

Furthermore, future studies should evaluate these relationships in an older population in order to evaluate the relationships between stress biomarker exposures and longer-term consequences, such as educational attainment or intelligence, which may be more indicative of long-term outcomes. In addition, dimension reduction methods may enable investigators to determine a clinically meaningful stress biomarker score that can summarize values.

Summary

Randomized controlled trials are the gold standard for causal inference, and methodological advances in both study design and analysis can improve our ability to draw meaningful inferences from these trials. This dissertation highlights methods that we can use to maximize inference from trials and provides examples of these methods. While this evaluation addresses multiple challenges in randomized trials (e.g., failed randomization, equipoise, generalizability etc.), each chapter seeks to maximize the inferences that can be drawn from trial data, which addresses the limitation of trials' high cost.

First, I conducted a systematic review of the ring trial design. Ring trial designs can provide a method to retain sufficient power despite high and unpredictable spatiotemporal clustering of the outcome, which is particularly useful for diseases in emergent and elimination transmission settings. Next, I assessed treatment heterogeneity of N+WSH, WSH, and N interventions on child growth based on pathogen and EED biomarker status. While traditional analysis of trials effectively provides an average treatment effect, the optimal allocation of this treatment and the impact of the intervention in subpopulations is often obscured. Through targeted learning analysis of treatment heterogeneity, I described replicable methods for assessment of the individual conditional average treatment effect and demonstrate that an optimal treatment rule based on EED and pathogen status could identify children who would most benefit from intervention. I found that myeloperoxidase and Campylobacter were associated with a greater effect of all interventions on growth. Finally, I assessed the relationship between child stress and child development. Through an observational analysis of trial data, I provide evidence that biomarkers of HPA axis activity are correlated with child development. This dissertation provides recommendations for researchers, replicable examples of methods, and contributions to subject-matter understanding.

It is the goal of global health researchers to maximize health outcomes of all individuals. The effective allocation of resources applies to research as much as it does to policy or clinical practice. Therefore, wasted resources in global health research can be equated to lives lost. It is imperative that global health researchers embrace the full spectrum of study design and analysis tools available in order to maximize efficiency.

V. REFERENCES

1. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG*. 2018;125(13):1716.

2. Höfler M. Causal inference based on counterfactuals. *BMC Medical Research Methodology*. 2005;5(1):28.

3. Spieth PM, Kubasch AS, Penzlin AI, et al. Randomized controlled trials - a matter of design. *Neuropsychiatr Dis Treat*. 2016;12:1341–1349.

4. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317(3):141–145.

5. Hemming K, Taljaard M, Moerbeek M, et al. Contamination: How much can an individually randomized trial tolerate? *Stat Med*. 2021;40(14):3329–3351.

6. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. 1. publ. London: Arnold; 2000 178 p.

7. Campbell MJ, Donner A, Klar N. Developments in cluster randomized trials and Statistics in Medicine. *Stat Med*. 2007;26(1):2–19.

8. Hayes RJ, Moulton LH. Cluster Randomised Trials. 0 ed. Chapman and Hall/CRC; 2009 (Accessed April 1, 2021).(https://www.taylorfrancis.com/books/9781584888178). (Accessed April 1, 2021)

9. Benjamin-Chung J, Arnold BF, Berger D, et al. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *International Journal of Epidemiology*. 2018;47(1):332–347.

10. Dean NE, Halloran ME, Longini IM. Design of vaccine trials during outbreaks with and without a delayed vaccination comparator. *Ann Appl Stat*. 2018;12(1):330–347.

11. Corbett MS, Higgins JPT, Woolacott NF. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Res Synth Methods*. 2014;5(1):79–85.

12. Ivers NM, Halperin IJ, Barnsley J, et al. Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials*. 2012;13(1):120.

13. Henao-Restrepo AM, Camacho A, Longini IM, et al. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). *Lancet*. 2017;389(10068):505–518.

14. Henao-Restrepo AM, Longini IM, Egger M, et al. Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the Guinea ring vaccination cluster-randomised trial. *Lancet*. 2015;386(9996):857–866.

15. Rose G. Sick individuals and sick populations. *International Journal of Epidemiology*. 2001;30(3):427–432.

16. VanderWeele TJ, Luedtke AR, van der Laan MJ, et al. Selecting Optimal Subgroups for Treatment Using Many Covariates. *Epidemiology*. 2019;30(3):334–341.

17. Bilkey GA, Burns BL, Coles EP, et al. Optimizing Precision Medicine for Public Health. *Front Public Health*. 2019;7:42–42.

18. Levy J, van der Laan M, Hubbard A, et al. A fundamental measure of treatment effect heterogeneity. *Journal of Causal Inference*. 2021;9(1):83–108.

19. Murphy SA. Optimal dynamic treatment regimes: *Dynamic Treatment Regimes*. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2003;65(2):331–355.

20. Zhao Y, Zeng D, Rush AJ, et al. Estimating Individualized Treatment Rules Using Outcome Weighted Learning. *Journal of the American Statistical Association*. 2012;107(499):1106–1118.

21. Robins J, Rotnitzky A. Discussion of "Dynamic treatment regimes: Technical challenges and applications." *Electron. J. Statist.* [electronic article]. 2014;8(1). (https://projecteuclid.org/journals/electronic-journal-of-statistics/volume-8/issue-1/Discussion-of-Dynamic-treatment-regimes-Technical-challenges-and-applications/10.1214/14-EJS908.full). (Accessed July 3, 2022)

22. Zhang B, A. Tsiatis A, Davidian M, et al. Estimating optimal treatment regimes from a classification perspective: Estimating optimal treatment regimes from a classification perspective. *Stat*. 2016;5(1):278–278.

23. Chakraborty B, Moodie EEM. Statistical Methods for Dynamic Treatment Regimes. New York, NY: Springer New York; 2013 (Accessed July 3, 2022).(http://link.springer.com/10.1007/978-1-4614-7428-9). (Accessed July 3, 2022)

24. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*. 1986;7(9–12):1393–1512.

25. van der Laan M, Coyle J, Hejazi N, et al. Targeted Learning in R: Causal Data Science with the tlverse Software Ecosystem. (https://tlverse.org/tlverse-handbook/optimal-individualized-treatment-regimes.html)

26. Gruber S, Van Der Laan MJ. Targeted maximum likelihood estimation: A gentle introduction. 2009;

27. van der Laan MJ, Gruber S. Collaborative double robust targeted maximum likelihood estimation. *Int J Biostat*. 2010;6(1):Article 17.

28. Black MM, Walker SP, Fernald LCH, et al. Early childhood development coming of age: science through the life course. *Lancet*. 2017;389(10064):77–90.

29. Null C, Stewart CP, Pickering AJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *Lancet Glob Health*. 2018;6(3):e316–e329.

30. Grantham-McGregor S, Cheung YB, Cueto S, et al. Developmental potential in the first 5 years for children in developing countries. *The Lancet*. 2007;369(9555):60–70.

31. Pickering AJ, Null C, Winch PJ, et al. The WASH Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea. *The Lancet Global Health*. 2019;7(8):e1139–e1146.

32. Franke HA. Toxic Stress: Effects, Prevention and Treatment. *Children (Basel)*. 2014;1(3):390–402.

33. McEwen BS. Effects of stress on the developing brain. *Cerebrum*. 2011;2011:14.

34. Engle PL, Fernald LCH, Alderman H, et al. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *Lancet*. 2011;378(9799):1339–1353.

35. National Scientific Council on the Developing Child. Excessive Stress Disrupts the Architecture of the Developing Brain. 2014.(https://developingchild.harvard.edu/)

36. Lu C, Black MM, Richter LM. Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *Lancet Glob Health*. 2016;4(12):e916–e922.

37. Walker SP, Wachs TD, Meeks Gardner J, et al. Child development: risk factors for adverse outcomes in developing countries. *The Lancet*. 2007;369(9556):145–157.

38. Hutton G, Chase C. Water Supply, Sanitation, and Hygiene. In: Mock CN, Nugent R, Kobusingye O, et al., eds. *Injury Prevention and Environmental Health*. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2017

39. Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting? Current evidence and policy implications. *Matern Child Nutr*. 2016;12 Suppl 1(Suppl 1):91–105.

40. Luby SP, Rahman M, Arnold BF, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *Lancet Glob Health*. 2018;6(3):e302–e315.

41. Gough EK, Moulton LH, Mutasa K, et al. Effects of improved water, sanitation, and hygiene and improved complementary feeding on environmental enteric dysfunction in children in

rural Zimbabwe: A cluster-randomized controlled trial. *PLoS Negl Trop Dis*. 2020;14(2):e0007963.

42. Galasso E, Weber AM, Stewart CP, et al. Effects of nutritional supplementation and home visiting on growth and development in young children in Madagascar: a cluster-randomised controlled trial. *Lancet Glob Health*. 2019;7(9):e1257–e1268.

43. Yackobovitch-Gavan M, Lebenthal Y, Lazar L, et al. Effect of Nutritional Supplementation on Growth in Short and Lean Prepubertal Children after 1 Year of Intervention. *J Pediatr*. 2016;179:154-159.e1.

44. Grembi JA, Lin A, Karim MA, et al. Effect of water, sanitation, handwashing and nutrition interventions on enteropathogens in children 14 months old: a cluster-randomized controlled trial in rural Bangladesh. *J Infect Dis*. 2020;

45. Watanabe K, Petri WAJ. Environmental Enteropathy: Elusive but Significant Subclinical Abnormalities in Developing Countries. *EBioMedicine*. 2016;10:25–32.

46. Ali A, Iqbal NT, Sadiq K. Environmental enteropathy. *Curr Opin Gastroenterol*. 2016;32(1):12–17.

47. Woolhouse ME, Dye C, Etard JF, et al. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc Natl Acad Sci U S A*. 1997;94(1):338–342.

48. Lipsitch M, Eyal N. Improving vaccine trials in infectious disease emergencies. *Science*. 2017;357(6347):153–156.

49. Cotter C, Sturrock HJW, Hsiang MS, et al. The changing epidemiology of malaria elimination: new strategies for new challenges. *Lancet*. 2013;382(9895):900–911.

50. Foege WH, Millar JD, Lane JM. Selective epidemiologic control in smallpox eradication. *Am J Epidemiol*. 1971;94(4):311–315.

51. Dean NE, Gsell P-S, Brookmeyer R, et al. Design of vaccine efficacy trials during public health emergencies. *Science Translational Medicine*. 2019;11(499):eaat0360.

52. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ*. 2001;322(7282):355–357.

53. Okell LC, Bousema T, Griffin JT, et al. Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nat Commun*. 2012;3:1237.

54. Finger F, Bertuzzo E, Luquero FJ, et al. The potential impact of case-area targeted interventions in response to cholera outbreaks: A modeling study. *PLOS Medicine*. 2018;15(2):e1002509.

55. Michel E, Gaudart J, Beaulieu S, et al. Estimating effectiveness of case-area targeted response interventions against cholera in Haiti. *Elife*. 2019;8.

56. Rebaudet S, Bulit G, Gaudart J, et al. The case-area targeted rapid response strategy to control cholera in Haiti: a four-year implementation study. *PLOS Neglected Tropical Diseases*. 2019;13(4):e0007263.

57. Bath D, Cook J, Govere J, et al. Effectiveness and cost-effectiveness of reactive, targeted indoor residual spraying for malaria control in low-transmission settings: a cluster-randomised, non-inferiority trial in South Africa. *Lancet*. 2021;397(10276):816–827.

58. Bridges DJ, Miller JM, Chalwe V, et al. Community-led Responses for Elimination (CoRE): a study protocol for a community randomized controlled trial assessing the effectiveness of community-level, reactive focal drug administration for reducing Plasmodium falciparum infection prevalence and incidence in Southern Province, Zambia. *Trials*. 2017;18(1):511.

59. Eisele TP, Silumbe K, Finn T, et al. Assessing the effectiveness of household-level focal mass drug administration and community-wide mass drug administration for reducing malaria parasite infection prevalence and incidence in Southern Province, Zambia: study protocol for a community randomized controlled trial. *Trials*. 2015;16:347.

60. Vilakati S, Mngadi N, Benjamin-Chung J, et al. Effectiveness and safety of reactive focal mass drug administration (rfMDA) using dihydroartemisinin-piperaquine to reduce malaria transmission in the very low-endemic setting of Eswatini: a pragmatic cluster randomised controlled trial. *BMJ Glob Health*. 2021;6(6).

61. Hsiang MS, Ntuku H, Roberts KW, et al. Effectiveness of reactive focal mass drug administration and reactive focal vector control to reduce malaria transmission in the low malaria-endemic setting of Namibia: a cluster-randomised controlled, open-label, two-by-two factorial design trial. *Lancet*. 2020;395(10233):1361–1373.

62. Sturrock HJW, Novotny JM, Kunene S, et al. Reactive Case Detection for Malaria Elimination: Real-Life Experience from an Ongoing Program in Swaziland. 2013;8(5):e63830.

63. Xu W, Su S, Jiang S. Ring vaccination of COVID-19 vaccines in medium- and high-risk areas of countries with low incidence of SARS-CoV-2 infection. *Clin Transl Med*. 2021;11(2):e331.

64. Grais R f, Conlan A j. k, Ferrari M j, et al. Time is of the essence: exploring a measles outbreak response vaccination in Niamey, Niger. *Journal of The Royal Society Interface*. 2008;5(18):67–74.

65. Lessler J, Metcalf CJE, Cutts FT, et al. Impact on Epidemic Measles of Vaccination Campaigns Triggered by Disease Outbreaks or Serosurveys: A Modeling Study. *PLOS Medicine*. 2016;13(10):e1002144.

66. Strebel PM, Cochi SL. Waving goodbye to measles. *Nature*. 2001;414(6865):695–696.

67. Andradóttir S, Chiu W, Goldsman D, et al. Reactive strategies for containing developing outbreaks of pandemic influenza. *BMC Public Health*. 2011;11(1):S1.

68. Chowell G, Fuentes R, Olea A, et al. The Basic Reproduction Number $\mathscr{R}0$ and Effectiveness of Reactive Interventions during Dengue Epidemics: The 2002 Dengue Outbreak in Easter Island, Chile. *Math Biosci Eng*. 2013;10(5–6):1455–1474.

69. Trentini F, Poletti P, Baldacchino F, et al. The containment of potential outbreaks triggered by imported Chikungunya cases in Italy: a cost utility epidemiological assessment of vector control measures. *Scientific Reports*. 2018;8(1):9034.

70. Ndeffo-Mbah ML, Durham DP, Skrip LA, et al. Evaluating the effectiveness of localized control strategies to curtail chikungunya. *Scientific Reports*. 2016;6(1):23997.

71. Hitchings MDT, Coldiron ME, Grais RF, et al. Analysis of a meningococcal meningitis outbreak in Niger – potential effectiveness of reactive prophylaxis. *PLOS Neglected Tropical Diseases*. 2019;13(3):e0007077.

72. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.

73. Ohrt C, Roberts KW, Sturrock HJW, et al. Information systems to support surveillance for malaria elimination. *Am J Trop Med Hyg*. 2015;93(1):145–152.

74. Kapesa A, Kweka EJ, Zhou G, et al. Utility of passive malaria surveillance in hospitals as a surrogate to community infection transmission dynamics in western Kenya. *Archives of Public Health*. 2018;76(1):39.

75. Owada K, Eckmanns T, Kamara K-BO, et al. Epidemiological Data Management during an Outbreak of Ebola Virus Disease: Key Issues and Observations from Sierra Leone. *Front Public Health*. 2016;4:163–163.

76. Ortuno-Gutierrez N, Younoussa A, Randrianantoandro A, et al. Protocol, rationale and design of PEOPLE (Post ExpOsure Prophylaxis for LEprosy in the Comoros and Madagascar): a cluster randomized trial on effectiveness of different modalities of implementation of post-exposure prophylaxis of leprosy contacts. *BMC Infect Dis*. 2019;19(1):1033.

77. Bausch DG, Piot P. Ebola Vaccines: Biomedical Advances, Human Rights Challenges: Commentary on: A Randomized, Blinded, Dose-Ranging Trial of an Ebola Virus Glycoprotein (EBOVGP) Nanoparticle Vaccine With Matrix-M Adjuvant in Healthy Adults. *The Journal of Infectious Diseases*. 2020;222(4):521–524.

78. Hitchings MDT, Grais RF, Lipsitch M. Using simulation to aid trial design: Ring-vaccination trials. *PLoS Negl Trop Dis*. 2017;11(3):e0005470.

79. Horne AD, Lachenbruch PA, Goldenthal KL. Intent-to-treat analysis and preventive vaccine efficacy. *Vaccine*. 2000;19(2–3):319–326.

80. The ring vaccination trial: a novel cluster randomised controlled trial design to evaluate vaccine efficacy and effectiveness during outbreaks, with special reference to Ebola. *BMJ*. 2015;351:h3740.

81. Jiang Z, Wang X, Xia J. Considerations on the clinical development of COVID-19 vaccine from trial design perspectives. *Hum Vaccin Immunother*. 2021;17(3):656–660.

82. Kretzschmar M, van den Hof S, Wallinga J, et al. Ring vaccination and smallpox control. *Emerg Infect Dis*. 2004;10(5):832–841.

83. Clemens J, Shin S, Ali M. New approaches to the assessment of vaccine herd protection in clinical trials. *Lancet Infect Dis*. 2011;11(6):482–487.

84. Halloran ME, Longini IM, Struchiner CJ. Design and analysis of vaccine studies. New York: Springer; 2010 387 p.

85. Coldiron ME, Assao B, Page A-L, et al. Single-dose oral ciprofloxacin prophylaxis as a response to a meningococcal meningitis epidemic in the African meningitis belt: A 3-arm, open-label, cluster-randomized trial. *PLoS Med*. 2018;15(6):e1002593.

86. Okebe J, Dabira E, Jaiteh F, et al. Reactive, self-administered malaria treatment against asymptomatic malaria infection: results of a cluster randomized controlled trial in The Gambia. *Malar J*. 2021;20(1):253.

87. Salazar-Austin N, Cohn S, Barnes GL, et al. Improving Tuberculosis Preventive Therapy Uptake: A Cluster-randomized Trial of Symptom-based Versus Tuberculin Skin Test-based Screening of Household Tuberculosis Contacts Less Than 5 Years of Age. *Clin Infect Dis*. 2020;70(8):1725–1732.

88. Echevarría J, Seas C, Carrillo C, et al. Efficacy and tolerability of ciprofloxacin prophylaxis in adult household contacts of patients with cholera. *Clin Infect Dis*. 1995;20(6):1480–1484.

89. Halperin SA, Bortolussi R, Langley JM, et al. A randomized, placebo-controlled trial of erythromycin estolate chemoprophylaxis for household contacts of children with culture-positive bordetella pertussis infection. *Pediatrics*. 1999;104(4):e42.

90. Welliver R, Monto AS, Carewicz O, et al. Effectiveness of oseltamivir in preventing influenza in household contacts: a randomized controlled trial. *JAMA*. 2001;285(6):748–754.

91. Herzog C, Berger R, Fernex M, et al. Intranasal interferon (rIFN-alpha A, Ro 22-8181) for contact prophylaxis against common cold: a randomized, double-blind and placebo-controlled field study. *Antiviral Res*. 1986;6(3):171–176.

92. Barnabas RV, Brown ER, Bershteyn A, et al. Hydroxychloroquine as Postexposure Prophylaxis to Prevent Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *Ann Intern Med*. 2021;174(3):344–352.

93. Ikematsu H, Hayden FG, Kawaguchi K, et al. Baloxavir Marboxil for Prophylaxis against Influenza in Household Contacts. *N Engl J Med*. 2020;383(4):309–320.

94. Kashiwagi S, Watanabe A, Ikematsu H, et al. Laninamivir octanoate for post-exposure prophylaxis of influenza in household contacts: a randomized double blind placebo controlled trial. *J Infect Chemother*. 2013;19(4):740–749.

95. Kashiwagi S, Watanabe A, Ikematsu H, et al. Long-acting Neuraminidase Inhibitor Laninamivir Octanoate as Post-exposure Prophylaxis for Influenza. *Clin Infect Dis*. 2016;63(3):330–337.

96. Nakano T, Ishiwada N, Sumitani T, et al. Inhaled Laninamivir Octanoate as Prophylaxis for Influenza in Children. *Pediatrics*. 2016;138(6).

97. Hayden FG, Gubareva LV, Monto AS, et al. Inhaled zanamivir for the prevention of influenza in families. Zanamivir Family Study Group. *N Engl J Med*. 2000;343(18):1282–1289.

98. van der Sande MA, Meijer A, Sen-Kerpiclik F, et al. Effectiveness of post-exposition prophylaxis with oseltamivir in nursing homes: a randomised controlled trial over four seasons. *Emerg Themes Epidemiol*. 2014;11:13.

99. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet*. 2002;359(9307):696–700.

100. Ram PK, DiVita MA, Khatun-e-Jannat K, et al. Impact of Intensive Handwashing Promotion on Secondary Household Influenza-Like Illness in Rural Bangladesh: Findings from a Randomized Controlled Trial. *PLoS One*. 2015;10(6):e0125200.

101. Hernández AV, Steyerberg EW, Habbema JDF. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *Journal of Clinical Epidemiology*. 2004;57(5):454–460.

102. Suess T, Remschmidt C, Schink SB, et al. The role of facemasks and hand hygiene in the prevention of influenza transmission in households: results from a cluster randomised trial; Berlin, Germany, 2009-2011. *BMC Infect Dis*. 2012;12:26.

103. Moerbeek M, Schie S van. What are the statistical implications of treatment non-compliance in cluster randomized trials: A simulation study. *Stat Med*. 2019;38(26):5071–5084.

104. Mitja O, Corbacho-Monne M, Ubals M, et al. A Cluster-Randomized Trial of Hydroxychloroquine for Prevention of Covid-19. *N Engl J Med*. 2021;384(5):417–427.

105. Hayden FG, Belshe R, Villanueva C, et al. Management of influenza in households: a prospective, randomized comparison of oseltamivir treatment with or without postexposure prophylaxis. *J Infect Dis*. 2004;189(3):440–449.

106. Lee YJ, Ellenberg JH, Hirtz DG, et al. Analysis of clinical trials by treatment actually received: is it really an option? *Stat Med*. 1991;10(10):1595–1605.

107. Frangakis C, Rubin D. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*. 1999;86(2):365–379.

108. Stuart EA, Bradshaw CP, Leaf PJ. Assessing the generalizability of randomized trial results to target populations. *Prev Sci*. 2015;16(3):475–485.

109. Dron L, Taljaard M, Cheung YB, et al. The role and challenges of cluster randomised trials for global health. *The Lancet Global Health*. 2021;9(5):e701–e710.

110. Rutterford C, Copas A, Eldridge S. Methods for sample size determination in cluster randomized trials. *Int J Epidemiol*. 2015;44(3):1051–1067.

111. Eisele TP, Bennett A, Silumbe K, et al. Impact of Four Rounds of Mass Drug Administration with Dihydroartemisinin-Piperaquine Implemented in Southern Province, Zambia. *Am J Trop Med Hyg*. 2020;103(2_Suppl):7–18.

112. Cowling BJ, Fung ROP, Cheng CKY, et al. Preliminary findings of a randomized trial of non-pharmaceutical interventions to prevent influenza transmission in households. *PLoS One*. 2008;3(5):e2101.

113. Smit M, Marinosci A, Nicoletti GJ, et al. Efficacy of pragmatic same-day ring prophylaxis for adult individuals exposed to SARS-CoV-2 in Switzerland (COPEP): protocol of an open-label cluster randomised trial. *BMJ Open*. 2020;10(11):e040110.

114. Tan DHS, Chan AK, Jüni P, et al. Post-exposure prophylaxis against SARS-CoV-2 in close contacts of confirmed COVID-19 cases (CORIPREV): study protocol for a cluster-randomized trial. *Trials*. 2021;22(1):224.

115. Jo B, Asparouhov T, Muthén BO. Intention-to-treat analysis in cluster randomized trials with noncompliance. *Stat Med*. 2008;27(27):5565–5577.

116. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol*. 1978;108(2):100–102.

117. Hsiang MS, Ntshalintshali N, Kang Dufour M-S, et al. Active Case Finding for Malaria: A 3-Year National Evaluation of Optimal Approaches to Detect Infections and Hotspots Through Reactive Case Detection in the Low-transmission Setting of Eswatini. *Clin Infect Dis*. 2020;70(7):1316–1325.

118. Smith JL, Auala J, Tambo M, et al. Spatial clustering of patent and sub-patent malaria infections in northern Namibia: Implications for surveillance and response strategies for elimination. *PLoS One*. 2017;12(8):e0180845.

119. Staples PC, Ogburn EL, Onnela J-P. Incorporating Contact Network Structure in Cluster Randomized Trials. *Scientific Reports*. 2015;5(1):17581.

120. Halloran ME, Auranen K, Baird S, et al. Simulations for designing and interpreting intervention trials in infectious diseases. *BMC Medicine*. 2017;15(1):223.

121. Weijer C, Grimshaw JM, Eccles MP, et al. The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials. *PLOS Medicine*. 2012;9(11):e1001346.

122. Choko AT, Roshandel G, Conserve DF, et al. Ethical issues in cluster randomized trials conducted in low- and middle-income countries: an analysis of two case studies. *Trials*. 2020;21(1):314.

123. Osrin D, Azad K, Fernandez A, et al. Ethical challenges in cluster randomized controlled trials: experiences from public health interventions in Africa and Asia. *Bull World Health Organ*. 2009;87(10):772–779.

124. Zani B, Gathu M, Donegan S, et al. Dihydroartemisinin-piperaquine for treating uncomplicated Plasmodium falciparum malaria. *Cochrane Database Syst Rev*. 2014;2014(1):CD010927.

125. Artemether-lumefantrine treatment of uncomplicated Plasmodium falciparum malaria: a systematic review and meta-analysis of day 7 lumefantrine concentrations and therapeutic response using individual patient data. *BMC Med*. 2015;13:227.

126. Poirot E, Skarbinski J, Sinclair D, et al. Mass drug administration for malaria. *Cochrane Database Syst Rev*. 2013;2013(12):CD008846.

127. Newby G, Hwang J, Koita K, et al. Review of mass drug administration for malaria and its operational challenges. *Am J Trop Med Hyg*. 2015;93(1):125–134.

128. Rid A, Miller FG. Ethical Rationale for the Ebola "Ring Vaccination" Trial Design. *Am J Public Health*. 2016;106(3):432–435.

129. Gsell P-S, Camacho A, Kucharski AJ, et al. Ring vaccination with rVSV-ZEBOV under expanded access in response to an outbreak of Ebola virus disease in Guinea, 2016: an operational and vaccine safety report. *Lancet Infect Dis*. 2017;17(12):1276–1284.

130. Chow S-C, Chang M. Adaptive design methods in clinical trials. 2nd ed. Boca Raton: Taylor & Francis; 2012 358 p.

131. Bellan SE, Pulliam JRC, Pearson CAB, et al. Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation study of trial design and analysis. *Lancet Infect Dis*. 2015;15(6):703–710.

132. Larsen DA, Lane S, Jennings-Bey T, et al. Spatio-temporal patterns of gun violence in Syracuse, New York 2009-2015. *PLoS One*. 2017;12(3):e0173001.

133. Tracy M, Braga AA, Papachristos AV. The Transmission of Gun and Other Weapon-Involved Violence Within Social Networks. *Epidemiologic Reviews*. 2016;38(1):70–86.

134. Kumar N, Oles W, Howell BA, et al. The role of social network support in treatment outcomes for medication for opioid use disorder: A systematic review. *J Subst Abuse Treat*. 2021;127:108367.

135. Park JJH, Grais RF, Taljaard M, et al. Urgently seeking efficiency and sustainability of clinical trials in global health. *Lancet Glob Health*. 2021;9(5):e681–e690.

136. Park JJH, Mogg R, Smith GE, et al. How COVID-19 has fundamentally changed clinical research in global health. *The Lancet Global Health*. 2021;9(5):e711–e720.

137. Barnabas RV, Brown E, Bershteyn A, et al. Efficacy of hydroxychloroquine for post-exposure prophylaxis to prevent severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection among adults exposed to coronavirus disease (COVID-19): a structured summary of a study protocol for a randomised controlled trial. *Trials*. 2020;21(1):475.

138. Bridges DJ. Community-led Responses for Elimination: Controlled Trial of Reactive Case Detection Versus Reactive Drug Administration (CORE Zambia. *ClinicalTrials.gov*. 2016;(https://clinicaltrials.gov/ct2/show/NCT02654912). (Accessed April 1, 2021)

139. Coldiron ME, Alcoba G, Ciglenecki I, et al. Ciprofloxacin for contacts of cases of meningococcal meningitis as an epidemic response: study protocol for a cluster-randomized trial. *Trials*. 2017;18(1):294.

140. Egsmose T, Ang'awa JO, Poti SJ. The use of isoniazid among household contacts of open cases of pulmonary tuberculosis. *Bull World Health Organ*. 1965;33(3):419–433.

141. Eisele TP, Bennett A, Silumbe K, et al. Short-term Impact of Mass Drug Administration With Dihydroartemisinin Plus Piperaquine on Malaria in Southern Province Zambia: A Cluster-Randomized Controlled Trial. *J Infect Dis*. 2016;214(12):1831–1839.

142. Fritz SA, Hogan PG, Hayek G, et al. Household versus individual approaches to eradication of community-associated Staphylococcus aureus in children: a randomized trial. *Clin Infect Dis*. 2012;54(6):743–751.

143. George CM, Monira S, Zohura F, et al. Effects of a Water, Sanitation and Hygiene Mobile Health Program on Diarrhea and Child Growth in Bangladesh: A Cluster-Randomized Controlled Trial of the CHoBI7 Mobile Health Program. *Clin Infect Dis*. 2020;

144. Masud J, Islam Bhuyian MS, Kumar Biswas S, et al. Diarrhoeal disease knowledge among diarrhoea patient housholds: findings from the randomised controlled trial of the Cholera-Hospital-Based-Intervention-for-7-days (CHoBI7) mobile health program. *Trop Med Int Health*. 2020;25(8):996–1007.

145. Medzihradsky OF, Kleinschmidt I, Mumbengegwi D, et al. Study protocol for a cluster randomised controlled factorial design trial to assess the effectiveness and feasibility of reactive focal mass drug administration and vector control to reduce malaria transmission in the low endemic setting of Namibia. *BMJ Open*. 2018;8(1):e019294.

146. Hsiang MS. Evaluation of Reactive Focal Mass Drug Administration (rfMDA) +/- Reactive Focal Vector Control (RAVC) in Namibia. *ClinicalTrials.gov*. 2015;(https://clinicaltrials.gov/ct2/show/NCT02610400). (Accessed April 1, 2021)

147. Iturriaga C, Eiffler N, Aniba R, et al. A cluster randomized trial of interferon ß-1a for the reduction of transmission of SARS-Cov-2: protocol for the Containing Coronavirus Disease 19 trial (ConCorD-19). *BMC Infect Dis*. 2021;21(1):814–814.

148. Low N, McCarthy A, Roberts TE, et al. Partner notification of chlamydia infection in primary care: randomised controlled trial and analysis of resource use. *BMJ*. 2006;332(7532):14–19.

149. Murphy TV, Chrane DF, McCracken GHJ, et al. Rifampin prophylaxis v placebo for household contacts of children with Hemophilus influenzae type b disease. *Am J Dis Child*. 1983;137(7):627–632.

150. Nanni O, Viale P, Vertogen B, et al. PROTECT Trial: A cluster-randomized study with hydroxychloroquine versus observational support for prevention or early-phase treatment of Coronavirus disease (COVID-19): A structured summary of a study protocol for a randomized controlled trial. *Trials*. 2020;21(1):689.

151. de Jong B, Hasker E, Assoumani Y, et al. Post ExpOsure Prophylaxis for LEprosy in the Comoros and Madagascar (PEOPLE). *ClinicalTrials.gov*. 2018;(https://ClinicalTrials.gov/show/NCT03662022). (Accessed August 25, 2021)

152. Sagliocca L, Amoroso P, Stroffolini T, et al. Efficacy of hepatitis A vaccine in prevention of secondary hepatitis A infection: a randomised trial. *Lancet*. 1999;353(9159):1136–1139.

153. Seddon JA, Garcia-Prats AJ, Purchase SE, et al. Levofloxacin versus placebo for the prevention of tuberculosis disease in child contacts of multidrug-resistant tuberculosis: study protocol for a phase III cluster randomised controlled trial (TB-CHAMP). *Trials*. 2018;19(1):693.

154. Calmy A, Labhardt ND. Efficacy of Pragmatic Same-day COVID-19 Ring Prophylaxis for Adult Individuals Exposed to SARS-CoV-2 in Switzerland (COPEP). *ClinicalTrials.gov*. 2020;(https://clinicaltrials.gov/ct2/show/NCT04364022). (Accessed April 1, 2021)

155. van der Sande MA. A Randomised Controlled Trial on the Effect of Post-exposure Oseltamivir Prophylaxis on Influenza Transmission in Nursing Homes (PEPpIE). *ClinicalTrials.gov*. 2010;(https://ClinicalTrials.gov/show/NCT01053377). (Accessed August 25, 2021)

156. Vasiliu A, Eymard-Duvernay S, Tchounga B, et al. Community intervention for child tuberculosis active contact investigation and management: study protocol for a parallel cluster randomized controlled trial. *Trials*. 2021;22(1):180.

157. Hsiang MS. Evaluation of Reactive Focal Mass Drug Administration for Malaria Elimination in Swaziland (fMDA). *ClinicalTrials.gov*. 2014;(https://clinicaltrials.gov/ct2/show/NCT02315690). (Accessed April 1, 2021)

158. Wamuti BM, Erdman LK, Cherutich P, et al. Assisted partner notification services to augment HIV testing and linkage to care in Kenya: study protocol for a cluster randomized trial. *Implement Sci*. 2015;10:23.

159. Cherutich P, Golden MR, Wamuti B, et al. Assisted partner services for HIV in Kenya: a cluster randomised controlled trial. *Lancet HIV*. 2017;4(2):e74–e82.

160. Wang R, DeGruttola V, Lei Q, et al. The vitamin D for COVID-19 (VIVID) trial: A pragmatic cluster-randomized design. *Contemp Clin Trials*. 2021;100:106176.

161. Wingfield T, Tovar MA, Huff D, et al. A randomized controlled study of socioeconomic support to enhance tuberculosis prevention and treatment, Peru. *Bull World Health Organ*. 2017;95(4):270–280.

162. Agrawal R. Safety And Efficacy Of Hydroxychloroquine For At Risk Population (SHARP) Against COVID-19 (SHARP COVID-19). *ClinicalTrials.gov*. 2020;(https://ClinicalTrials.gov/show/NCT04342156). (Accessed August 25, 2021)

163. Bardin M. Trial to Evaluate the Efficacy and Safety of Nitazoxanide (NTZ) for Post-Exposure Prophylaxis of COVID-19 and Other Viral Respiratory Illnesses in Elderly Residents of Long-Term Care Facilities (LTCF). *ClinicalTrials.gov*. 2020;(https://ClinicalTrials.gov/show/NCT04343248). (Accessed August 25, 2021)

164. Bennett A, Vanisaveth V. Targeting High Risk Populations With Enhanced Reactive Case Detection in Southern Lao Peoples Democratic Republic (COMBAT). *ClinicalTrials.gov*. 2020;(https://clinicaltrials.gov/ct2/show/NCT04416945). (Accessed April 1, 2021)

165. Borrie M. COVID-19 PEP- High-risk Individuals in Long-term and Specialized Care - Canada. *ClinicalTrials.gov*. 2020;(https://ClinicalTrials.gov/show/NCT04397328). (Accessed August 25, 2021)

166. Bracchi A, Flego M. PREvent Viral Exposure aNd Transmission Study: a SARS-CoV-2 PEP Study (PREVENT). *ClinicalTrials.gov*. 2021;(https://ClinicalTrials.gov/show/NCT04842331). (Accessed August 25, 2021)

167. Elvira M. Airborne Preventive Measures to Reduce New TB Infections in Household Contacts (TBMask). *ClinicalTrials.gov*. 2021;(https://ClinicalTrials.gov/show/NCT04938596). (Accessed August 25, 2021)

168. Gadisa E. Evaluation of Targeted Mass Drug Administration for Malaria in Ethiopia. *ClinicalTrials.gov*. 2020;(https://clinicaltrials.gov/ct2/show/NCT04241705). (Accessed April 1, 2021)

169. Giles J. Hydroxychloroquine Post Exposure Prophylaxis for Coronavirus Disease (COVID-19). *ClinicalTrials.gov*. 2020;(https://ClinicalTrials.gov/show/NCT04318444). (Accessed August 25, 2021)

170. Malin R. A Study to Test Whether BI 767551 Can Prevent COVID-19 in People Who Have Been Exposed to SARS-CoV-2. *ClinicalTrials.gov*. 2021;(https://ClinicalTrials.gov/show/NCT04894474). (Accessed August 25, 2021)

171. McGeer A. Control of COVID-19 Outbreaks in Long Term Care. *ClinicalTrials.gov*. 2020;(https://ClinicalTrials.gov/show/NCT04448119). (Accessed August 25, 2021)

172. Sued O. Efficacy and Safety of Nitazoxanide for Post Exposure Prophylaxis of COVID-19 in Household Contacts (PENTZ). *ClinicalTrials.gov*. 2021;(https://ClinicalTrials.gov/show/NCT04788407). (Accessed August 25, 2021)

173. Sudfeld CR, McCoy DC, Danaei G, et al. Linear growth and child development in low- and middle-income countries: a meta-analysis. *Pediatrics*. 2015;135(5):e1266-1275.

174. Katona P, Katona-Apte J. The Interaction between Nutrition and Infection. *Clinical Infectious Diseases*. 2008;46(10):1582–1588.

175. Giallourou N, Medlock GL, Bolick DT, et al. A novel mouse model of Campylobacter jejuni enteropathy and diarrhea. *PLOS Pathogens*. 2018;14(3):e1007083.

176. Donowitz Jeffrey R., Haque Rashidul, Kirkpatrick Beth D., et al. Small Intestine Bacterial Overgrowth and Environmental Enteropathy in Bangladeshi Children. *mBio*. 7(1):e02102-15.

177. Donowitz JR, Pu Z, Lin Y, et al. Small Intestine Bacterial Overgrowth in Bangladeshi Infants Is Associated With Growth Stunting in a Longitudinal Cohort. *Official journal of the American College of Gastroenterology | ACG* [electronic article]. 2022;117(1). (https://journals.lww.com/ajg/Fulltext/2022/01000/Small_Intestine_Bacterial_Overgrowth_in.26.aspx)

178. Lin A, Ali S, Arnold BF, et al. Effects of Water, Sanitation, Handwashing, and Nutritional Interventions on Environmental Enteric Dysfunction in Young Children: A Cluster-randomized, Controlled Trial in Rural Bangladesh. *Clinical Infectious Diseases*. 2020;70(5):738–747.

179. Lin A, Ercumen A, Benjamin-Chung J, et al. Effects of Water, Sanitation, Handwashing, and Nutritional Interventions on Child Enteric Protozoan Infections in Rural Bangladesh: A Cluster-Randomized Controlled Trial. *Clin Infect Dis*. 2018;67(10):1515–1522.

180. Ercumen A, Benjamin-Chung J, Arnold BF, et al. Effects of water, sanitation, handwashing and nutritional interventions on soil-transmitted helminth infections in young children: A cluster-randomized controlled trial in rural Bangladesh. *PLOS Neglected Tropical Diseases*. 2019;13(5):e0007323.

181. Ercumen A, Mertens A, Arnold BF, et al. Effects of Single and Combined Water, Sanitation and Handwashing Interventions on Fecal Contamination in the Domestic Environment: A Cluster-Randomized Controlled Trial in Rural Bangladesh. *Environ. Sci. Technol.* 2018;52(21):12078–12088.

182. Arnold BF, Null C, Luby SP, et al. Cluster-randomised controlled trials of individual and combined water, sanitation, hygiene and nutritional interventions in rural Bangladesh and Kenya: the WASH Benefits study design and rationale. *BMJ Open*. 2013;3(8):e003476–e003476.

183. Curtis VA, Danquah LO, Aunger RV. Planned, motivated and habitual hygiene behaviour: an eleven country review. *Health education research*. 2009;24(4):655–673.

184. Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction: an overview. *Food Nutr Bull*. 2015;36(1 Suppl):S76–S87.

185. Liu J, Gratz J, Amour C, et al. Optimization of quantitative PCR methods for enteropathogen detection. *PloS one*. 2016;11(6):e0158199.

186. Liu J, Platts-Mills JA, Juma J, et al. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. *The Lancet*. 2016;388(10051):1291–1301.

187. Platts-Mills JA, Liu J, Rogawski ET, et al. Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. *The Lancet Global Health*. 2018;6(12):e1309–e1318.

188. Kosek MN. Causal Pathways from Enteropathogens to Environmental Enteropathy: Findings from the MAL-ED Birth Cohort Study. *EBioMedicine*. 2017;18:109–117.

189. Amour C, Gratz J, Mduma E, et al. Epidemiology and Impact of Campylobacter Infection in Children in 8 Low-Resource Settings: Results From the MAL-ED Study. *Clin Infect Dis*. 2016;63(9):1171–1179.

190. Stephensen CB. Burden of Infection on Growth Failure. *The Journal of Nutrition*. 1999;129(2):534S-538S.

191. Cogill B. Anthropometric Indicators Measurement Guide. *Food and Nutrition Technical Assistance Project*. 2003;

192. de Onis M, Onyango AW, Van den Broeck J, et al. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull*. 2004;25(1 Suppl):S27-36.

193. Kuhn M. The caret Package. 2019;(https://topepo.github.io/caret/)

194. Fitzgerald C. Campylobacter. *Clin Lab Med*. 2015;35(2):289–298.

195. Rogawski ET, Liu J, Platts-Mills JA, et al. Use of quantitative molecular diagnostic methods to investigate the effect of enteropathogen infections on linear growth in children in low-resource settings: longitudinal analysis of results from the MAL-ED cohort study. *The Lancet Global Health*. 2018;6(12):e1319–e1328.

196. Young KT, Davis LM, DiRita VJ. Campylobacter jejuni: molecular biology and pathogenesis. *Nature Reviews Microbiology*. 2007;5(9):665–679.

197. Black RE, Levine MM, Clements ML, et al. Experimental Campylobacter jejuni infection in humans. *Journal of infectious diseases*. 1988;157(3):472–479.

198. Reti KL, Tymensen LD, Davis SP, et al. Campylobacter jejuni increases flagellar expression and adhesion of noninvasive Escherichia coli: effects on enterocytic Toll-like receptor 4 and CXCL-8 expression. *Infection and immunity*. 2015;83(12):4571–4581.

199. Masanta WO, Heimesaat MM, Bereswill S, et al. Modification of intestinal microbiota and its consequences for innate immune response in the pathogenesis of campylobacteriosis. *Clinical and Developmental Immunology*. 2013;2013.

200. Kalischuk LD, Leggett F, Inglis GD. Campylobacter jejuni induces transcytosis of commensal bacteria across the intestinal epithelium through M-like cells. Gut Pathog 2010; 2: 14; PMID: 21040540. 2010.

201. Riddle MS, Gutierrez RL, Verdu EF, et al. The chronic gastrointestinal consequences associated with Campylobacter. *Current gastroenterology reports*. 2012;14(5):395–405.

202. McCormick BJJ, Murray-Kolb LE, Lee GO, et al. Intestinal permeability and inflammation mediate the association between nutrient density of complementary foods and biochemical measures of micronutrient status in young children: results from the MAL-ED study. *Am J Clin Nutr*. 2019;110(4):1015–1025.

203. McCormick BJJ, Lee GO, Seidman JC, et al. Dynamics and Trends in Fecal Biomarkers of Gut Function in Children from 1-24 Months in the MAL-ED Study. *Am J Trop Med Hyg*. 2017;96(2):465–472.

204. Grant SS, Hung DT. Persistent bacterial infections, antibiotic tolerance, and the oxidative stress response. *Virulence*. 2013;4(4):273–283.

205. Gomez JE, McKinney JD. M. tuberculosis persistence, latency, and drug tolerance. *Tuberculosis*. 2004;84(1–2):29–44.

206. Monack DM, Mueller A, Falkow S. Persistent bacterial infections: the interface of the pathogen and the host immune system. *Nature Reviews Microbiology*. 2004;2(9):747–765.

207. Shonkoff JP, Boyce WT, Bush NR, et al. Translating the Biology of Adversity and Resilience Into New Measures for Pediatric Practice. *Pediatrics*. 2022;e2021054493.

208. Walker SP, Wachs TD, Grantham-McGregor S, et al. Inequality in early childhood: risk and protective factors for early child development. *Lancet*. 2011;378(9799):1325–1338.

209. Engle PL, Black MM, Behrman JR, et al. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet*. 2007;369(9557):229–242.

210. Condon EM. Chronic Stress in Children and Adolescents: A Review of Biomarkers for Use in Pediatric Research. *Biol Res Nurs*. 2018;20(5):473–496.

211. Johnson SB, Riley AW, Granger DA, et al. The science of early life toxic stress for pediatric practice and advocacy. *Pediatrics*. 2013;131(2):319–327.

212. Booth A, Carver K, Granger DA. Biosocial Perspectives on the Family. *Journal of Marriage and Family*. 2000;62(4):1018–1034.

213. van der Knaap LJ, Riese H, Hudziak JJ, et al. Glucocorticoid receptor gene (NR3C1) methylation following stressful events between birth and adolescence. The TRAILS study. *Transl Psychiatry*. 2014;4(4):e381.

214. Aschbacher K, O'Donovan A, Wolkowitz OM, et al. Good stress, bad stress and oxidative stress: insights from anticipatory cortisol reactivity. *Psychoneuroendocrinology*. 2013;38(9):1698–1708.

215. Montuschi P, Barnes PJ, Roberts LJ 2nd. Isoprostanes: markers and mediators of oxidative stress. *FASEB J*. 2004;18(15):1791–1800.

216. Pizzino G, Irrera N, Cucinotta M, et al. Oxidative Stress: Harms and Benefits for Human Health. *Oxid Med Cell Longev*. 2017;2017:8416763.

217. Il'yasova D, Spasojevic I, Wang F, et al. Urinary biomarkers of oxidative status in a clinical model of oxidative assault. *Cancer Epidemiol Biomarkers Prev*. 2010;19(6):1506–1510.

218. McEwen BS. Allostasis and allostatic load: implications for neuropsychopharmacology. *Neuropsychopharmacology*. 2000;22(2):108–124.

219. Brietzke E, Kauer Sant'anna M, Jackowski A, et al. Impact of childhood stress on psychopathology. *Braz J Psychiatry*. 2012;34(4):480–488.

220. Loman MM, Gunnar MR, Early Experience S and Neurobehavioral Development Center. Early experience and the development of stress reactivity and regulation in children. *Neurosci Biobehav Rev*. 2010;34(6):867–876.

221. Kagan J. The role of parents in children's psychological development. *Pediatrics*. 1999;104(1 Pt 2):164–167.

222. Kagan J, Reznick JS, Snidman N. The physiology and psychology of behavioral inhibition in children. *Child Dev*. 1987;58(6):1459–1473.

223. Gunnar MR, Fisher PA. Bringing basic research on early experience and stress neurobiology to bear on preventive interventions for neglected and maltreated children. *Development and psychopathology*. 2006;18(3):651–677.

224. Coe CL, Kramer M, Czéh B, et al. Prenatal stress diminishes neurogenesis in the dentate gyrus of juvenile rhesus monkeys. *Biol Psychiatry*. 2003;54(10):1025–1034.

225. Fernald LCH, Gunnar MR. Poverty-alleviation program participation and salivary cortisol in very low-income children. *Social Science & Medicine*. 2009;68(12):2180–2189.

226. Huth-Bocks AC, Levendosky AA, Bogat GA. The effects of domestic violence during pregnancy on maternal and infant health. *Violence Vict*. 2002;17(2):169–185.

227. Levendosky AA, Bogat GA, Martinez-Torteya C. PTSD symptoms in young children exposed to intimate partner violence. *Violence Against Women*. 2013;19(2):187–201.

228. Reilly EB, Gunnar MR. Neglect, HPA axis reactivity, and development. *Int J Dev Neurosci*. 2019;78:100–108.

229. Afifi T, Boman J, Fleisher W, et al. The relationship between child abuse, parental divorce, and lifetime mental disorders and suicidality in a nationally representative adult sample. *Child abuse & neglect*. 2009;33:139–47.

230. Fernald LC, Grantham-McGregor SM, Manandhar DS, et al. Salivary cortisol and heart rate in stunted and nonstunted Nepalese school children. *Eur J Clin Nutr*. 2003;57(11):1458–1465.

231. Fernald LC, Grantham-McGregor SM. Stress response in school-age children who have been growth retarded since early childhood. *Am J Clin Nutr*. 1998;68(3):691–698.

232. Labella MH, Eiden RD, Tabachnick AR, et al. Infant neurodevelopmental outcomes of prenatal opioid exposure and polysubstance use. *Neurotoxicol Teratol*. 2021;86:107000.

233. Schmidt KL, Merrill SM, Gill R, et al. Society to cell: How child poverty gets "Under the Skin" to influence child development and lifelong health. *Developmental Review*. 2021;61:100983.

234. Godoy LD, Rossignoli MT, Delfino-Pereira P, et al. A Comprehensive Overview on Stress Neurobiology: Basic Concepts and Clinical Implications. *Frontiers in Behavioral Neuroscience* [electronic article]. 2018;12. (https://www.frontiersin.org/article/10.3389/fnbeh.2018.00127)

235. Jiang S, Postovit L, Cattaneo A, et al. Epigenetic Modifications in Stress Response Genes Associated With Childhood Trauma. *Front Psychiatry*. 2019;10:808–808.

236. Wolf EJ, Miller MW, Sullivan DR, et al. A classical twin study of PTSD symptoms and resilience: Evidence for a single spectrum of vulnerability to traumatic stress. *Depression and anxiety*. 2018;35(2):132–139.

237. Amstadter AB, Moscati A, Maes HH, et al. Personality, cognitive/psychological traits and psychiatric resilience: A multivariate twin study. *Personality and Individual Differences*. 2016;91:74–79.

238. Asok A, Bernard K, Roth TL, et al. Parental responsiveness moderates the association between early-life stress and reduced telomere length. *Dev Psychopathol*. 2013;25(3):577–585.

239. Fisher PA, Stoolmiller M, Gunnar MR, et al. Effects of a therapeutic intervention for foster preschoolers on diurnal cortisol activity. *Psychoneuroendocrinology*. 2007;32(8):892–905.

240. Bernard K, Dozier M, Bick J, et al. Intervening to enhance cortisol regulation among children at risk for neglect: Results of a randomized clinical trial. *Development and Psychopathology*. 2015;27(3):829–841.

241. Tofail F, Fernald LC, Das KK, et al. Effect of water quality, sanitation, hand washing, and nutritional interventions on child development in rural Bangladesh (WASH Benefits Bangladesh): a cluster-randomised controlled trial. *The Lancet Child & Adolescent Health*. 2018;2(4):255–268.

242. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374(2065):20150202.

243. Granger DA, Kivlighan KT, Fortunato C, et al. Integration of salivary biomarkers into developmental and behaviorally-oriented research: problems and solutions for collecting specimens. *Physiol Behav*. 2007;92(4):583–590.

244. Granger DA, Kivlighan KT, el-Sheikh M, et al. Salivary alpha-amylase in biobehavioral research: recent developments and applications. *Ann N Y Acad Sci*. 2007;1098:122–44.

245. Edelman S, Shalev I, Uzefovsky F, et al. Epigenetic and genetic factors predict women's salivary cortisol following a threat to the social self. *PLoS One*. 2012;7(11):e48597.

246. Zhang TY, Labonté B, Wen XL, et al. Epigenetic Mechanisms for the Early Environmental Regulation of Hippocampal Glucocorticoid Receptor Gene Expression in Rodents and Humans. *Neuropsychopharmacology*. 2013;38(1):111–123.

247. Curtis MJ, Bond RA, Spina D, et al. Experimental design and analysis and their reporting: new guidance for publication in BJP. *British Journal of Pharmacology*. 2015;172(14):3461–3471.

248. Fagan TC, Conrad KA, Mayshar PV, et al. Single versus triplicate measurements of blood pressure and heart rate. *Hypertension*. 1988;11(3):282–284.

249. Rehman S, Nelson VL. Blood Pressure Measurement. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2022

250. Hamadani JD, Tofail F, Hilaly A, et al. Use of family care indicators and their relationship with child development in Bangladesh. *J Health Popul Nutr*. 2010;28(1):23–33.

251. Lin A, Mertens A, Li A, et al. WASH Benefits Bangladesh Child Stress and Child Development Analysis Plan. 2021;(osf.io/hzb6m)

252. Wood SN, Pya N, Säfken B. Smoothing Parameter and Model Selection for General Smooth Models. *Journal of the American Statistical Association*. 2016;111(516):1548–1563.

253. Lin A, Mertens AN, Arnold BF, et al. Telomere length is associated with growth in children in rural Bangladesh. *Elife*. 2021;10:e60389.

254. Hastie T, Tibshirani R. Generalized additive models for medical research. *Stat Methods Med Res*. 1995;4(3):187–196.

255. Nychka D. Bayesian Confidence Intervals for Smoothing Splines. *Journal of the American Statistical Association*. 1988;83(404):1134–1143.

256. Benjamini Y, Drai D, Elmer G, et al. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*. 2001;125(1–2):279–284.

257. Ferreira JA. The Benjamini-Hochberg method in the case of discrete test statistics. *Int J Biostat*. 2007;3(1):Article 11.

258. Gunnar MR, Vazquez DM. Low cortisol and a flattening of expected daytime rhythm: Potential indices of risk in human development. *Development and Psychopathology*. 2001;13(3):515–538.

259. Hostinar CE, Sullivan RM, Gunnar MR. Psychobiological mechanisms underlying the social buffering of the hypothalamic–pituitary–adrenocortical axis: A review of animal models and human studies across development. *Psychological Bulletin*. 2014;140(1):256–282.

260. Cicchetti D, Handley ED. Methylation of the glucocorticoid receptor gene, nuclear receptor subfamily 3, group C, member 1 (NR3C1), in maltreated and nonmaltreated children: Associations with behavioral undercontrol, emotional lability/negativity, and externalizing and internalizing symptoms. *Dev Psychopathol*. 2017;29(5):1795–1806.

261. Lin A, Mertens AN, Rahman MdZ, et al. Effects of drinking water, sanitation, handwashing and nutritional interventions on stress physiology, oxidative stress, and epigenetic

programming in young children living in rural Bangladesh: A randomized clinical trial. *medRxiv*. 2021;2021.11.24.21266798.