

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Causes and Consequences of Child Growth Failure in Low Resource Settings: A Targeted Learning Individual Participant Meta-analysis of 35 Cohorts and Trials

### Permalink

<https://escholarship.org/uc/item/9hx5n31f>

### Author

Mertens, Andrew

### Publication Date

2020

Peer reviewed|Thesis/dissertation

Causes and Consequences of Child Growth Failure in Low Resource Settings: A Targeted  
Learning Individual Participant Meta-analysis of 35 Cohorts and Trials

By

Andrew N Mertens

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 M. Colford, Jr., Chair

Professor Alan E. Hubbard

Professor Kirk R. Smith

Professor Benjamin F. Arnold

Spring 2020

Causes and Consequences of Child Growth Failure in Low Resource Settings: A Targeted  
Learning Individual Participant Meta-analysis of 35 Cohorts and Trials

Copyright 2020  
by  
Andrew N Mertens  
All rights reserved

## Abstract

# Causes and Consequences of Child Growth Failure in Low Resource Settings: A Targeted Learning Individual Participant Meta-analysis of 35 Cohorts and Trials by

Andrew N Mertens

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor John M. Colford, Jr., Chair

**Background:** Every child deserves an equal chance at good health and cognitive development, but growth failure from undernutrition and disease rob many poor children of this opportunity in the first 2 years of their life. Child malnutrition kills roughly three million children every year, and many more suffer long-term health consequences from impaired physical growth and cognitive development, especially if malnutrition occurs during the first 1000 days of life. Wasting and stunting, two markers of child malnutrition, are strongly associated with increased mortality and decreased cognitive development, and are still highly prevalent in developing countries, even though their reduction are targets of the World Health Organization Millennium Development Goals. Trials on nutritional interventions have had null results or not clinically meaningful effect sizes. Most existing research on exposures associated with stunting and wasting rely on small cohorts or large cross-sectional surveys focused on stunting and wasting prevalence. Pooled analyses of data from many longitudinal studies offer unique opportunities to identify novel patterns in child undernutrition across combinations of time, space, and human characteristics, which is not possible in smaller datasets. The objective of this research was to answer the following questions by pooling information across multiple cohorts: when and where growth failure occurs, what characteristics best identify children at risk, and how early growth failure affects later growth and mortality. Answering these questions may help inform generalizable rules for child monitoring and develop preventative intervention to stop the initial onset of wasting and stunting.

**Methods:** In this dissertation, I conduct an individual-participant-data meta-analysis of 35 different longitudinal cohorts and trials from 15 different countries to identify patterns in wasting incidence and recovery across child ages, across regions, and across seasons. I use machine learning and targeted learning methods to identify key child, parental, and household characteristics associated with wasting and stunting, adjusting for potential confounders. In addition to estimating relative risks of wasting between levels of characteristics, I estimate more policy-relevant population attributable fractions and a novel variable importance metric. In the first chapter I introduce the motivation for the analyses and give an overview of individual participant data meta-analysis methods and

advantages. In the second chapter I examine longitudinal patterns in wasting incidence, recovery, seasonality, and concurrence with stunting by using a subset of 18 monthly-measured cohorts, and I compare differences in the epidemiology of wasting across regions. In the third chapter I estimate the relative importance of the causes of early life growth failure using targeted learning methodologies. Using machine learning to flexibly adjust for measured confounders, I estimate and rank order associations between 36 child, parental, and household characteristics and measures of growth failure in 31 cohorts and trials. In the fourth chapter I estimate the associations between early growth failure and subsequent adverse outcomes, including mortality. In the fifth chapter I discuss key subject-area advancements in knowledge from these analyses and offer the public health implications.

**Significance:** This dissertation is currently the largest scale individual-participant data analysis of longitudinal patterns in and causes of child malnutrition in the first 1000 days of life. Using longitudinal, high quality anthropometry measurements, I examined patterns in wasting onset, recovery, seasonality, and concurrence with stunting not possible from national health surveys or accurate from small cohorts. The data were large in scope as well as scale, with data from 35 separate research sites (94,019 children and 645,869 total measurements), 18 with monthly measured cohorts (10,854 children and 187,215 anthropometry measurements). I was able to examine effect modification of wasting incidence by region and season. The results from chapter two found a high prevalence of wasting incidence at birth and the first three months of life, especially in South Asia. This is earlier than the paradigm in the nutrition field that most wasting episodes occur in children older than six months, after the cessation of exclusive breastfeeding and the start of crawling behavior. My findings suggest that interventions to prevent child acute malnutrition may have the largest impact if delivered earlier, including maternal interventions during pregnancy and early lactation and child interventions through the first months of life. Development of maternally focused, prenatal interventions represents an important opportunity to prevent child growth failure in lower-or-middle income countries (LMICs), especially as I found that early growth failure had a strong relationship with mortality. Future research into potential tradeoffs in prioritizing exclusive breast feeding versus the prevention or treatment of wasting in the first 6 months is needed. I also found high seasonality in wasting in synchrony with rainfall patterns, possibly due to seasonality in food availability and infectious diseases, indicating the timing of intervention delivery should target seasonal peaks in wasting. The strongest determinants of child growth failure were maternal anthropometry and child birth size. Future research should use the key determinants identified to find the optimal subgroups of children who would benefit most from postnatal interventions to start and reduce the substantial global health burden of child growth failure.

## Dedication

For Becky, I am so grateful for your continued support, advice, and inspiration.

# Table of contents

|                                                                                                       |     |
|-------------------------------------------------------------------------------------------------------|-----|
| Acknowledgments.....                                                                                  | v   |
| 1 Introduction.....                                                                                   | 1   |
| 1.1 Motivation.....                                                                                   | 1   |
| 1.2 Specific Aims.....                                                                                | 2   |
| 1.3 Individual-Participant-Data Meta-Analyses in the Field of Global Health.....                      | 2   |
| 2 Longitudinal analysis of child wasting and concurrence with stunting in low-resource settings ..... | 6   |
| 2.1 Background.....                                                                                   | 6   |
| 2.2 Objectives .....                                                                                  | 8   |
| 2.3 Methods.....                                                                                      | 8   |
| 2.4 Results.....                                                                                      | 12  |
| 2.5 Discussion.....                                                                                   | 24  |
| 3 Causes of child growth failure in low resource settings .....                                       | 26  |
| 3.1 Background.....                                                                                   | 26  |
| 3.2 Objectives .....                                                                                  | 27  |
| 3.3 Methods.....                                                                                      | 27  |
| 3.4 Results.....                                                                                      | 32  |
| 3.5 Discussion.....                                                                                   | 39  |
| 4 Consequences of child growth failure in low resource settings .....                                 | 41  |
| 4.1 Background.....                                                                                   | 41  |
| 4.2 Objectives .....                                                                                  | 42  |
| 4.3 Methods.....                                                                                      | 43  |
| 4.4 Results.....                                                                                      | 44  |
| 4.5 Discussion.....                                                                                   | 51  |
| 5 Conclusion .....                                                                                    | 54  |
| 5.1 Key Findings.....                                                                                 | 54  |
| 5.2 Scientific Conclusions .....                                                                      | 60  |
| References.....                                                                                       | 60  |
| Appendix A: Supplementary material for Chapter 2 .....                                                | 78  |
| Appendix B: Supplementary material for Chapter 3.....                                                 | 86  |
| Appendix C: Supplementary material for Chapter 4.....                                                 | 105 |
| Appendix D: Research and Methodological Lessons .....                                                 | 110 |

# List of figures

|                                                                                                                                                                                            |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 2.1   Summaries of included cohorts.....                                                                                                                                            | 13 |
| Figure 2.2   WLZ, Prevalence and incidence of wasting by age and region. ....                                                                                                              | 15 |
| Figure 2.3   Mean weight-for-length Z-scores by age and season. ....                                                                                                                       | 18 |
| Figure 2.4   Wasting incidence rate and recovery. ....                                                                                                                                     | 21 |
| Figure 2.5   Co-occurrence of wasting, stunting, and underweight.....                                                                                                                      | 23 |
| Figure 3.1   Cohort sample sizes and exposures measured. ....                                                                                                                              | 33 |
| Figure 3.2   Rank-ordered associations between child, parental, and household characteristics and population attributable differences in anthropometry Z-scores. ....                      | 35 |
| Figure 3.3   Effect of key exposures on the trajectories, timing, and severity of growth failure .....                                                                                     | 38 |
| Figure 4.1   Age at death and preceding anthropometry measurements among recorded deaths within the included KI cohorts.....                                                               | 45 |
| Figure 4.2   Adjusted differences in linear growth velocity (in centimeters) across 3-month age bands, by quartile of weight-for-length z-score (WLZ) in the preceding three months.....   | 46 |
| Figure 4.3   Relative risk of stunting onset after age 6 months between children who experienced measures of early wasting compared to children who did not experience early wasting. .... | 47 |
| Figure 4.4   Risks of persistent wasting due to early life growth failure.....                                                                                                             | 48 |
| Figure 4.5   Risks of concurrent wasting and stunting due to early life growth failure. ....                                                                                               | 49 |
| Figure 4.6   Risks of mortality due to early life growth failure.....                                                                                                                      | 50 |
| Figure 4.7   Risks of mortality by WLZ- or MUAC-defined wasting episodes.....                                                                                                              | 51 |
| Figure 4.8   Risks of mortality and serious growth failure due to early-life growth failure, stratified by region. ....                                                                    | 51 |
| Figure A.1   KI cohort selection.....                                                                                                                                                      | 78 |
| Figure A.2   Geographic location of KI cohorts.....                                                                                                                                        | 79 |
| Figure A.3   Comparison of cohort anthropometry to population-based samples.....                                                                                                           | 81 |
| Figure A.4   Prevalence of persistent wasting, severe wasting, and underweight by region.....                                                                                              | 81 |
| Figure A.5   Month of birth affects seasonal patterns in WLZ .....                                                                                                                         | 83 |
| Figure A.6   Comparison of underweight and stunting prevalence at birth with and without gestational age correction .....                                                                  | 84 |
| Figure B.1   Example forest plot of cohort-specific and pooled parameter estimates.....                                                                                                    | 87 |
| Figure B.2   Heatmap of significance and direction across exposure-outcome combinations.....                                                                                               | 89 |
| Figure B.3   Age-stratified population attributable differences in length-for-age Z-scores.....                                                                                            | 91 |
| Figure B.4   Age-stratified population attributable differences in weight-for-length Z-scores.....                                                                                         | 93 |
| Figure B.5   Mediation by birth characteristics of associations between parental anthropometry and child Z-scores at 24 months. ....                                                       | 93 |

|                                                                                                                                                                                                                |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure B.6   Rank-ordered associations between child, parental, and household characteristics and population attributable fractions of stunting and wasting. ....                                              | 94  |
| Figure B.7   Regionally-stratified population attributable differences in length-for-age Z-scores. ....                                                                                                        | 96  |
| Figure B.8   Regionally-stratified population attributable differences in weight-for-length Z-scores.....                                                                                                      | 97  |
| Figure B.9   Difference between adjusted and unadjusted Z-score effects by number of selected adjustment variables.....                                                                                        | 98  |
| Figure B.10   Assessing sensitivity of estimates to unmeasured confounding using E-values .....                                                                                                                | 100 |
| Figure C.1   Risks of subsequent linear-growth failure across categories of prior WLZ.....                                                                                                                     | 105 |
| Figure C.2   Risks of mortality due to growth failure before 24 months. ....                                                                                                                                   | 106 |
| Figure C.3   Unadjusted risks of mortality due to growth failure before 6 months.....                                                                                                                          | 106 |
| Figure C.4   Sensitivity analysis on the effect of excluding children without anthropometry measures in the first month of life on estimated associations between growth failure under 6mo and mortality ..... | 107 |
| Figure D.1   Anthropometry measurement quality checking of KI individual-participant datasets .....                                                                                                            | 114 |
| Figure D.2   Comparing population attributable differences estimated using targeted learning using either a one-step or two-step approach.....                                                                 | 121 |

## List of tables

|                                                                                   |     |
|-----------------------------------------------------------------------------------|-----|
| Table A.1  Summary of KI cohorts used in Chapter 2.....                           | 84  |
| Table B.1  Summary of KI cohorts used in Chapter 3 .....                          | 100 |
| Table B.2  List of potential causes of growth failure analyzed in Chapter 3 ..... | 102 |
| Table C.1  Summary of KI cohorts used in Chapter 4.....                           | 108 |

# Acknowledgments

I would first like to acknowledge my advisor, Jack Colford, for his consistent mentoring and support throughout my graduate training, and for pushing me forward at each step of my career. Jack's accessibility, insightfulness, advice, and support have been critical to my development as an epidemiologist. He has and continues to provide me with numerous opportunities to build my skillset, grow my professional network, become a mentor myself, and challenge myself in new research inquiries. He is truly a one of a kind mentor. I look forward to a long and fruitful research and teaching collaboration through my post-doctoral career.

I am also very grateful to Ben Arnold for mentoring me through every step of my graduate career. I have learned so much from you and every piece of this dissertation is a product of your guidance and teaching. Even as I step into independence, I hope to continue learning from you through a career-long research partnership.

I want to thank Jade Benjamin-Chung for her brilliant contributions in developing this research as well as her continued leadership in best practices in research transparency and analysis reproducibility.

Thank you to my other committee members, Alan Hubbard and Kirk Smith, for their valuable input on this dissertation. Thank you to Jeremy Coyle for continued statistical support, which I would be lost in a maze of R errors without.

I would also like to thank all the primary investigators of the studies analyzed for this dissertation, as well as the hundreds of investigators involved in study design, data collection, and data sharing. Without your generous sharing of data, as well as feedback on this research, this dissertation would not have been possible. Thank you to Parul Christian, Ken Brown, Thea Norman, Vishak Subramoney, and Jonas Häggström for their partnership through the Bill and Melinda Gates Foundation Knowledge Integration effort, their continued support in pushing forward this research effort, and their guidance in developing impactful research.

Lastly, thank you to my UC Berkeley peers who have assisted in all facets of this work: Oleg Sofrygin, Wilson Cai, Anna Nguyen, Nolan Pokpongkiat, Wendy Jilek, Esther Chung, Sonali Dayal, Nima Hejazi, Ivana Malenica, and Haodong Li.

# 1 Introduction

## 1.1 Motivation

Undernutrition contributes to 45% of deaths in children under 5 years of age according to Black et al. in 2013.<sup>1</sup> Stunting is an indicator of impaired linear growth from chronic undernutrition and is associated with increased mortality and decreased cognitive development and educational attainment.<sup>2</sup> Wasting, or dangerously low weight for a given height, results from a loss of muscle and fat tissue from acute malnutrition or infection, and is associated with an even higher increase in mortality.<sup>3,4</sup> Wasting and stunting may also be key developmental origins of adult disease.<sup>2,3,4</sup> Worldwide, an estimated 7.5% of children under 5 are wasted and 22.2% are stunted, primarily in the developing world. Wasting prevalence has not decreased over the last 30 years, and reductions in stunting to date are insufficient to meet the WHO 2025 target.<sup>5,6</sup>

Randomized trials on nutritional interventions on stunting and wasting have shown inconsistent and small effects.<sup>7</sup> Elucidating the characteristics of *who* become stunted and wasted, *when* and *where* onset occurs, and *what* interventions are effective in *which* children is critical to reducing stunting and wasting. To answer these questions precisely, sufficient data are needed to identify fine-grained patterns in wasting and stunting across children. Most prior research is either from cross-sectional surveys, where the timing of onset or recovery cannot be determined, or from small cohorts with insufficient data to examine stunting or wasting at specific ages or within important subgroups. I conducted an individual-participant data (IPD) meta-analysis from a global set of cohorts and trials to elucidate patterns in wasting across regions, child ages, and season, to detect exposures associated with stunting and wasting. I used data from 108,336 children from 35 cohorts or trials from lower-or-middle income (LMIC) countries, conducted between 1969 and 2017, and aggregated for the Bill & Melinda Gates Foundation's Knowledge Integration (KI) project.<sup>8</sup> The Bill & Melinda Gates Foundation launched the KI effort to collect study data from around the world and analyze the studies to answer key questions related to child growth and development, including who experiences growth failure, when and where onset occurs, and what characteristics best identify children at risk. All the studies have repeated anthropometry measurements from birth to age 2 years, and measure child, parental, and household characteristics potentially associated with poor growth.

A goal of the KI effort is to define precise intervention strategies to prevent preterm birth, growth faltering, wasting, and poor child development. The aggregated set of IPD is, to my knowledge, the largest longitudinal dataset on child growth and development in low-resource settings ever assembled. Similar studies include fewer children,<sup>9–11</sup> used cross-sectional data,<sup>10</sup> or are meta-analyses of published effects.<sup>12,13</sup> The KI data offer a unique opportunity to calculate the incidence of wasting, wasting recovery, and concurrent wasting and stunting within narrow age intervals, and still have sufficient power to examine differences across geographic regions. The IPD also included a large set of variables I was able to harmonize across studies, including gender, birth size, measures of SES, breastfeeding practice, food security, and diarrheal disease. Access to individual participant data allows me to estimate the effect of exposures within subgroups with consistent definitions across studies, which is not possible in a traditional meta-analysis.<sup>14</sup>

This dissertation is the largest longitudinal analyses of stunting and wasting to my knowledge, with standardized exposure and outcome definitions to reduce bias and sufficient data to analyze rare subgroups. An exploration of the timing and location of wasting episodes helps inform when and where intervention efforts can have the most impact, and a systematic analysis of conditions that precede episode onset helps target preventative interventions to the children at highest risk. In the resource-constrained LMIC public health sector, targeting interventions to children who benefit most will maximize health improvements.

## 1.2 Specific Aims

In this dissertation, I conduct an individual-participant data meta-analysis of longitudinal studies of child growth in low-resource settings to better understand the patterns, causes, and consequences of child wasting and stunting. My specific aims are as follows:

1. To explore regional, seasonal, and age-specific patterns in the epidemiology of wasting (incidence, prevalence, and recovery) in the first 24 months of children's lives (Chapter 2).
2. To apply targeted learning to estimate and rank the relative contributions of child, parental, and household characteristics to measures of child growth failure (Chapter 3).
3. To apply targeted learning to estimate the risk of mortality and serious growth failure across children with different measures of growth failure (Chapter 4).

## 1.3 Individual-Participant-Data Meta-Analyses in the Field of Global Health

Here, I report results on a literature review of individual-participant data (IPD) meta-analysis methods and applications to global health research, prior to conducting an IPD meta-analysis of KI data to answer my specific aims. IPD meta-analyses, where original participant-level data from all included studies is aggregated and re-analyzed, is considered the gold standard of meta-analyses.<sup>15</sup> While IPD analyses require more effort than traditional aggregate data (AD) meta-analyses, which extract reported summary measures from published research, they have several advantages, including the ability to standardize modeling approaches across studies, and examine effects in consistently defined subgroups. As sharing of scientific data becomes more common,<sup>16</sup> the number of IPD analyses published each year has risen from 5 or less in the early 1990's to an average of 29 a year from 2005-2009<sup>17</sup>, with an average increase of 3.7 additional IPD analyses per year from 1987-2012.<sup>18</sup> As sharing of data and IPD analyses become more common, and if their results are to be held up as the strongest evidence to answer a scientific question, advancements in IPD meta-analysis methods and data-sharing standards are becoming increasingly important.<sup>17,19</sup>

This is especially true in two research domains covered in this dissertation: observational analyses and global health research. Of 829 articles on IPD analyses published before August 2012, 13.8% were on observational studies.<sup>18</sup> This may be because it is easier to code and

analyze randomized control trial data in a standardized way than it is to harmonize data from disparate observational studies into comparable exposures and outcomes. Additionally, while IPD analyses, by definition, involve data from numerous sites and predominately from multiple countries, most have focused on high-income populations. IPD analyses conducted in low-resource settings are primarily on HIV,<sup>20–32</sup> with a few on tuberculosis<sup>33–36</sup> and malaria<sup>37–40</sup>, and two on neglected tropical diseases.<sup>41,42</sup> Among nutrition IPD analyses, one was on zinc supplementations effect on child diarrhea,<sup>43</sup> one on formula supplementation's effect on child cognitive development,<sup>44</sup> one on maternal vitamin B12 supplementation's effect on infant birthweight,<sup>45</sup> one on the effect of fetal growth restriction and preterm birth on later child growth failure,<sup>46</sup> and one on the associations between child growth failure and mortality.<sup>9</sup>

As with a traditional AD meta-analyses, the goal of an IPD meta-analysis is to answer a particular research question with more precise estimates of effect sizes through a weighted pooling of estimates from a set of studies designed to answer the research question, which increases statistical power and generalizability. Generally, the process of conducting an IPD meta-analyses begins the same as a traditional AD meta-analysis. The investigator searches the literature using pre-specified search terms and identifies all published studies that fit the selection criteria. In an AD meta-analysis, the investigator extracts the primary estimates of interest, often treatment effects, from the selected publications, while in an IPD meta-analysis the investigator contacts the authors of the selected publications and requests the raw, individual-level data. Ideally, the investigator will receive data from all selected studies, but realistically a subset of authors will respond and share data. The investigator will then analyze the contributed data.

The individual-participant dataset analyzed in chapters 2-4 of this dissertation differs from most IPD meta-analyses in that it was assembled before the research questions were finalized. Instead of identifying a research question and contacting the authors of publications on the question of interest, the KI effort aggregated public and Bill and Melinda Gates Foundation-funded studies on child growth and development, harmonized the data, and made the data available to research partners to analyze to better understand three complex and interrelated outcomes: preterm birth, growth failure and impaired neurocognitive development. The data varied, from trials, to country-level Demographic and Health Surveys, to historic and recent cohorts, and studies were from both high- and low-income settings. I used a subset of longitudinal, frequently measured studies of children's anthropometry from birth to two years of age that were conducted in low and middle income countries collected by the KI effort. This approach reduced the data gathering and harmonization effort needed to create the individual-participant dataset used in this dissertation, but there may be public datasets, or data the primary investigator would have agreed to share, that met my inclusion criteria not included in the KI data. This IPD meta-analysis approach will likely become more common as data-sharing increases and more large, multi-study datasets are available as a public asset used to answer multiple research questions. The KI IPD analysis was a large IPD meta-analysis with 35 studies included, compared to the median IPD meta-analysis size of 8 studies.<sup>18</sup>

Estimates from IPD analyses and AD meta-analysis will match if the underlying data used are the same and if the modeling approaches to estimate the treatment effects are equivalent. Discrepancies usually arise because IPD data sets include different data than AD, either if different patients are excluded, if different variable cleaning rules are used in coding the

raw data, or if IPD analyses use a more appropriate effect measure (such as a risk ratio instead of an odds ratio). In a Cochrane systematic review of IPD analyses, the precision and size of effect varies compared to equivalent AD estimates, but with no systematic pattern.<sup>47</sup> However, in a simulation study IPD analyses consistently had more power than equivalent AD analyses.<sup>48</sup> The relative advantage of using an IPD analysis over an AD analysis will depend on the disease, treatment of interest, clinical question, and data availability.<sup>49</sup>

The primary advantage of IPD analyses is the ability to standardize subgroups across studies to uncover heterogeneity in treatment effects. Subgroup effects can only be estimated in AD analyses if all studies report them with standardized subgroup definitions (such as effects of treatment, stratified by gender). Alternatively, effect modification can be assessed in AD analyses with meta-regressions, but only if the modifier is continuous, and the mean of the modifier in populations used in published analyses varies.<sup>50</sup> However, the meta-regression approach has low power and is vulnerable to ecological fallacy, while IPD analyses allow more unbiased estimates of effect modification through more precise and standardized subgroup definitions. An IPD meta-analysis of 5 RCTs of anti-lymphocyte antibody induction therapy for renal transplant patients found no difference in treatment effect in patients with elevated antibodies compared to non-elevated when compared across studies using an AD approach (i.e. no differences in trials where mean antibody levels were higher compared to trials where mean antibody levels were lower).<sup>51</sup> However, when comparing subgroups within trials, treatment was significantly more effective in the high antibody subgroups. Additionally, a simulation found IPD analyses had a 90.8% power to detect true subgroup interactions, while meta-regression using AD estimates had a 10.8% power to detect interactions.<sup>48</sup> Therefore, IPD analysis reduces any ecological bias that may be present when using aggregated measures reported in trial publications.<sup>51</sup> The large size of an individual-participant dataset also allows intervention effects to be identified in rarer subgroups, a first step towards more individualized medicine. While IPD will not necessarily include enough observations to be “big data” in a computational sense, by definition, it will be big data for the scientific field it is conducted within as it is an aggregation of studies of normal size for the field.<sup>52</sup> Beyond the precise intervention effects that can be estimated from an IPD meta-analysis of trials, the big-data nature of IPD allows estimations of causal effects from observational data that emulates a trial when a trial is not feasible.<sup>53</sup>

Along with allowing analyses with consistently defined subgroups, IPD meta-analyses allow for more comparable estimates across studies than AD meta-analyses. With subject-level data, outcome and covariate information can be coded in a standardized way, and the same models can be used to estimate parameters of interest across all studies. While differences in analytical approach may not matter in unadjusted treatment effect estimated from well-randomized trials, if the analysis is observational, or confounding is present, or the analysis is less straightforward (e.g. time-to-event analyses), differing modeling approaches may bias AD meta-analyses.<sup>49</sup> Access to IPD allows estimates from each study to be adjusted for a standard set of confounders, like subject age and sex, or for all measured potential confounders within each study, which is the approach I took in Chapters 3 and 4. If data are longitudinal, as in the KI analyses, IPD also allows analysis within standardized, narrow bands of time or subject age that would not be reported in all published studies. IPD analyses also allow investigators to include patients who were excluded in the original studies. In an IPD meta-analysis in soft tissue sarcoma, 99% of the 344 patients that had been excluded from published individual trial analyses

were included in the shared individual-participant datasets, and the meta-analysis estimates were no longer significant after reinstating all exclusions.<sup>54</sup> Missing data can be observed and accounted for at the individual level, and multiple imputation or other methods to account for missing exposure or covariate information.<sup>55,56</sup> Crucially, missing data should only be imputed within studies, as imputation across studies may lead to bias because associations between covariates likely vary across different study populations.<sup>55</sup>

IPD analyses can also reduce outcome reporting bias, where published literature only presents results for a subset of outcomes where the treatment effect was significant. Thirty-four percent of 283 Cochrane reviews had at least one trial where there was evidence of outcome reporting bias, and of 42 significant meta-analyses, 8 would not have remained significant had all outcomes been reported. Also, IPD analyses can ensure that outcomes are standardized across studies and allows the IPD investigator to check the relative quality of contributed datasets. In the analyses in Chapter 2-4 of this dissertation, children's growth measures were standardized across all studies using the 2006 World Health Organization growth standards.<sup>57</sup>

The primary disadvantage of an IPD meta-analysis is the amount of time, money, and research complexity required to contact study authors, receive, check, understand, and harmonize contributed dataset, and re-analyze study data rather than just extracting published AD estimates. If the key question of interest is the population-level effect of an intervention on a standardized outcome measurement, and published estimates are from high quality trial, and subgroup estimates are not of interest, the time required to conduct an IPD analysis may be excessive for minimal gains in scientific validity.<sup>58</sup> Another serious disadvantage of IPD analyses is the possibility of bias if the studies shared with the IPD investigator are systematically different from datasets that were requested but not shared. Similar to publication bias in AD meta-analyses, "data contribution" bias is likely to occur if null studies are harder to find, or if investigators are more likely to share results from studies with significant treatment effects.

As IPD meta-analyses require all the work of a traditional meta-analysis to find and understand relevant studies, but also require data aggregation, standardization, and analysis, some have argued that IPD is not worth the time investment,<sup>58</sup> and IPD analyses often find similar primary results as AD analyses.<sup>59</sup> Using the example of the Bill and Melinda Gates Foundation Knowledge Integration (KI) program and the analyses presented in chapters 2-4 of this dissertation, I lay out applied examples of the challenges and promise of individual participant data analyses. In the appendix of this dissertation, I discuss the often overlooked, unglamorous, and yet critical details of the data harmonization process, as well as the novel application of targeted estimation using machine learning to IPD analyses.<sup>60</sup> Although time and resource intensive, IPD analyses allow estimation of important subgroup effects, can lead to public big-data resources, and can be used as an opportunity to externally replicate published results.

## 2 Longitudinal analysis of child wasting and concurrence with stunting in low-resource settings

### Preface

The work presented in this chapter is the product of myself and a large number of collaborators, and we plan on publishing this work in a similar form after submission of this dissertation, with myself as first author. The Bill and Melinda Gates Foundation (BMGF) funded a University of California, Berkeley team led by Professors John Colford, Alan Hubbard, and Mark van der Laan to investigate longitudinal patterns in child growth using the KI data. The research questions and analysis plan for this chapter was developed by me, Benjamin Arnold (UC San Francisco), Jade Benjamin Chung (UC Berkeley), John Colford, Kenneth H. Brown (BMGF), and Parul Christian (BMGF). The following UC Berkeley post-doctoral researchers, graduate students, and undergraduate students assisted in researching the included studies, statistical coding, data analysis, and figure creation: Jeremy Coyle, Oleg Sofrygin, Wilson Cai, Anna Nguyen, Nolan Pokpongkiat, Wendy Jilek, Esther O Chung, Sonali Dayal, Nima Hejazi, Ivana Malenica, and Haodong Li. Ryan Hafen (BMGF) provided visualization suggestions and Vishak Subramoney (BMGF) assisted in data harmonization. Thea Norman and Jonas Häggström (BMGF) assisted in coordinating research with BMGF subject-area experts and securing data-use agreements with primary investigators of studies contributed to KI. Benjamin Arnold will be senior author on the planned publication, and had the largest contribution to developing the research methods and drafting the manuscript after myself.

### 2.1 Background

Undernutrition is estimated to contribute to 45% of deaths in children under 5 years of age.<sup>61,1</sup> Wasting, or dangerously low weight for a given length, is a form of undernutrition that results from a loss of muscle and fat tissue from acute malnutrition or infection.<sup>3,4,62</sup> An estimated 52 million children worldwide (7.7% of children under 5 years) are wasted, with over half living in South Asia; and prevalence has not decreased in 30 years.<sup>5,6</sup> Wasting in early childhood, defined as a weight-for-length Z-score (WLZ)  $< -2$ , is a serious condition, as children with less tissue mass have weakened immune systems and are vulnerable to disease, leading to an increased mortality risk, especially if the wasting is severe (WLZ  $< -3$ ).<sup>1,3,4,9,63</sup> Severe wasting can be successfully treated using nutrition rehabilitation with ready-to-use-therapeutic foods and appropriate antimicrobial agents,<sup>64,65</sup> but relapse and mortality continue to be high. Also, randomized trials have shown inconsistent and small effects of nutritional interventions in preventing moderate wasting.<sup>7</sup> Although reducing mortality from wasting through treatment is impactful and cost effective,<sup>64,66</sup> prevention is important as children who recover from wasting have a higher risk of poor long-term growth and neurocognitive development.<sup>1,67,68</sup> Wasting may also be a risk factor for stunting.<sup>63,69,70</sup> Research and programs often consider wasting and stunting separately, but similar risk factors underlie each

condition,<sup>71,72,73</sup> wasting may reduce linear growth,<sup>74</sup> and concurrent wasting and stunting heightens mortality risk far beyond the mortality risk of either condition in isolation.<sup>3,9,63,75</sup> The failure to implement large scale programs to prevent wasting may reflect a historic focus of programs and research on the treatment of severe wasting rather than prevention, or an incomplete understanding of the optimal time to intervene.

Wasting is thought to occur primarily during the critical period for adequate, appropriate and safe complementary feeding from ages 6-24 months.<sup>76,77</sup> Episodes of illness including diarrhea and pneumonia may be followed by wasting or periods of food insecurity could cause more wasting, although a better understanding of the pathways to wasting and therefore the optimal ages/times at which to intervene is needed.<sup>78</sup> A more complete understanding of the epidemiology of wasting and how it varies by age is key to develop and target preventive interventions.<sup>7,79,13,80,81</sup>

Unlike the cumulative process of linear growth failure leading to stunting, wasting varies considerably over time both within-individuals and within-populations.<sup>70,82</sup> This dynamic nature means that the number of wasting episodes within a year is poorly captured in cross-sectional, survey-based estimates of wasting: as many as 13 times the number of children who are wasted at one point in time (prevalence) may experience a period of being wasted within a year (incidence).<sup>83</sup> Wasting varies within populations in relation to seasonal changes in food insecurity or disease,<sup>79,84–86</sup> but not consistently between countries; the wasting peaks were not always in the hungry season<sup>78</sup> and there may be an interaction between season and child age.<sup>87</sup> The highly seasonal nature of wasting means surveys of prevalence taken at different times have limited comparability. A longitudinal analysis could elucidate the optimal timing of intervention delivery to prevent seasonal peaks.<sup>78</sup> Overall, there is limited research about incident episodes of wasting,<sup>1,6</sup> and longitudinal anthropometry data from multiple cohorts gives a unique opportunity to examine the life-course of wasting episodes across a wide range of populations.

Here, I report a pooled analysis of 18 longitudinal cohorts from 10 low- and middle-income countries (LMICs) in South Asia, Sub-Saharan Africa, and Latin America, that measured length and weight monthly among children 0-24 months of age. My primary objectives were to produce the first large-scale estimates of wasting incidence and recovery and to quantify the temporal and regional variation. I also assessed the concurrence of wasting and stunting and compared estimates of wasting prevalence and incidence. I hypothesized that a careful examination of wasting incidence and recovery would provide new insights into the timing and burden of wasting and that incidence-based summaries of wasting could differ from summaries based on prevalence.

To test these hypotheses, I analyzed data from the Bill & Melinda Gates Foundation's Knowledge Integration (KI) database, which includes observations on millions of participants from varied studies on child growth and development. I selected longitudinal cohorts for analysis that met 5 criteria: 1) conducted in low- or middle-income countries; 2) enrolled children between birth and age 24 months and measured their length and weight repeatedly over time; 3) did not restrict enrollment to acutely ill children; 4) enrolled at least 200 children; 5) collected anthropometry measurements at least monthly (Figure A.1). Inclusion criteria were defined to select cohorts representative of general populations in low- and middle-income countries

(LMICs) with sufficiently frequent measurements to investigate the acute, dynamic nature of wasting.

## 2.2 Objectives

- 1) Estimate if patterns in wasting vary across different regions and child ages, and whether these patterns vary if measuring wasting prevalence or wasting incidence.
- 2) Estimate the degree and timing of wasting concurrence with child stunting.
- 3) Estimate if seasonal trends in rainfall are associated with child wasting.

## 2.3 Methods

### Study designs and inclusion criteria

I included all longitudinal observational studies and randomized trials available through the KI project on April 2018 that met five inclusion criteria (Figure A.1): 1) conducted in low- or middle-income countries; 2) enrolled children between birth and age 24 months and measured their length and weight repeatedly over time; 3) did not restrict enrollment to acutely ill children; 4) enrolled at least 200 children; 5) collected anthropometry measurements at least monthly. The frequency of measurements was assessed by calculating the median days between measurements. The pre-specified analysis protocol stipulated that if randomized trials found effects of interventions on growth, the analysis would only include the control arm only; yet, all intervention trials that met the inclusion criteria had null effects on growth, so all arms were included. I included all children under 24 months of age, assuming months were 30.4167 days. I excluded extreme measurements of WLZ  $> 5$  or  $< -5$ , of height-for-age Z-scores (HAZ)  $< -6$  or  $> 6$ , and of weight-for-age Z-scores (WAZ)  $< -6$  or  $> 5$ , consistent with 2006 WHO Growth Standards recommendations.<sup>88</sup> I checked for cohort-wide anthropometry measurement quality by plotting Z-score densities and calculating the proportion of length measurements where length decreased beyond the expected technical error of measurement compared to the last measurement on a child. One cohort, MAL-ED Pakistan, was excluded because measurements exhibited a multimodal WLZ distribution with scores binned at  $-2$ ,  $-1.5$ , and  $-1$  instead of continuously distributed.

### Outcome definitions

I used the following outcome measures in the analysis:

Weight-for-Length Z-scores were calculated using the 2006 WHO growth standards,<sup>57</sup> and mean WLZ was calculated within strata of interest. I used the medians of triplicate measurements of lengths and weights of children from pre-2006 cohorts to re-calculate Z scores to the 2006 standard.

Prevalent wasting was defined as the proportion of measurements within a specific stratum (e.g., age) below the 2006 WHO standard  $-2$  WLZ, and analogously below the 2006 WHO standard  $-3$  WLZ for severe wasting. For each age, I included children with WLZ measurements within one month before and after that age in the point prevalence estimate (e.g., for point prevalence at 6 months, I include children aged 5-7 months).

Incident wasting episodes were defined as a change in WLZ from above  $-2$  Z in the prior measurement to below  $-2$  Z in the current measurement. Similarly, I defined severe wasting episodes using  $-3$  Z cutoff. I assumed a 60-day washout period between episodes of wasting before a new episode of wasting could occur (Figure 2.4a). Children were considered at risk for wasting at birth, so children born wasted were considered to have an incident episode of wasting at birth. Children were also assumed to be at risk of wasting at the first measurement in studies that enrolled children after birth, so children wasted at the first measurement in a non-birth cohort were assumed to have incident wasting occurring at the age halfway between birth and the first measurement.

Incidence proportion of wasting was calculated during a defined age range (e.g. 6-12 months) as the proportion of children at risk who became wasted during the age period (the onset of new episodes).

Recovery from wasting was defined as a WLZ change from below to above  $-2$  Z among children who are currently wasted or severely wasted. I required a child to maintain WLZ above  $-2$  for 60 days to be considered “recovered”. Children were only considered “at risk” for recovery if their prior measurement was below  $-2$  WLZ. I measured the proportion of children who recover from moderate wasting ( $WLZ < -2$ ) within 30, 60, and 90 days of the onset of the episode. I assumed recovery from wasting was spontaneous because I did not have consistent information on referral guidelines across cohorts in the analysis, but some children with moderate wasting may have been referred to clinical facilities for outpatient treatment.

Wasting duration was estimated by counting the days between the onset of wasting and recovery within an individual child. I assumed that the episode started or ended at the midpoint between measurements. For example, if a child was not wasted at age 40 days, wasted at age 70 days, and not wasted at age 100 days, the duration of the wasting episode was:  $(70-40)/2 + (100-70)/2 = 30$  days. I calculated the median duration of wasting episodes among cohorts where children recovered during the study period from  $>50\%$  of observed wasting episodes, as the duration of episodes that extended beyond the study period is unknown. Therefore, the median duration could only be calculated when less than half of episodes were censored. Confidence intervals were calculated using the quantile method.

Incidence rate of wasting was calculated during defined age ranges as the number of incident episodes of wasting per 1,000 child-days at risk during the age range. Children were considered “at risk” for incident wasting episodes if they were not currently wasted and were classified as recovered from any prior wasting episode (beyond the 60-day washout period). Therefore, wasted children, or children within the washout period, did not contribute to the person-time at risk for wasting used to calculate wasting incidence. To calculate person-time at risk of wasting, I assumed that the onset of a wasting episode occurred when the child’s age was at the midpoint between the measurement of  $WLZ \geq -2$  and the measurement of  $WLZ < -2$ , and conversely the time of recovery occurred when the child’s age was at the midpoint between the last measurement of  $WLZ < -2$  and the first measurement of  $WLZ \geq -2$  within the 60-day washout period (Figure 2.4a).

Persistent wasting was defined as a  $>50\%$  longitudinal prevalence of below  $-2$  WLZ (wasting), and analogously  $>50\%$  longitudinal prevalence of below  $-3$  WLZ for persistent severe wasting.<sup>69</sup>

Concurrent prevalence of wasting and stunting was defined as the proportion of measurements at a specific age when a child was both wasted and stunted at the same measurement. For each age, I included children with WLZ and LAZ measurements within one month before and after that age in the point prevalence estimate (e.g., for point prevalence at 6 months, I included children aged 5-7 months).

Prevalent underweight was defined as the proportion of measurements at a specific age below the 2006 WHO standard  $-2$  WAZ. For each age, I included children with WAZ measurements within one month before and after that age in the point prevalence estimate (e.g., for point prevalence at 6 months, I included children aged 5-7 months).

### Subgroups of interest

I stratified the above outcomes of interest within the following subgroups: Child age, grouped into one, three, or six month intervals, (depending on the outcome); the region of the world (Asia, sub-Saharan Africa, Latin America); the month of the year, and the combinations of the above categories.

### Fixed and random effects models

I estimated wasting outcomes within specific cohorts and pooled estimates within each age strata using random effects models. I estimated overall pooled effects, and pooled estimates specific to South Asian, African, or Latin American cohorts. I pooled all statistics using random effects models. This modeling approach assumes that studies are randomly drawn from a hypothetical population of longitudinal studies that could have been conducted on children's linear growth in the past or future.

Random effects models assume that the true population outcomes  $\theta$  are normally distributed with mean  $\mu$  and variance  $\tau^2$  (i.e., that  $\theta \sim N(\mu, \tau^2)$ ). To estimate outcomes in this study, the random effects model is defined as follows for each study in the set of  $i = 1, \dots, k$  studies:

$$y_i = \mu + u_i + e_i$$

where  $y_i$  is the observed outcome in study  $i$ ,  $u_i$  is the random effect for study  $i$ , and  $e_i$  is the estimated outcome for study  $i$ , and  $e_i$  is the sampling error within study  $i$ . The model assumes that  $u_i \sim N(0, \tau^2)$  and  $e_i \sim N(0, v_i)$ , where  $v_i$  is the study-specific sampling variance. I fit random effects models using the restricted maximum likelihood estimator.<sup>89,90</sup> If a model failed to converge, models were fit using a maximum likelihood estimator instead. The estimate of  $\mu$  is the estimated mean outcome in the hypothetical population of studies on this topic (i.e., the estimated outcome pooling across study cohorts). I pooled binary outcomes on the logit scale and then back-transformed estimates after pooling to constrain confidence intervals between 0 and 1.

### Seasonality analysis

I compared mean WLZ over day of the year, over child birth day, and over seasons defined by rainfall. I estimated mean WLZ by cohort over day of the year and child birth day using cubic splines (Figure 2.3a, c).<sup>91</sup> Splines of WLZ over day of the year were plotted over country-level monthly mean rainfall averaged over the years a study measured child anthropometry below ages 24 months (Figure 2.3a). Rainfall data was extracted from the University of East Anglia Climate Research Unit CRU TS dataset (version 3.23, 23 October 2015).<sup>92</sup>

The season of peak rainfall was defined as the three-month period with the highest mean rainfall. Mean differences in WLZ between three-month quarters was estimated using linear regression models. I compared the consecutive three months of the maximum average rainfall over the study period, as well as the three months prior and the three months after the maximum-rainfall period, to a reference level of the three months opposite the calendar year of the maximum-rainfall period. I used all WLZ measurements of children under two years of age.

Estimates were unadjusted for other covariates because I assumed that seasonal effects on WLZ were exogenous and could not be confounded. Mean differences in WLZ were pooled across cohorts using random-effects models, with cohorts grouped by the Walsh and Lawler seasonality index.<sup>93</sup> Cohorts from years with a seasonal index  $\geq 0.8$  were classified as occurring in locations with high seasonality, cohorts with a seasonal index  $< 0.8$  and  $\geq 0.5$  were classified as occurring in locations with medium seasonality, and cohorts with a seasonal index  $< 0.5$  were classified as occurring in locations with low seasonality, with cutpoints chosen to keep cohorts from the same country grouped together.<sup>93</sup>

$$Seasonal\ Index = \frac{1}{R} \sum_{n=1}^{n=12} |X_n - \frac{R}{12}|$$

Where R = total annual precipitation and  $X_n$  = monthly precipitation.

#### Estimation of mean LAZ, WAZ, and WLZ by age in Demographic and Health Surveys and KI cohorts

I downloaded standard DHS individual recode files for each country from the DHS program website (<https://dhsprogram.com/>). I used the most recent standard DHS datasets for the individual women's, household, and length and weight datasets from each country, and I estimated age-stratified mean LAZ, WAZ, and WLZ from ages 0 to 24 months within each DHS survey, accounting for the complex survey design and sampling weights. See Benjamin-Chung et. al (2019) for additional details on the DHS data cleaning and analysis. I compared DHS estimates with mean LAZ, WAZ, and WLZ by age in the KI study cohorts with penalized cubic-splines with bandwidths chosen using generalized cross-validation.<sup>91</sup>

#### **Statistical analysis**

Replication scripts are available here (URL pending). All analyses were conducted in R version 3.6.1. All pooling was completed using the `rma()` function from the "metafor" package

in the R language (version 3.6.1).<sup>94</sup> Fitted smoothers used in manuscript figures were fit using cubic splines.<sup>91</sup>

### Sensitivity analyses

I estimated incidence rates and of wasting after excluding children born or enrolled wasted (Figure 2.4b). The rationale for this sensitivity analysis is that incident cases at birth imply a different type of intervention (i.e., prenatal) compared with postnatal onset of wasting. I estimated the overall and region stratified prevalence of persistent wasting and of underweight and severe wasting by age (Figure A.4). I also examined the effect of shorter (30 day) and longer (90) day washout period when determining if a child was again at risk when estimating wasting incidence and wasting recovery rates (Figure A.5). Within cohorts that measured child gestational age at birth, I estimated the prevalence of stunting and underweight at birth both uncorrected and corrected for gestational age at birth using the Intergrowth standards (Figure A.6).<sup>95</sup>

## 2.4 Results

### **Datasets**

Eighteen cohorts met the inclusion criteria, encompassing 10,854 children and 187,215 total anthropometry measurements from birth to age 24 months (Figure 2.1). I calculated weight-for-length Z-scores (WLZ) using WHO 2006 growth standards.<sup>57</sup> I dropped 381 unrealistic measurements (0.2%) of WLZ ( $> 5$  or  $< -5$  Z-score), following the WHO recommendations.<sup>88</sup> Cohorts ranged in size from 215 children in the CONTENT cohort<sup>96</sup> to 2,920 children from the MRC Keneba cohort<sup>97</sup> (Figure 2.1). Unless otherwise indicated, I calculated estimates within cohorts and then pooled estimates across cohorts using random effects models fit with restricted maximum likelihood estimation, which is more conservative than fixed-effects models when cohort-specific estimates are heterogeneous.<sup>89,90</sup> Cohorts were distributed throughout South Asia, Africa, and Latin America, but did not cover entire regions (Figure A.2). I assessed the representativeness of the KI cohorts by comparing Z-score measurements to population-based samples measured in DHS. I found that in Latin America, both weight and length were higher in KI cohorts than in population-based DHS samples, but WLZ distributions were similar (Figure A.3). WAZ distributions were nearly identical between the Asian and African KI cohorts and population-based DHS samples, but the mean LAZ was lower in KI cohorts than in their target populations, and thus mean WLZ was higher (Figure A.3). LAZ decreased from birth to 6 months in KI cohorts but held steady over this age range in the DHS data, so the KI cohort estimates of early WLZ recovery may be higher than estimates from population-based samples.

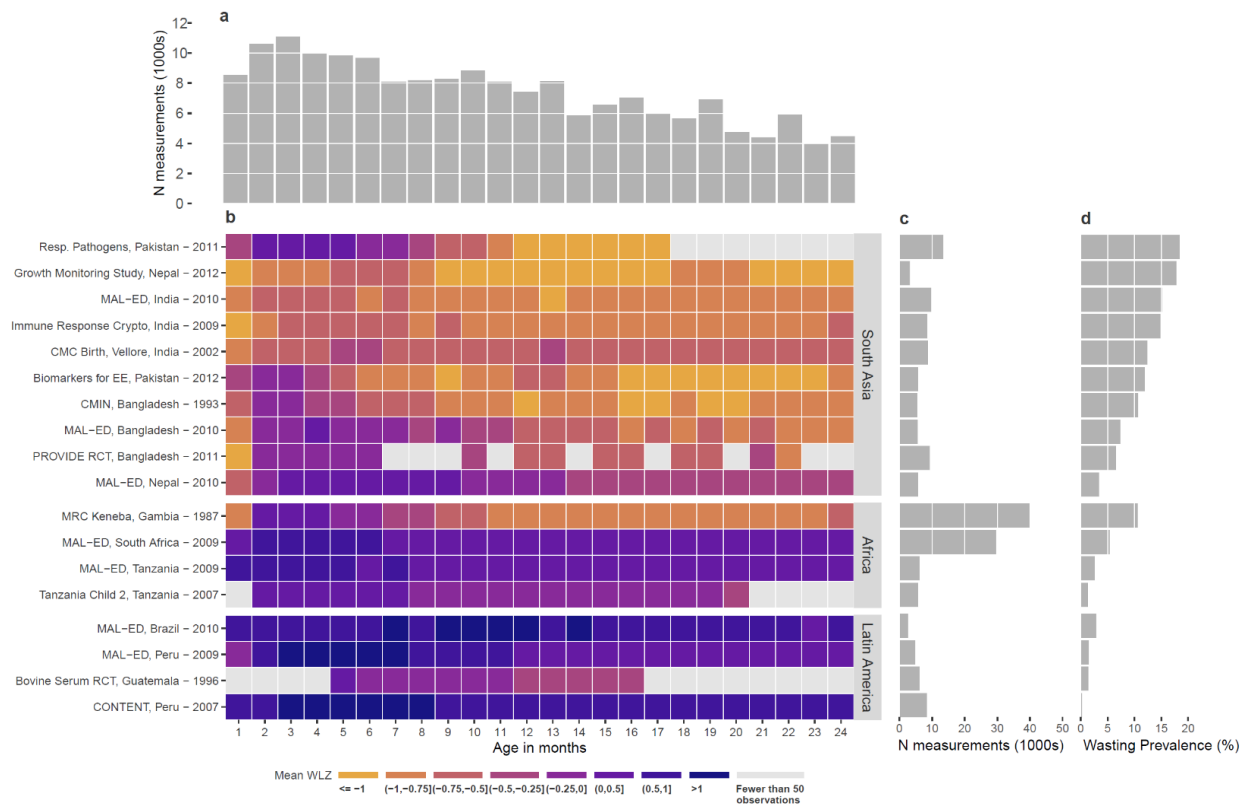


**Figure 2.1 | Summaries of included cohorts.**

(a) Number of observations (1000s) across studies at each age in months. (b) Mean weight-for-length Z-scores by age in months within each included cohort. Cohorts are sorted by geographic region and overall mean weight-for-length Z-score. The country of each cohort and the start year is printed by each cohort name. (c) The number of observations included in each cohort. (d) Overall wasting prevalence by cohort, defined as proportion of measurements with weight-for-length z-score  $< -2$ .

## Regional differences in wasting

Across all observations, mean WLZ was near zero at birth and then increased over the first 6 months before decreasing until 12 months (Figure 2.2a). Region-specific patterns in Latin America and South Asia reflected the overall pattern, although children in South Asian cohorts had much lower WLZ throughout the age range. Mean WLZ among children from African cohorts was similar to children from Latin American cohorts at birth but then steadily decreased until 12 months (Figure 2.2a).

Children were wasted for 15,543 (8.7%) measurements, and severely wasted for 3,290 (1.8%) measurements. The highest wasting prevalence was at birth, which may result from intrauterine growth restriction or preterm birth, with a random-effects pooled prevalence of 11.9% (95% CI: 7.0, 19.5), in contrast to prior research findings that show peaks in wasting prevalence occurring in children between 6-24 months old (Figure 2.2b).<sup>87,98–100</sup> Across all three regions, there was a consistent pattern in prevalence decreasing from birth to three months and then increasing until children were 12 months old, but in South Asia and Latin America wasting prevalence was higher at birth than at 12 months. Severe wasting followed a similar pattern but was much rarer (Figure A.4). Wasting onset was highest during the first 3 months (Figure 2.2c). Overall, almost half (47.8%) of children who ever experienced wasting in the first two years of life experienced the first onset of wasting in their first 3 months, and an estimated 14.8% (95% CI: 9.12, 23.0) of all children experienced wasting by age 3 months. Prevalence measures of wasting masked how common it was for children to experience wasting in their first 24 months. After birth, a maximum of 7.0% (95% CI: 4.4, 11.0) of children were wasted at any specific age, but 33.3% (95% CI: 21.1, 48.3) of children experienced at least one wasting episode by age 24 months, and in South Asian cohorts the proportion was 50.5% (95% CI: 41.6, 59.3, Figure 2.2b, c).

Cohorts from South Asia, where low birthweight is common,<sup>101</sup> had a high wasting prevalence at birth, suggesting that poor maternal nutrition or maternal illness are key regional drivers of wasting.<sup>102</sup> There may be regional differences in wasting etiology, as prior research has found that severely wasted children respond slower to treatment in India than in Africa, potentially due to the younger ages of wasted children, more severe wasting episodes, or the use of WLZ to identify severe wasting.<sup>71,103–106</sup> WLZ cutoffs may capture different severities of growth failure across different populations due to population differences in body composition; children with thinner abdomens will have lower WLZ scores even with identical body fat and muscle composition as children with a thicker abdomens.<sup>107,108</sup>

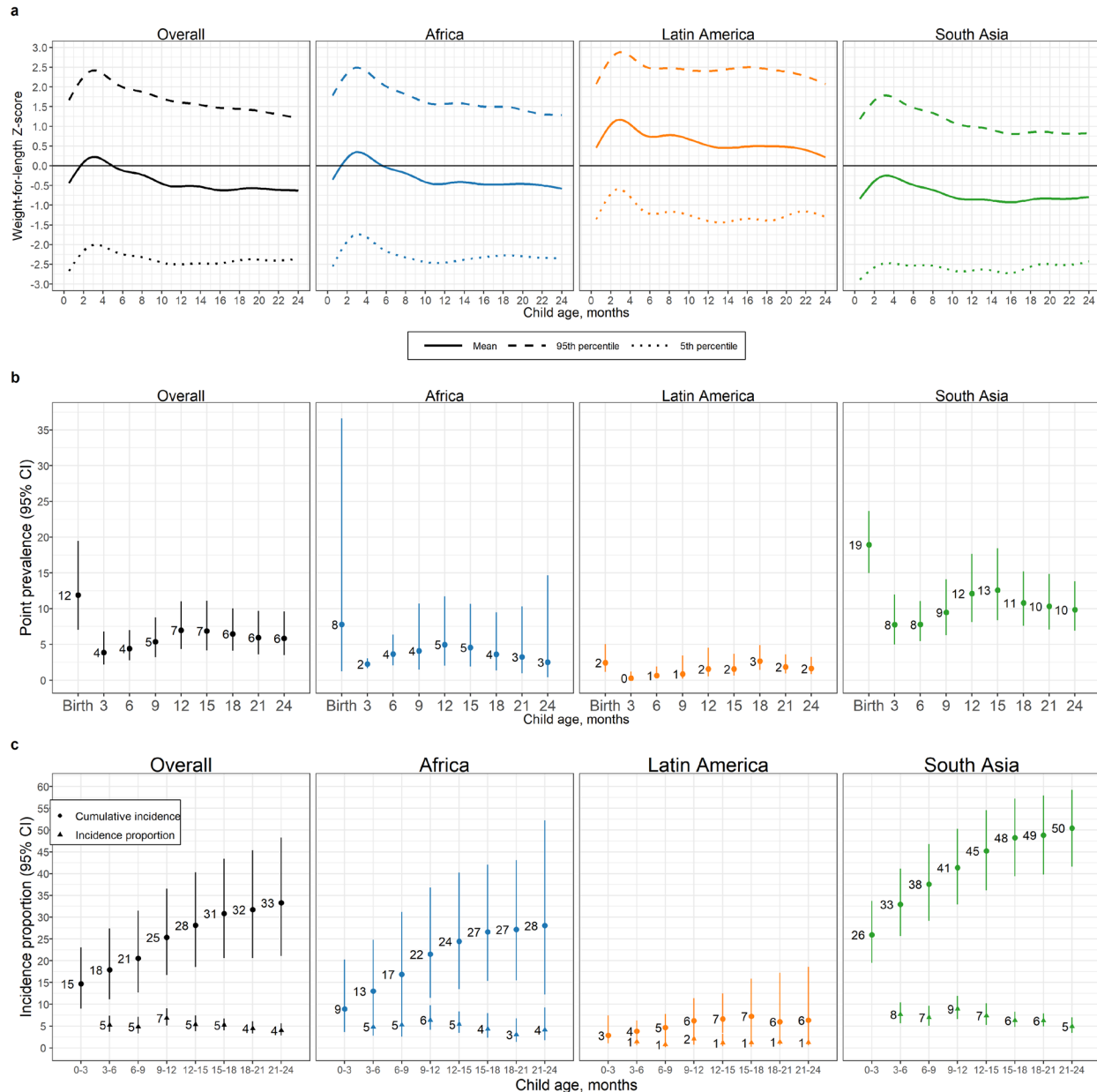


**Figure 2.2 | WLZ, Prevalence and incidence of wasting by age and region.**

(a) Mean, 5th and 95th percentile of weight-for-length Z-score (WLZ) by age in 18 included longitudinal cohorts, overall (N=18 studies; N=13,349-52,389 observations per month) and stratified by region (Africa: N=4 studies, N=1,067-5,428 observations; Latin America: N=4 studies, N=417-1358 observations, South Asia 10 studies, N=2,223-4,142 observations). Spline fits with approximate 95% simultaneous confidence intervals.

(b) Age-specific wasting prevalence, defined as WLZ < -2, overall (N=3,985-9,353 children) and stratified by region (Africa: 1,701-5,017 children, Latin America: N=209-996 children, South Asia: N=1,994-3,605 children). Estimates include children with WLZ measurements within one month before and after that age in the point prevalence estimate, averaging multiple measurements if available for a child. I pooled estimates

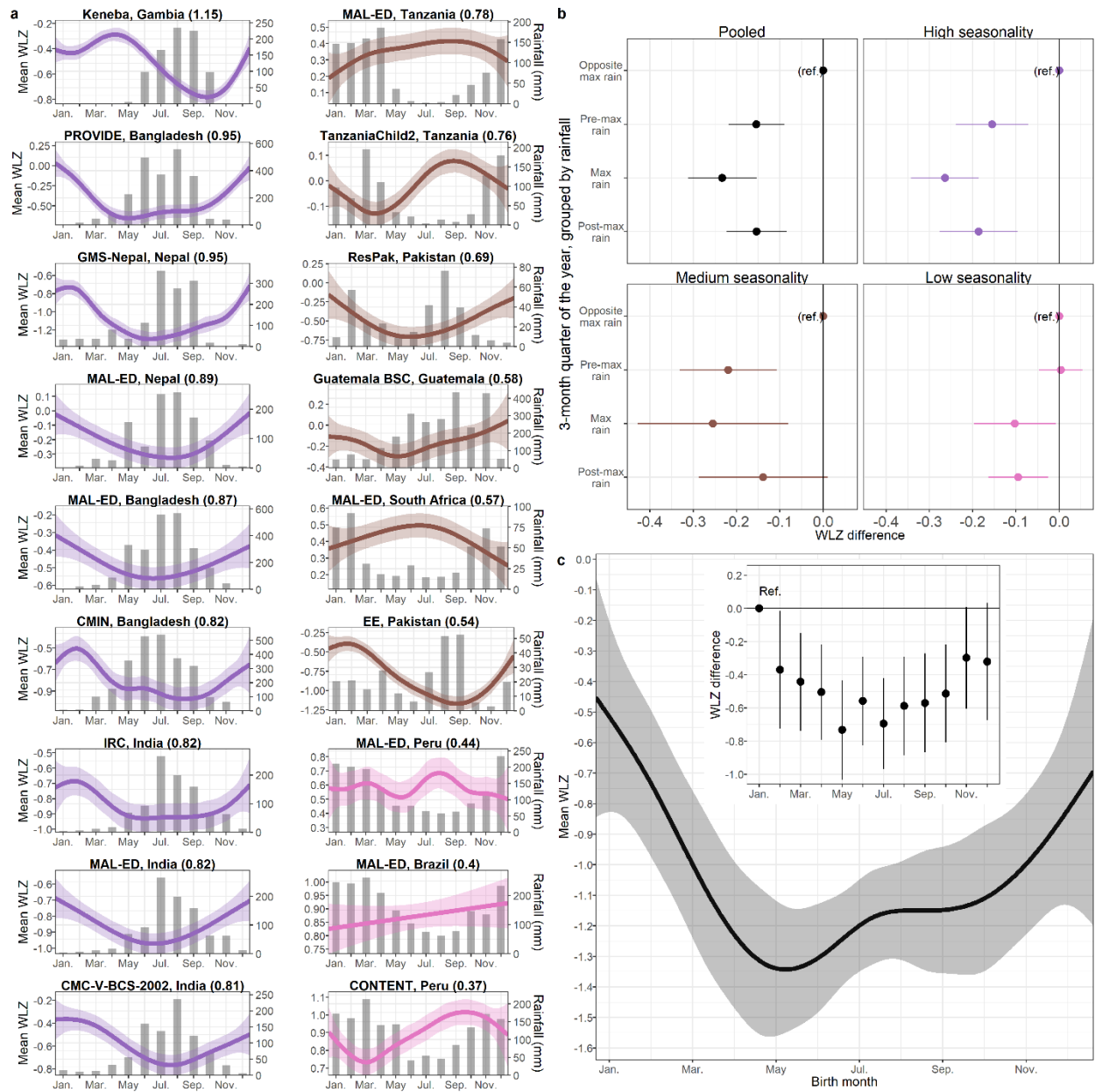
across studies using random effects models fit with restricted maximum likelihood estimation. Vertical bars indicate 95% confidence intervals.

(c) Age-specific wasting incidence overall (N=5,848-9,841 children) and stratified by region (Africa: N=2,249-5,259 children, Latin America: N=572-1,010 children, South Asia: N=2,916-3,832 children). Cumulative incidence measures the proportion of children who have ever experienced wasting since birth, while the new incident cases represent the proportion of children at risk who had an episode of wasting begin during the age range. I pooled estimates across studies using random effects models fit with restricted maximum likelihood estimation. Vertical bars indicate 95% confidence intervals. Although cumulative incidence cannot decrease over time, the estimated pooled cumulative incidence decreased between the 15-18 month and 18-21 month age bands in the Latin America panel because fewer children were measured between 18-21 months.

### Seasonality of wasting

I joined monthly rainfall totals to cohorts by location and year, and for each cohort I estimated a seasonality index based on rainfall (details in Methods).<sup>92,93</sup> I estimated monthly average WLZ and examined seasonal changes with respect to rainfall, under the hypothesis that seasonal changes in food availability tied to rainfall and harvests could cause seasonal wasting. Mean WLZ varied dramatically by calendar date in almost all cohorts, with a consistent minimum mean WLZ coinciding with peak rainfall (Figure 2.3a). Mean WLZ was -0.23 (95% CI: -0.31, -0.15) lower during the three months period of peak rainfall, compared to the mean WLZ in the opposite three-month period of the year, when pooled across cohorts (Figure 2.3b). Average seasonal decline in WLZ was -0.26 (95% CI: -0.34, -0.19) in cohorts with a seasonality index > 0.8, but some cohorts evidenced seasonal WLZ declines of more than 0.5 Z. (Figure 2.3b).

South Asian cohorts had temporally consistent rainfall patterns, so I estimated WLZ at birth by calendar month, pooled across the 10 South Asian cohorts. WLZ at birth ranged from -0.5 Z to -1.3 Z, depending on the month the child was born (Figure 2.3c), which suggests that seasonally-influenced intrauterine growth restriction or preterm birth was a major determinant of a child's WLZ at birth. Birth month also influenced the effect of season on WLZ trajectories that persisted through a child's second year of life (Figure A.5).



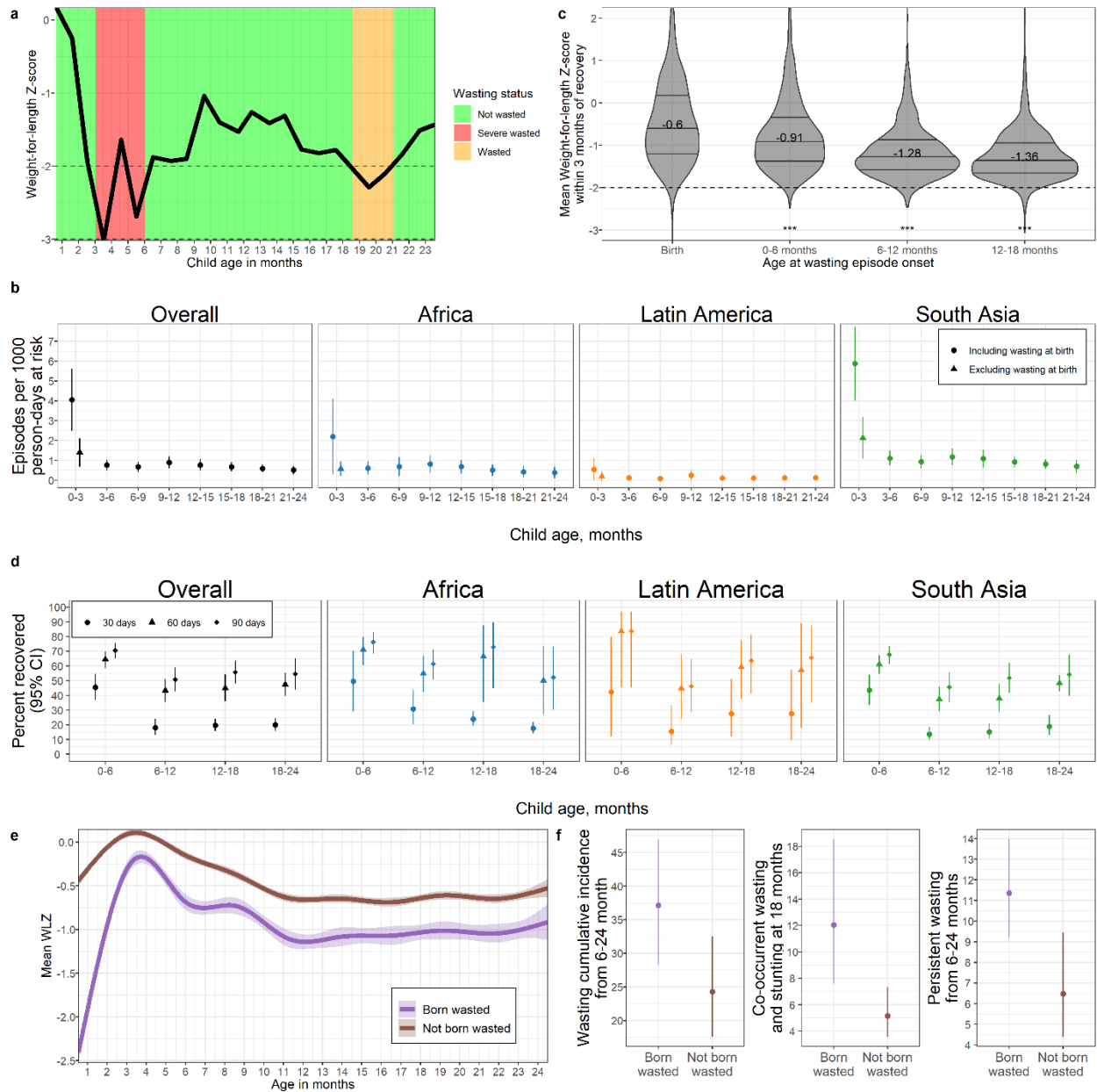
### Figure 2.3 | Mean weight-for-length Z-scores by age and season.

- (a) Cubic splines of mean WLZ over day of the year, superimposed over histograms of monthly mean rainfall over study periods, with the seasonality index printed in parentheses beside the cohort name, and splines colored by high ( $\geq 0.8$ ), medium ( $< 0.8$  and  $\geq 0.5$ ), and low seasonality categories ( $< 0.5$ ). Sample sizes range from 215 children (8,339 measurements) in CONTENT to 2,396 (29,596 measurements) in the Tanzania Child cohort. Rainfall data was extracted from the University of East Anglia Climate Research Unit CRU TS dataset (version 3.23, 23 October 2015).<sup>92</sup> Spline fits with approximate 95% simultaneous confidence intervals are colored by category of seasonal index (high:  $\geq 0.8$ , medium:  $< 0.8$  and  $\geq 0.5$ , low:  $< 0.5$ ) defined by rainfall, and the Walsh and Lawler seasonal index is printed within parentheses next to cohort names.<sup>93</sup> Higher seasonality indices indicate more heterogeneity in rainfall across months of the year. Y-axis scales vary to best illustrate the seasonal averages for individual cohorts.
- (b) Mean differences in child WLZ between quarters of the year defined around the 3-month period with the highest mean rainfall, pooled across cohorts in panel (a) overall (N=18 cohorts, 2,545-40,115 observations) and by seasonality index (high: N=9 cohorts, 5,372-40,115 observations, medium: N=6 cohorts, 5,698-29,518 observations, low: N=3 cohorts, 2,545-8,427 observations). Cohort-specific mean differences were estimated using linear regressions, unadjusted for other covariates. Cohort-specific estimates of seasonal differences in WLZ were pooled using random-effects models.<sup>93</sup> The three-month period opposite the three months of maximum rainfall was used as the reference level (e.g., if June-August was the period of maximum rainfall, the reference level is child mean WLZ during January-March).
- (c) Cubic spline of mean WLZ at birth across birth months among 2,334 children with WLZ measured at birth in ten South Asian cohorts, with mean differences in WLZ at birth by birth month in the inset plot. Estimates are from children with WLZ measured at birth in South Asian cohorts in both the primary and the inset plot. Shaded region around spline fit indicates its 95% simultaneous confidence interval. In the inset plot, the estimation of mean differences in WLZ based on month of birth use the same methods as in Panel (b) to estimate cohort-specific difference, then pool using fixed effect models due to data sparsity. January is used as the reference level

## Wasting incidence and recovery

There was high variability in wasting status across longitudinal measurements and frequent fluctuations of many children's WLZ measurements around the  $-2$  Z threshold. I thus defined unique wasting episodes by imposing a 60-day recovery period when a child's WLZ measurements needed to remain above  $-2$  to be considered recovered and at risk for a future episode (Figure 2.4a). Children who were born wasted were considered to have an incident episode of wasting at birth. Incidence at all ages was higher in South Asia, with a highest incidence rate in the first 3 months, even when excluding episodes of wasting at birth (Figure 2.4b).

Ultimately, most children recovered from moderate or severe wasting episodes, reinforcing its status as an acute condition. Children met my definition of "recovered" in 91.5% of episodes of moderate and 82.5% of episodes of severe wasting. Pooling across cohorts, children recovered from 39.2% (95% CI: 33.9, 44.7) of wasting episodes within 30 days, 64.6% (95% CI: 58.8, 70.0) within 60 days, and 71.0% (95% CI: 65.5, 76.0) within 90 days. Median episode length was 42 (95% CI: 39, 44) days. I examined the distribution of WLZ in the three-month period after a child was considered recovered from wasting, stratified by the age at which the episode occurred. I found that children recovered to a higher WLZ from at-birth wasting or wasting episodes in the 0-6 month age window compared with episodes that occurred at older ages (Figure 2.4c). Consistent with larger WLZ increases during recovery from wasting at younger ages, a larger proportion of children recovered within 30, 60, and 90 days if the wasting episode occurred before age 6 months (Figure 2.4d). South Asian cohorts had lower rates of recovery in the 6-18 month age period compared with other regions (Figure 2.4d). On average, however, the WLZ of children born wasted did not catch up to the WLZ of children not born wasted, and children who were born wasted but recovered had a higher cumulative incidence of wasting after 6 months of age (37.1% cumulative incidence in children born wasted [95% CI: 28.3, 47.0] vs. 24.3% cumulative incidence in children not born wasted [95% CI: 17.6, 32.5], Figure 2.4e-f).



## Figure 2.4| Wasting incidence rate and recovery.

(a) Example child WLZ trajectory and wasting episode classifications. Green intervals mark when the child was not wasted, the yellow interval marks when a child was moderately wasted, and the red interval marks when a child was severely wasted. The age of wasting onset was assumed to occur halfway between a measurement of  $WLZ < -2$  and the previous measurement of  $WLZ \geq -2$ . Recovery from an episode of wasting or severe wasting occurred when a child had measurements of  $WLZ \geq -2$  for at least 60 days, with the age of recovery assumed to be halfway between the last measurement of  $WLZ < -2$  and the first measurement of  $WLZ \geq -2$ .

(b) Wasting incidence rate per 1,000 days at risk, stratified by age and region (Overall:  $N=484,732-769,072$  child-days per estimate, Africa:  $N=245,046-407,940$  child-days, Latin America:  $N=44,300-101,468$  child-days, South Asia:  $N=178,112-302,470$  child-days). Estimates are stratified by child age and geographic region. Wasting incidence was calculated as described in panel (a), and the person-time used in the denominator was determined by the number of days a child was at risk of a new wasting episode (not wasted or in the 60-day recovery period). Incidence rate was estimated by pooling information over cohorts using a random effects estimator; vertical lines indicate 95% confidence intervals.

(c) The distribution of children's mean WLZ in the three months following recovery from wasting, with inter-quartile range marked. Children wasted before age 6 months experienced larger improvements in WLZ compared with children wasted at older ages ( $p < 0.001$ ). The analysis uses 4,234 observations of 2,325 children who recovered from wasting episodes, with 1,219 observations at birth, 615 observations from 0-6 months, 793 observations from 6-12 months, and 937 observations from 12-18 months. Distributions of WLZ after recovery are stratified by age at wasting onset (birth, 0-6m, 6-12m, 12-24m). The horizontal lines within the violins mark the interquartile range, with the median WLZ labeled, and the dashed line marks the  $-2$  Z threshold used to define wasting.

(d) Recovery proportion from wasting within 30, 60, and 90 days of episode onset ( $N=18$  cohorts, 5,589 wasting episodes). Proportion of children who recovered from a wasting episode within 30, 60, and 90 days of episode onset, stratified by region and age at the onset of the wasting episode. Recovery was defined as 2 or more consecutive measures of  $WLZ > -2$ . Proportions estimated by pooling information over cohorts using a random effects estimator; vertical lines indicate 95% confidence intervals.

(e) Mean WLZ by age shows that children born wasted ( $N=809$  children, 14,233 observations) did not catch up to children not born wasted ( $N=3,306$  children, 14,233 observations). Spline curves of children's WLZ over the first two years of life, stratified by wasting status at birth, show that children born wasted experienced large gains in WLZ in the first four months, but on average their WLZ did not catch up to children not born wasted. Estimates are stratified by wasting status at birth, which includes the first measure of a child within 7 days of birth.

(f) Estimates of cumulative incidence of wasting from 6-24 months, prevalence of co-occurrent wasting and stunting at 18 months, and proportion of children persistently wasted from 6-24 months of age among children born wasted ( $N=809$ ) compared to those who were not ( $N=3306$ ).

## Persistent and concurrent growth failure

The longitudinal cohorts enabled me to identify a subset of children who experienced persistent wasting during their first 24 months. I used a pragmatic definition<sup>69</sup> that classified children as persistently wasted if  $\geq 50\%$  of their WLZ measurements from birth to 24 months fell below  $-2$ , which could result from many short wasting episodes or a few longer episodes. Among 9,788 children with at least four measurements, 3.7% (95% CI: 2.2, 6.1) were persistently wasted. Persistent wasting was highest in South Asia, with 6.9% (95% CI: 4.7, 10.2) persistently wasted over the first 24 months of life (Figure A.4). Among children wasted at birth, 11.4% of children were persistently wasted after 6 months (95% CI: 9.2, 14.0) while 6.5% of children not enrolled wasted were persistently wasted after 6 months (95% CI: 4.4, 9.4). Chapter 3 of this dissertation reports key determinants of persistent wasting and Chapter 4 reports its consequences for mortality.

Wasted children who are simultaneously stunted (length-for-age  $Z < -2$ ) are at substantially higher risk of mortality than all other growth phenotypes.<sup>9</sup> I characterized the epidemiology of concurrent wasting and stunting across the cohorts and by region to understand when concurrent wasting and stunting first manifests and when it is most common among children under 24 months. Overall, 1,508 children (14%) were both wasted and stunted during at least one measurement, and 233 (2%) were both severely wasted and severely stunted during at least one measurement. Children who were ever wasted in the first 6 months were 1.3 (95% CI: 1.2, 1.3) times more likely to be stunted and 3.8 (95% CI: 3.2, 4.6) times more likely to be both wasted and stunted at age 24 months. Concurrent wasting and stunting was most common in South Asia, with peak prevalence at ages 12-18 months (Figure 2.5a), driven primarily by increases in stunting prevalence as children aged (Benjamin-Chung et al. 2020). Among anthropometry measurements of wasted children, a large proportion were also classified as underweight (37.9%) or stunted and underweight (39.9%); the highest proportion of only-wasted children occurred at birth (6.2% of all children, Figure 2.5b). Chapter 4 of this dissertation reports the relationship between WLZ and risk of subsequent linear growth failure.

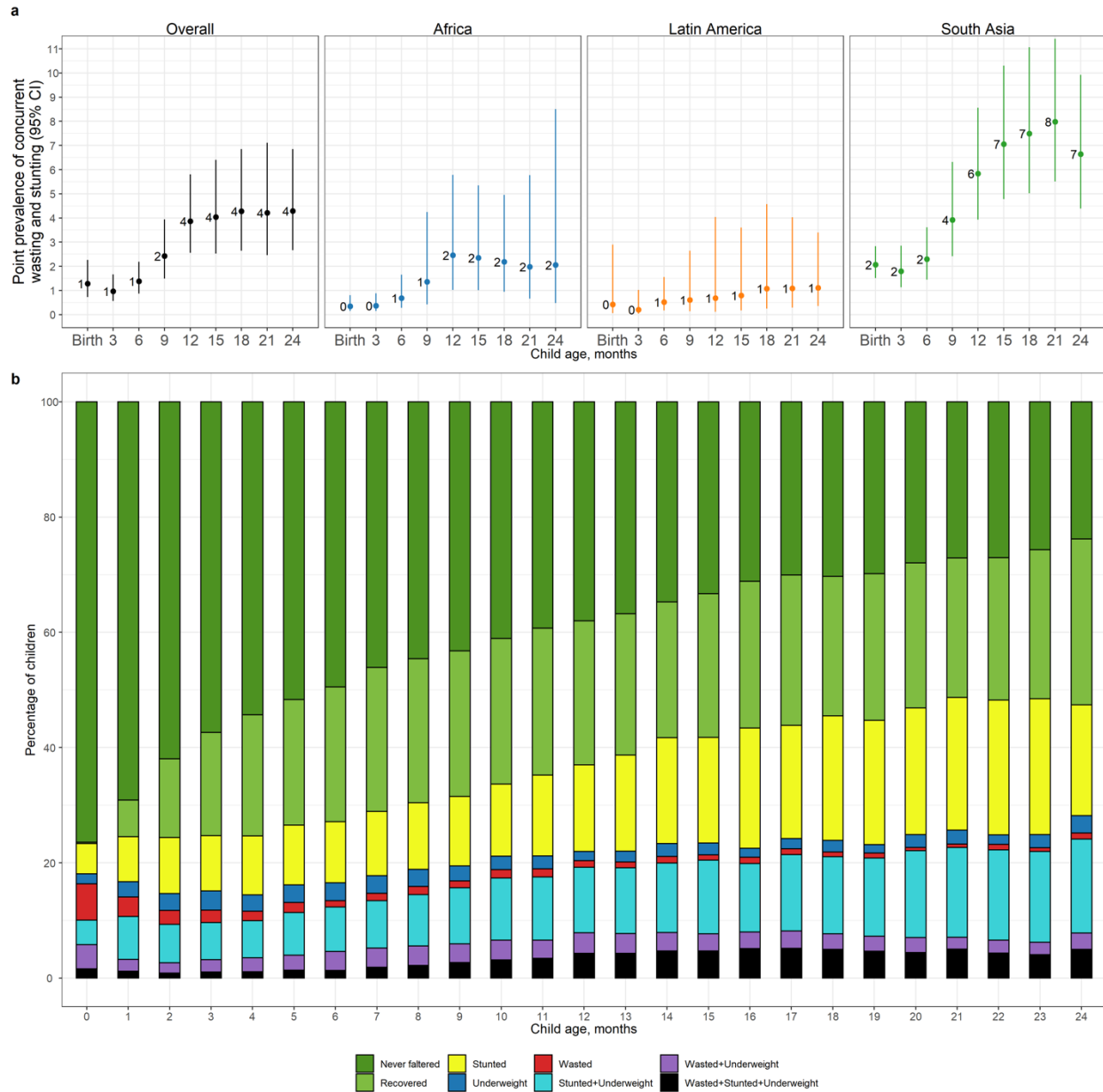


**Figure 2.5| Co-occurrence of wasting, stunting, and underweight.**

(a) Age-specific prevalence of co-occurrent wasting and stunting overall (N=3,984 to 9,346 children) and stratified by region (Africa: N=1,700-5,014 children, Latin America: N= 290-996 children, South Asia: N=1,994-3,601 children). Estimates were pooled across cohorts using random effects models fit with restricted maximum likelihood estimation. All children who were wasted and stunted were also underweight.

(b) Proportion of children with measures of growth failure, alone or combined. Children never faltered at an age in months had never been wasted, stunted, or underweight, and children in the “recovered” category were not wasted, stunted, or underweight but had experienced at least one of these conditions previously. All children who were wasted and stunted were also underweight. The number of children contributing to each

age ranges from 3,623 to 8,652 children, with 10, 836 total children. Proportions in each category were calculated within cohorts, pooled using random effects, and scaled so percentages added to 100%.

## 2.5 Discussion

There were several limitations to this analysis. Z-score distributions in the cohorts were broadly similar to measurements from nationally representative surveys, but South Asian cohorts had lower LAZ (Figure A.3), which could have led to higher rates of wasting recovery in early life compared with the general population. Children were not measured at the same ages across all cohorts, and not every child was measured from birth to age 24 months, so the population underlying each age-specific estimate varied. I assumed recovery from moderate wasting was spontaneous because I did not have consistent data across cohorts on child treatment referral; wasting recovery may thus be higher in these highly monitored cohorts than in the general population due to treatment referrals. I used weight-for-length Z-score to define wasting, as it was measured in all cohorts, but mid-upper arm circumference (MUAC) as an alternative wasting measure has been found in some settings to be a better predictor of subsequent mortality.<sup>109–111</sup> Case fatality rates for severe wasting are high; if incident wasting and mortality occurred between monthly visits I could have under-estimated incidence rates, and prevalence estimates could be lower from survivorship bias.<sup>112,113</sup> Additionally, the WHO Child Growth Standards overestimate wasting at birth among children born preterm, but I did not adjust Z-scores for gestational age in the primary analyses because only five included cohorts measured gestational age via either recall of last menstrual period or newborn examination.<sup>114</sup> Adjustment for gestational age among cohorts that measured it reduced at-birth stunting prevalence by one percent and at-birth underweight prevalence by two percent (Figure A.6). Even if preterm birth accounts for some of at-birth wasting, the small birth size — by any cause — documented in this analysis raises concern because of its consequences for child growth failure and mortality during the first 24 months (Mertens et al. 2019).

My results provide new insights into the dynamic process of child wasting with important implications for preventive interventions. WLZ varied dramatically by season in most cohorts, and in South Asia at-birth WLZ varied by 0.7 Z-scores (95% CI: 0.4, 1.0) depending on birth month with consequences that persisted throughout the first 24 months (Figure 2.3, Figure A.5). These results provide strong support for the development of seasonally targeted interventions to prevent wasting in food-insecure populations. The longitudinal analyses show that the wasting burden is likely far larger than can be quantified with prevailing methods based on cross-sectional, prevalence surveys. By age 24 months, 50% of children in South Asian cohorts experienced at least one wasting episode and 7% were persistently wasted — new insights only possible from the longitudinal data analyses. Crucially, I found that onset of wasting was highest from birth to age 3 months, and among children who experience any wasting during their first 24 months, 48% experience their first wasting episode during this early period. Wasting treatment programs have traditionally focused on children from 6-59 months. My results should motivate preventive and therapeutic programs to consider extending efforts to mothers during pregnancy and early lactation, and to children during the first months of life. If preventive or therapeutic

interventions focus on ages earlier than 6 months, then they must integrate carefully with current recommendations for exclusive breastfeeding.

## **Conclusion**

These findings suggest that, among children who experience wasting in the first two years of life, the first occurrence of wasting occurs earlier than prior research has indicating, and many children are wasted at birth, especially in South Asia and during seasonal rainfall. Interventions to prevent child acute malnutrition may have the largest impact if delivered earlier, including in utero and through the first months of life, and if targeted to reduce seasonal peaks in wasting.

# 3 Causes of child growth failure in low resource settings

## Preface

The work presented in this chapter is the product of myself and a large number of collaborators, and we plan on publishing a subset of this research and the research presented in chapter 4 as a single manuscript after submission of this dissertation, with myself as first author. The Bill and Melinda Gates Foundation (BMGF) funded a University of California, Berkeley team led by Professors John Colford, Alan Hubbard, and Mark van der Laan to investigate risk factors for poor child growth using the KI data. The research questions and analysis plan for this chapter was developed by me, Benjamin Arnold (UC San Francisco), Jade Benjamin Chung (UC Berkeley), John Colford, Kenneth H. Brown (BMGF), and Parul Christian (BMGF). Mark van der Laan, Alan Hubbard, Jeremy Coyle (UC Berkeley), and Oleg Sofrygin (UC Berkeley) led the KI-funded development of the new generation of targeted learning statistical software used for analyses in this chapter, and the following UC Berkeley graduate students assisted in statistical programming: Wilson Cai, Nima Hejazi, Ivana Malenica, and Haodong Li. The following UC Berkeley graduate students, and undergraduate students assisted in researching the methods of exposure measurement used in included studies, statistical coding, data analysis, and figure creation: Anna Nguyen, Nolan Pokpongkiat, Wendy Jilek, Esther O Chung, and Haodong Li. Ryan Hafen (BMGF) provided visualization suggestions and Vishak Subramoney (BMGF) assisted in data harmonization. Thea Norman and Jonas Häggström (BMGF) assisted in coordinating research with BMGF subject-area experts and securing data-use agreements with primary investigators of studies contributed to KI. Benjamin Arnold will be senior author on the planned publication using analyses in chapters 3 and 4, and had the largest contribution to developing the research methods and drafting the manuscript after myself.

## 3.1 Background

Worldwide, 22% of children under age 5 years are stunted, a measure of chronic growth failure, and 7% are wasted, a measure of acute growth failure, with most of the burden occurring in low- and middle- income countries (LMIC).<sup>61</sup> Both may contribute to child mortality and adult morbidity.<sup>115,116</sup> Despite widespread recognition of the importance of growth failure to global public health,<sup>115,116</sup> preventive interventions in LMICs have had limited success.<sup>7,117,118</sup> A range of nutritional interventions, targeting all stages of fetal and child development, including nutrition education, food and micronutrient supplementation during pregnancy, promotion of exclusive breastfeeding for 6 months and continued breastfeeding for 2 years, and food and micronutrient supplementation during weaning, have been found to have a beneficial effect on child growth.<sup>64,119,120</sup> However, postnatal breastfeeding interventions and nutritional interventions delivered to children who have begun complementary feeding have only had small effects on population-level stunting and wasting burdens.<sup>119–122</sup> Additionally, water, sanitation, and hygiene (WASH) interventions, which aim to reduce childhood infections that may heighten the risk of wasting and stunting in non-emergency settings<sup>123,124</sup>, have had no effect on child growth in several recent large randomized control trials.<sup>80,81,125,126</sup> It may be that these preventative interventions have had small effect sizes because of an incomplete understanding of

how to intervene, when in early life children's growth responds the best to interventions, and what the characteristics are of children who would most benefit.<sup>78</sup>

Modest effects of interventions to prevent stunting and wasting in recent decades have spurred renewed interest in efforts to combine rich data sources with advances in statistical methodology to more deeply understand the key causes of child growth failure.<sup>7,13,60,80,81</sup> Understanding the relationship between the causes and timing of growth failure is also crucial because children who falter early could be at higher risk for more severe growth failure later. In Chapter two of this dissertation, I report that the highest rate of incident wasting occurs by age 3 months, and a companion paper conducted on the same set of studies used in this dissertation also found that the highest rate of incident stunting occurs by age 3 months (Benjamin-Chung et al. 2020, in preparation). Behaviors associated with higher risk of stunting or wasting could be targeted by future interventions, and interventions could be optimized to encourage behavior change before the age at which growth failure occurs. Characteristics associated with higher risk could also be used to identify children at risk of growth failure who might benefit most from preventative interventions.

Here, I report a pooled analysis of 35 longitudinal cohorts in 15 low- and middle-income countries in South Asia, Sub-Saharan Africa, Latin America, and Eastern Europe, initiated between 1969 and 2014. Cohorts were assembled as part of the Knowledge Integration (KI) initiative of the Bill & Melinda Gates Foundation which includes a database of millions of participants from studies on childbirth, growth and development. My objective was to estimate relationships between child, parental, and household characteristics and measures of child chronic and acute growth failure, including length-for-age and weight-for-length Z-scores, stunting, wasting, and length and weight velocities from birth to age 24 months. Details on the estimation of growth failure outcomes are included in Chapter 2 of this dissertation. I estimated mean differences in Z-scores and relative risks of indicators of growth failure across levels of these characteristics, which capture the individual effects of exposure, but I also rank-ordered the characteristics by population attributable difference and fractions. These population intervention effects better capture the public health importance of each characteristics by estimating the effect of a population shift in exposure, thereby comparing the observed exposure distribution with a counterfactual target distribution hypothetically achievable by interventions.<sup>127,128</sup>

## 3.2 Objectives

- 1) Rank order child, parental, and household exposures by population-attributable associations with measures of growth failure.
- 2) Estimate if associations between exposures and growth failure vary by region, child age, measure of growth failure, or severity of growth failure.
- 3) Explore the relative importance of child size at birth compared to postnatal growth failure in determining child growth trajectories.

## 3.3 Methods

### 1. Study designs and inclusion criteria

I included all longitudinal observational studies and randomized trials available through the KI project on April 1, 2018 that met five inclusion criteria: 1) conducted in low- or middle-income countries; 2) enrolled children between birth and age 24 months and measured their length and weight repeatedly over time; 3) did not restrict enrollment to acutely ill children; 4) enrolled at least 200 children; 5) collected anthropometry measurements at least every three months (Figure

B.1, Table B.1). Inclusion criteria ensured I could rigorously evaluate the timing and onset of growth failure among children who were broadly representative of populations in low- and middle-income countries. I included all children under 24 months of age, assuming months were 30.4167 days, and I considered a child's first measure recorded by age 7 days as their anthropometry at birth.

## **2. Outcome definitions**

Anthropometry Z-scores were calculated using the 2006 WHO standards.<sup>129</sup> I used the medians of triplicate measurements of heights and weights of children from pre-2006 cohorts to recalculate Z-scores to the 2006 standard. I excluded extreme measurements of WLZ > 5 or < -5, WAZ > 5 or < -6, and LAZ > 6 or < -6, consistent with WHO growth standard recommendations.<sup>88</sup> See Benjamin-Chung (2019) for details on cohort inclusion and assessment of anthropometry measurement quality. Children with length-for-age Z-scores (LAZ), weight-for-length Z-scores (WLZ), or weight-for-age Z-scores (WAZ) < -2 were classified as stunted, wasted, or underweight, respectively. Children with LAZ, WLZ, or WAZ < -3 were classified as severely stunted, severely wasted, or severely underweight, respectively. Children with  $\geq 50\%$  of WLZ measurements < -2 over a defined age range were classified as persistently wasted. Children were assumed to never recover from stunting episodes, but children were classified as recovered from wasting episodes (and at risk for a new episode of wasting) if their measured WLZ was  $\geq -2$  for at least 60 days (details in Mertens et. al (2019)).

## **3. Estimating relationships between child, parental, and household exposures and measures of growth failure**

### **3.1 Exposure definitions**

I selected the exposures of interest based on variables present in multiple cohorts that met my inclusion criteria, were found to be important predictors of stunting and wasting in prior literature and could be harmonized across cohorts for pooled analyses. Table B.2 lists all exposures included in the analysis, as well as exposure categories used across cohorts, and the total number of children in each category. Figure B.1 visualizes exposure presence by cohort. For parental education and asset-based household wealth, I categorized to levels relative to the distribution of educational attainment within each cohort. Continuous biological characteristics (gestational age, birth weight, birth height, parental weight, parental height, parental age) were classified based on a common distribution, pooling data across cohorts. My rationale was that the meaning of socio-economic variables is culturally dependent, whereas biological variables should have a more universal meaning.

### **3.2 Risk set definition**

For exposures that occur or exist before birth, I considered the child at risk of incident outcomes at birth. Therefore, I classified children who were born stunted (or wasted) as incident episodes of stunting (or wasting) when estimating the relationship between household characteristics, paternal characteristics, and child characteristics like gestational age, sex, birth order, and birth location.

For postnatal exposures (e.g., breastfeeding practices, WASH characteristics, birth weight), I excluded episodes of stunting or wasting that occurred at birth. Children who were born wasted could enter the risk set for postnatal exposures if they recovered from wasting during the study period (see Mertens et al. 2019 for details). This restriction ensured that for postnatal exposures, the analysis only included postnatal, incident episodes. Children born or enrolled wasted were included in the risk set for the outcome of recovery from wasting within 90 days for all exposures (prenatal and postnatal).

### 3.3 Estimating differences in outcomes across categories of exposures

I estimated measures of association between exposures and growth failure outcomes by comparing outcomes across categories of exposures in four ways:

Mean difference of the comparison levels of the exposure ( $A=a$ ) on LAZ, WLZ at birth, 6 months, and 24 months compared to a reference level ( $A=0$ ), and conditional on covariates  $W$ . The Z-scores used were the measures taken closest to the age of interest and within one month of the age of interest, except for Z-scores at birth which only included a child's first measure recorded by age 7 days.

$$(1) \theta_{A=a} = E(Y | A = a, W) - E(Y | A = 0, W)$$

Where  $\theta_{A=a}$  is the difference between level  $a$  of exposure  $A$  and the reference level  $A=0$ , adjusted for  $W$ .

Prevalence ratios (PR) between comparison levels of the exposure, compared to the reference level at birth, 6 months, and 24 months. Prevalence was estimated using anthropometry measurements closest to the age of interest and within one month of the age of interest, except for prevalence at birth which only included measures taken on the day of birth. The parameter of interest is:

$$(2) \phi_{A=a} = \log E(Y | A = a, W) - \log E(Y | A = 0, W)$$

Where  $e^{\phi_{A=a}}$  is the PR comparing level  $a$  of exposure  $A$  to the reference level, adjusted for  $W$ .

Cumulative incidence ratios (CIR) between comparison levels of the exposure, compared to the reference level, for the incident onset of outcomes between birth and 24 months, 6-24 months, and birth-6 months. The parameter of interest is:

$$(3) \Psi_{A=a} = \log E(Y | A = a, W) - \log E(Y | A = 0, W)$$

Where  $e^{\Psi_{A=a}}$  is the CIR comparing level  $a$  of exposure  $A$  to the reference level  $A = 0$ , adjusted for the confounder set  $W$ .

Mean Z-scores by continuous age, stratified by levels of exposures, from birth to 24 months were fit within individual cohorts using cubic splines with the bandwidth chosen to minimize the median Akaike information criterion across cohorts.<sup>91</sup> I estimated splines separately for each exposure category. I pooled spline curves across cohorts into a single prediction, offset by mean Z-scores at one year, using random effects models.<sup>130</sup>

### 3.4 Estimating population attributable parameters

I estimated three measures of the population-level effect of exposures on growth failure outcomes:

Population attributable difference was defined as the change in population mean Z-score if the entire population's exposure was set to an ideal reference level. For each exposure, I chose reference levels as the category with the highest mean LAZ or WLZ across cohorts. The parameter of interest is:

$$(4) \theta_{PAD} = E\{E(Y | A, W) - E(Y | A = 0, W)\}$$

Note that  $\theta_{PAD}$  is very similar to the mean difference, but instead of comparing a counterfactual difference of Z-scores for a population completely exposed ( $A=1$ ) with an unexposed population ( $A=0$ ), I compared the Z-scores observed under the empirical distribution of exposure ( $A$ ), with a counterfactual in which every child is exposed to the lowest-risk reference level ( $A=0$ ).

Population attributable fraction (PAF) was defined as the proportional reduction in cumulative incidence if the entire population's exposure was set to an ideal low risk reference level. I estimated the PAF for the prevalence of stunting and wasting at birth, 6, and 24 months and cumulative incidence of stunting and wasting from birth to 24 months, 6-24 months, and from birth to 6 months. For each exposure, I chose the reference level as the category with the lowest risk of stunting or wasting. The parameter of interest is:

$$(5) \theta_{PAF} = \frac{E\{E(Y | A, W) - E(Y | A=0, W)\}}{E\{E(Y | A, W)\}} \times 100\%$$

Optimal individualized intervention impact I employed a variable importance measure (VIM) methodology to estimate the impact of an optimal individualized intervention on an exposure. The optimal intervention on an exposure was determined through estimating individualized treatment regimes, which give an individual-specific rule for the lowest-risk level of exposure based on individuals' measured covariates. The impact of the optimal individualized intervention is derived from the VIM, which is the predicted change in the population-mean outcome from the observed if every child's exposure was shifted to their optimal level. This differs from the attributable difference and PAF parameters in that I did not specify the reference level, and the reference level could vary across participants. The estimated individual lowest risk level is denoted  $A = d(W)$ , the dynamic lowest-risk level estimated from covariates  $W$ . The parameters of interest are:

$$(6) \text{Opt}_{PAD} = E\{E(Y | A, W) - E(Y | A = d(W), W)\}$$

$$(7) \text{Opt}_{PAF} = \frac{E\{E(Y | A, W) - E(Y | A=d(W), W)\}}{E\{E(Y | A, W)\}} \times 100\%$$

## 4. Estimation approach

### Estimation of cohort-specific effects

For each exposure, I used directed acyclic graphs (DAGs) to identify potential confounders from the broader set of exposures used in the analysis.<sup>131</sup> I did not condition on characteristics that were assumed to be intermediate on the causal path between any exposure and the outcome.<sup>132</sup>

For missing covariate observations, I imputed missing measurements as the median (continuous variables) or mode (categorical variables) among all children within each cohort, and analyses included an indicator variable for missingness in the adjustment set. When calculating the median for imputation, I used children as independent units rather than measurements so that children with more frequent measurements were not over-represented.

To flexibly adjust for potential confounders and reduce the risk of model misspecification, I estimated the parameters listed in section 4 using targeted maximum likelihood estimation (TMLE), a two-stage estimation strategy that incorporates state-of-the-art machine learning algorithms (super learner) while still providing valid statistical inference.<sup>60,133</sup> The super learner is an ensemble machine learning method that uses cross-validation to select a weighted combination of predictions from a library of algorithms. Super learner models were used to estimate  $E[Y_a]$ , the mean outcome under the counterfactual that the population exposure level was set to  $A = a$ . Under the randomization assumption (no unmeasured confounding) and the positivity assumption (exposure level  $A = a$  is possible in all strata of  $W$ ), this is estimated as follows:

$$(8) \quad \psi_a = E_W\{E_{Y|A,W}(Y|A = a, W)\}$$

I estimated these parameters by using super learner models to fit the relevant likelihood factors  $E_{Y|A,W}(Y|A = a, W)$  and  $P(A = a|W)$ , and then updating the likelihood fit using TMLE. I then used the delta method to estimate relative risks and attributable risks relative to a prespecified baseline level of  $A = 0$ .

I included in the super learner library simple means, main effects generalized linear models, LASSO penalized regressions,<sup>134</sup> generalized additive models,<sup>135</sup> and gradient boosting machines.<sup>136</sup> The super learner was fit to maximize the 10-fold cross-validated area under the receiver operating characteristic curve (AUC) for binomial outcomes, and minimize the 10-fold cross-validated mean-squared error (MSE) for continuous outcomes. That is, the super learner was fit using 9/10 of the data, while the AUC/MSE was calculated on the remaining 1/10 of the data. Each fold of the data was held out in turn and the cross-validated performance measure was calculated as the average of the performance measures across the ten folds. This approach is practically appealing and robust in finite samples, since this cross-validation procedure utilizes unseen sample data to measure the estimator's performance. Also, the super learner is asymptotically optimal in the sense that it is guaranteed to outperform the best possible algorithm included in the library as sample size grows. The initial estimator obtained via super learner is subsequently updated to yield an efficient double-robust semi-parametric substitution estimator of the parameter of interest. To estimate the R2 of models including multiple exposures, I fit super learner models, without the targeted learning step, and within each cohort measuring the exposures. I then pooled cohort-specific R2 estimates using fixed effects models.

Unadjusted PRs and CIRs between the reference level of each exposure and comparison levels were estimated using logistic regressions.<sup>137</sup> Unadjusted mean differences for continuous outcomes were estimated using linear regressions.

I did not estimate relative risks between a higher level of exposure and the reference group if there were 5 or fewer cases in either stratum. In such cases, I still estimated relative risks between other exposure strata and the reference strata if those strata were not sparse.

I estimated influence curve-based, clustered standard errors to account for repeated measures in the analyses of recovery from wasting or progression to severe wasting. I assumed that the children were the independent units of analysis unless the original study had a clustered design, in which case the unit of independence in the original study were used as the unit of clustering. I used clusters as the unit of independence for the iLiNS-Zinc, Jivita-3, Jivita-4, Probit, and SAS Complementary Feeding trials. I estimated 95% confidence intervals for incidence using the normal approximation. All analyses were conducted using R version 3.6.1.

## 5. Pooling parameters

I pooled adjusted estimates from individual cohorts using random effects models, fit using restricted maximum likelihood estimation. The pooling methods are detailed in Chapter 2. All parameters were pooled directly using the cohort-specific estimates of the same parameter, except for population attributable fractions. Pooled PAFs were calculated from random-effects pooled population attributable risks (PARs), and pooled outcome prevalence in the population using the following formulas:<sup>138</sup>

$$(9) \quad PAF = \frac{PAR}{Outcome\ prevalence} \times 100$$

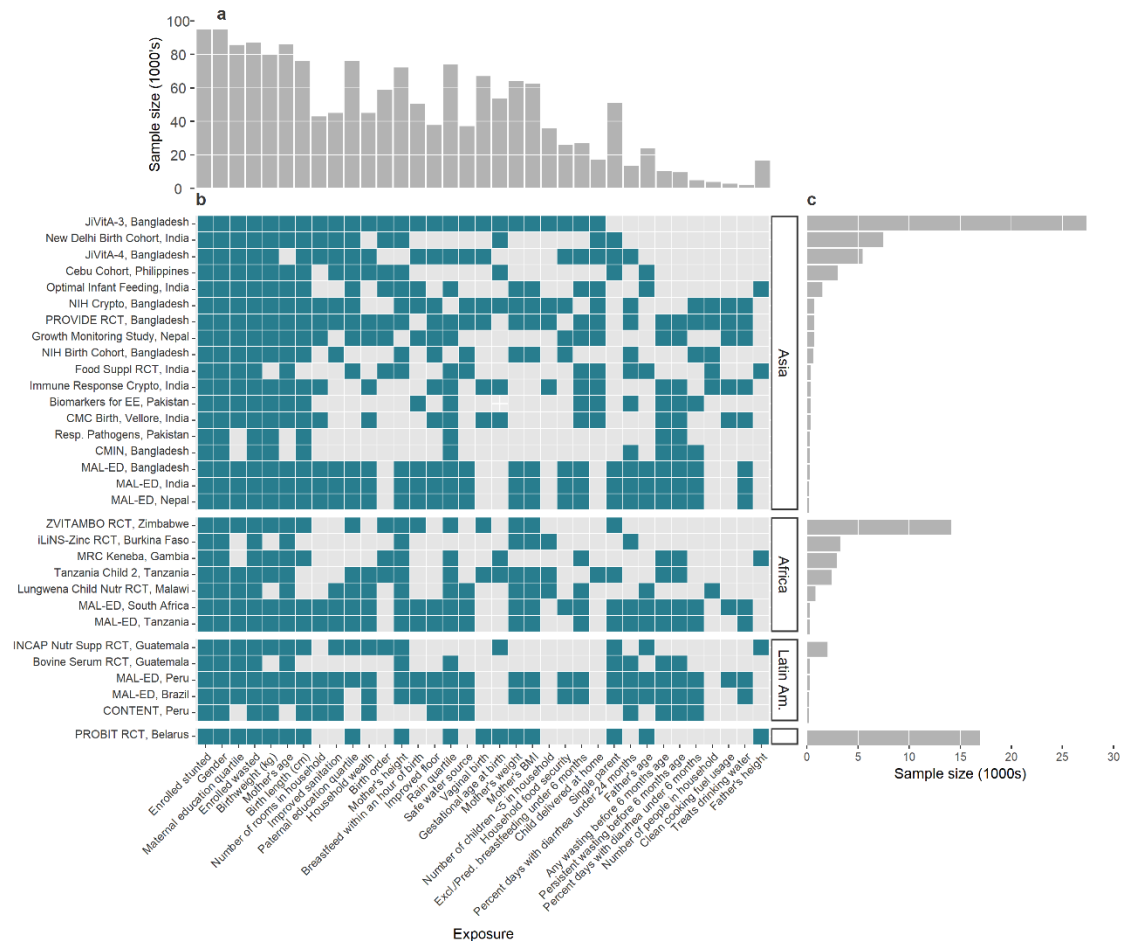
$$(10) \quad PAF\ 95\%CI = \frac{PAR\ 95\%CI}{Outcome\ prevalence} \times 100$$

For PAFs of exposures on the cumulative incidence of wasting and stunting, the pooled cumulative incidence was substituted for the outcome prevalence in the above equations. I used this method instead of direct pooling of PAFs because, unlike PAFs, PARs are unbounded with symmetrical confidence intervals.

## 3.4 Results

### Datasets

Thirty-one cohorts from 15 countries met inclusion criteria, and 94,019 children and 645,869 total measurements were included in this analysis (Figure 3.1, Table B.1). Cohorts were distributed throughout South Asia, Africa, and Latin America, with a single European cohort from Belarus. I calculated length-for-age Z-scores (LAZ), weight-for-age Z-scores (WAZ), and



**(a)** Total number of children with a measured exposure, sorted from left to right by number of cohorts measuring the exposure. **(b)** Presence of exposure variables in the KI data by within each included cohort. Cohorts are sorted by geographic region and sample size. **(c)** The number of observations of child anthropometry included in each cohort.

## Rank ordered causes of growth failure

I selected exposures of interest based on important predictors of stunting and wasting from prior literature that were measured in multiple cohorts and could be harmonized across cohorts for pooled analyses (Figure 3.1, Table B.2). All reported estimates were adjusted for all other measured exposures that I assumed were not on the causal pathway between the exposure of interest and the outcome. For example, the association between maternal height and stunting was not adjusted for a child's birthweight because low maternal height could increase stunting risk through lower child birthweight for gestational age.<sup>140</sup> Parameters were estimated using targeted maximum likelihood estimation, a doubly-robust, semiparametric method that allows for valid inference while adjusting for potential confounders using ensemble machine learning (details in Methods).<sup>133</sup> I estimated cohort-specific parameters, adjusting for measured covariates within each cohort, and then pooled estimates across cohorts using random effects models (Figure B.1).<sup>89,90</sup> When estimating relative risks, Z-score differences, and attributable risk parameters, I chose the reference as the mode of the level of lowest risk across cohorts. I also estimated the effects of optimal dynamic treatment interventions, where no a-priori reference level of low risk was specified, and each child's individual low-risk level of exposure was estimated from covariates. Timing of exposures varied, from parental and household characteristics present before birth, to fetal or at-birth exposures, and postnatal exposures including breastfeeding and diarrheal disease. I estimated only associations for growth failure occurring after exposure measurements to ensure time-ordering of exposures and outcomes.

Longer child birth length, higher maternal weight, earlier child birth order, higher maternal educations, and more rooms in the household were five of the top six population-level predictors of higher LAZ and WLZ at 24 months, as rank-ordered by population attributable difference, the estimated shift in population mean Z-score if the whole population had their exposure shifted from observed levels to the lowest-risk reference level (Figure 3.2a, b). The pooled, cross-validated R<sup>2</sup> for models that included these five key determinants, plus child sex and birthweight, was 0.29 for LAZ (N= 15 cohorts, 22,193 children) and 0.09 for WLZ (N=15 cohorts, 20,927 children). The dry season of the year was also an important predictor of higher WLZ, and taller mother's height was an important predictor of higher LAZ. Mother's height was a stronger predictor of both LAZ and WLZ than father's height, which may reflect that maternal anthropometric status integrates across multiple distal and proximate causes, such as family socio-economic status (SES), fetal growth environment, and breastmilk quality. Maternal height and weight and child characteristics measured at birth were the strongest predictors of LAZ and WLZ at age 24 months; beyond those, key predictors of higher Z-scores included markers of better household socioeconomic status (e.g., number of rooms in the home, parental education, clean cooking fuel use, household wealth index) and having a cesarean birth, which may reflect healthcare access or larger fetal size. The findings underscore the importance of prenatal exposures for child growth outcomes, and at the population-level growth failure may be difficult to shift without broad improvements in standard of living.<sup>141,142</sup> Exclusive or predominant breastfeeding before 6 months of age, which was not a major predictor of Z-scores at 24 months, was more strongly associated with higher WLZ than with higher LAZ at 6 months of age (Figures B.2,B.3,B.4).

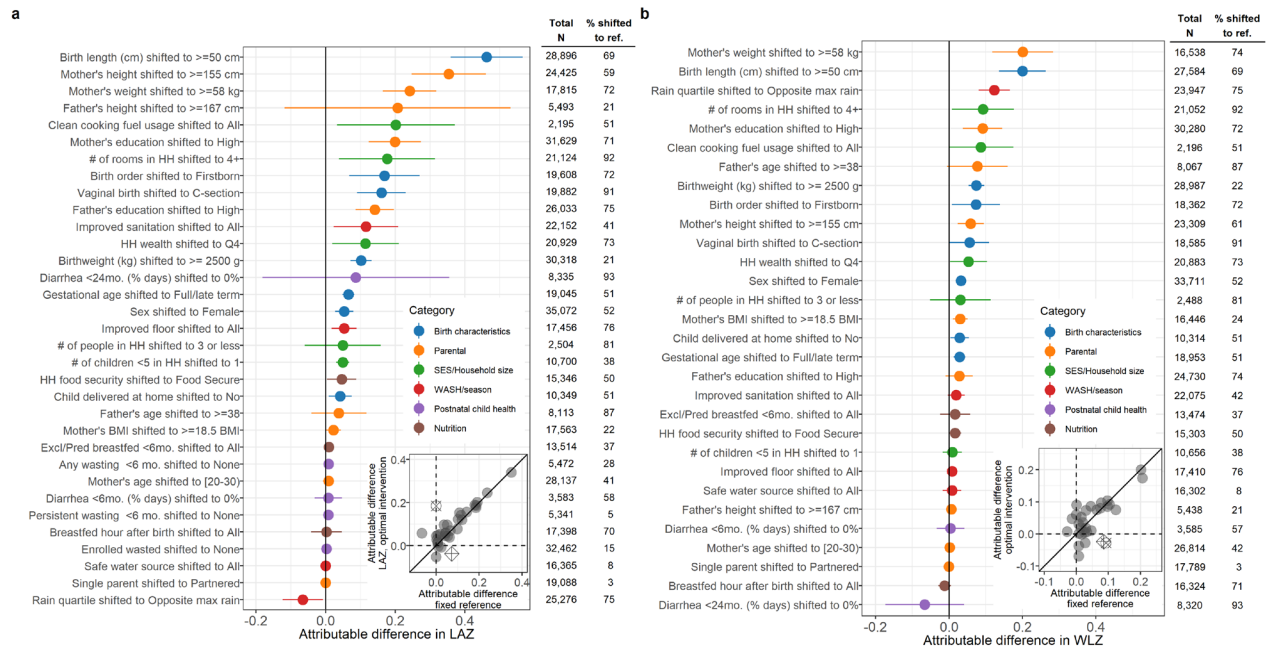


**Figure 3.2 | Rank-ordered associations between child, parental, and household characteristics and population attributable differences in anthropometry Z-scores.**

**(a)** Exposures, ordered by pooled population attributable difference on child LAZ at 24 months. The population attributable difference is the expected difference in population mean Z-score if all children had the reference level of the exposure rather than the observed distribution. Estimates are colored by broad category of exposure. Columns on the right list the total number of children with each exposure measured, and the percent of children with their exposure shifted to the reference level in the estimation of population attributable differences. Inset figures plot the population attributable difference on the X-axis, and the optimal intervention attributable difference on the Y-axis, where the level the exposure is shifted to can vary by child. The optimal intervention attributable differences, which are not estimated with an a-priori specified low-risk reference level, were generally close to the static attributable differences, indicating that the chosen reference levels were the lowest risk strata in most or all cohorts. The largest outliers are marked: the crossed circle marks the single parent variable and the diamond marks the father's height.

**(b)** Exposures, rank ordered by population attributable difference on child WLZ at 24 months. The crossed circle marks the father's age variable and the diamond marks the birth order variable.

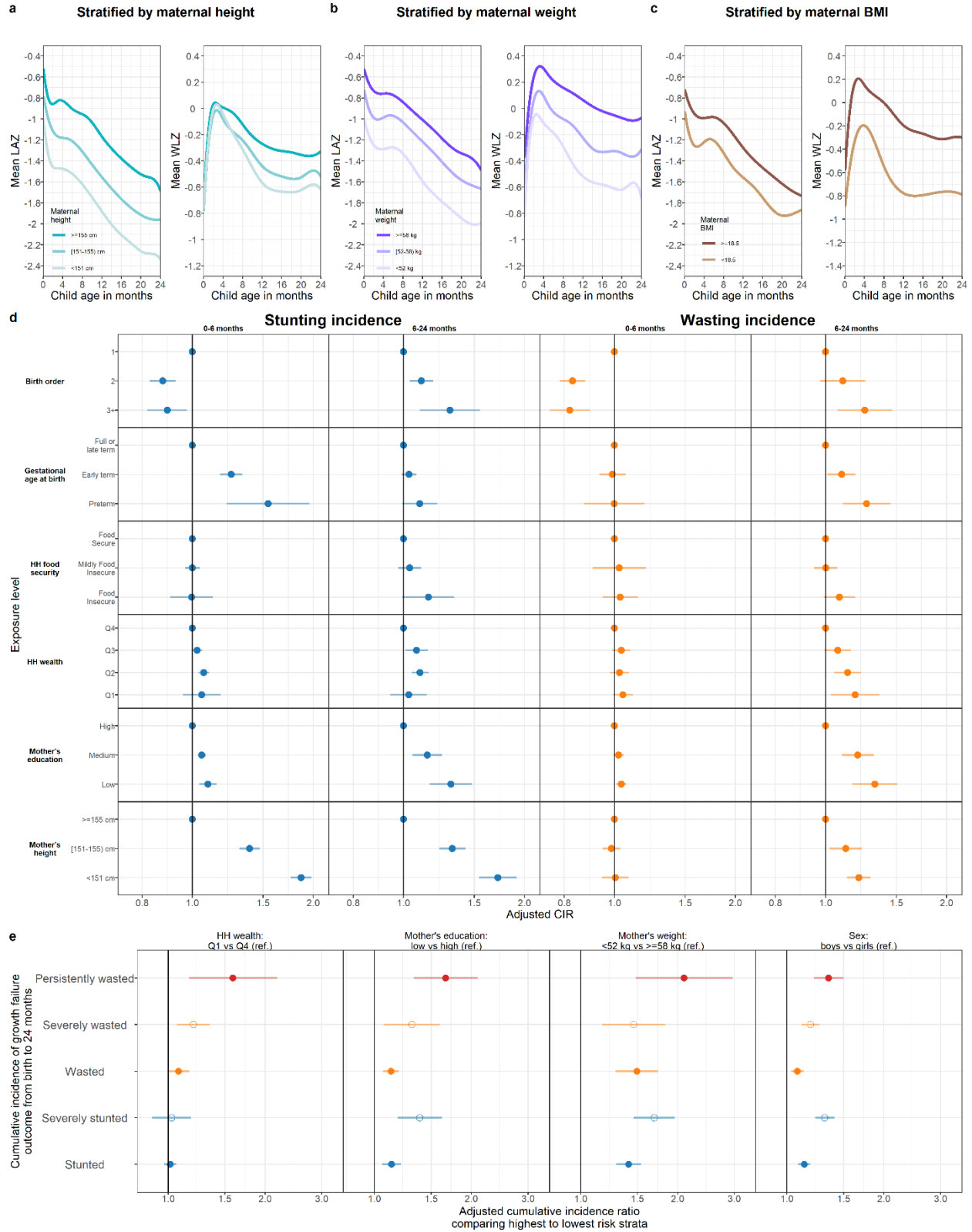
In both panels, reference levels are printed next to the name of the exposure. Estimates are adjusted for all other measured exposures not on the causal pathway.

Maternal anthropometric status can influence child Z-scores by affecting fetal growth and birth size.<sup>143,144</sup> In a secondary analysis, I estimated the association between parental anthropometric status and child Z-scores controlling for child birth characteristics, which showed that the relationship of maternal anthropometric status to child Z-scores was only partially mediated by child birth characteristics (Figure B.5). Maternal weight and BMI could directly affect postnatal health through breastmilk quality, or reflect family poverty, genetics, undernutrition, or food insecurity, or family lifestyle and diet.<sup>145,146</sup>

The strongest predictors of stunting and wasting estimated through population attributable fractions closely matched those identified for child LAZ and WLZ at 24 months (Figure B.6), suggesting that information embedded in continuous and binary measures of child growth provide similar inference with respect to identifying public-health relevant causes. The magnitude of population attributable effects was relatively modest. For example, if all children were born to taller mothers (heights  $\geq 155$  cm) compared to the observed distribution of maternal height, one of the largest predictors of stunting, I estimate it would reduce the incidence of stunting by age 24 months by 19% (Figure B.6a). Patterns in associations across growth outcomes were broadly consistent, except for preterm birth, which had a stronger association with stunting outcomes than wasting outcomes, and rainy season, which had a stronger association with wasting outcomes than stunting outcomes (Figure B.2). Direction of associations did not vary across regions, but magnitude did, notably with male sex less associated with low LAZ in South Asia, and higher parental education and larger household size more associated with higher WLZ in South Asia (Figures B.7, B.8).

### **Age-varying effects on growth failure**

Maternal height and weight consistently arose as key predictors of population attributable differences in child growth failure (Figure 3.2), so I sought to elucidate the longitudinal relationship between maternal anthropometry and child growth. I estimated trajectories of mean LAZ and WLZ stratified by maternal height, weight, and BMI. I found that maternal height strongly influenced at-birth LAZ, but that LAZ progressed along similar trajectories through age 24 months regardless of maternal height (Figure 3.3a), with similar though slightly less pronounced differences when stratified by maternal weight (Figure 3.3b). By contrast, children born to taller mothers had similar WLZ at birth and WLZ trajectories until age 3-6 months, when they diverged substantially (Figure 3.3a); WLZ trajectory differences were even more pronounced when stratified by maternal weight (Figure 3.3b). Maternal BMI strongly influenced WLZ, but not LAZ, at birth (Figure 3.3c). The findings illustrate how maternal status strongly influences where child growth trajectories start, but that growth trajectories evolve in parallel, seeming to respond similarly to postnatal insults independent of their starting point.



### **Figure 3.3 | Effect of key exposures on the trajectories, timing, and severity of growth failure**

**(a)** Child LAZ and WLZ trajectories, stratified by categories of maternal height (N=462,089 measurements, 71,286 children, 23 studies). Mean trajectories estimated using cubic splines in individual studies and then curves were pooled using random effects.<sup>130</sup> Curves estimated from all anthropometry measurements of children taken from birth to 24 months of age within studies that measured the measure of maternal anthropometry.

**(b)** Child LAZ and WLZ trajectories, stratified by categories of maternal weight (N=381,170 measurements, 63,551 children, 17 studies).

**(c)** Child LAZ and WLZ, stratified by categories of maternal BMI (N=373,393 measurements, 61,933 children, 17 studies).

**(d)** Associations between key exposures and wasting cumulative incidence, stratified by the age of the child during wasting incidence, show that many exposures strongly associated with wasting are only associated with wasting occurring after 6 months. The cumulative incidence of stunting or severe stunting from 6-24 months is among children who had not become stunted by 6 months, while the cumulative incidence of wasting or severe wasting from 6-24 months is among children who were not wasted at 6 months or who recovered from a wasting episode after 6 months. Persistent wasting from 6-24 months is among all children with at least 4 measurements between 6 months and 24 months. Children born stunted, and wasting episodes starting at birth, are excluded from the analysis of gestational age.

**(e)** Associations between key exposures and growth failure of different severities, showing that many exposures strongly associated with wasting and stunting have stronger associations with severe stunting and wasting or persistent wasting. Contrasts are between the highest and lowest risk exposure category of each exposure, which are printed in each panel title.

Children who were stunted by age 3 months exhibited a different longitudinal growth trajectory from those who were stunted later (Benjamin-Chung et al. 2020). I hypothesized that causes of growth failure could differ by age of growth failure onset. For key exposures identified in the population attributable effect analyses, I conducted analyses stratified by age of onset and in many cases found age-varying effects (Figure 3.3d). For example, most measures of socioeconomic status were associated with incident wasting or stunting only after age 6 months, and higher birth order lowered growth failure risk under age 6 months, but increased risk thereafter. Stronger relationships between key socio-demographic characteristics and child wasting and stunting as children age likely reflects the accumulation of insults that result from a child's household conditions, particularly as children begin complementary feeding, exploring their environment, and potentially face higher levels of food insecurity in homes with multiple children.<sup>147</sup> When viewed across multiple definitions of child growth failure, most causes had stronger associations with severe stunting, severe wasting, or persistent wasting (> 50% of measurements < -2 WLZ), rarer but more serious outcomes, than with incidence of any wasting or stunting (Figure 3.3e). Additionally, the characteristics strongly associated with lower

probability of recovering from a wasting episode in 90 days (birth size, small maternal stature, lower maternal education, later birth order, and male sex) were also characteristics associated with higher risk of wasting prevalence and cumulative incidence (Figure B.2). Age-specific risk factors were generally similar for different measures of linear growth and wasting (Figure B.2).

### 3.5 Discussion

There were several limitations to this analysis. Measurement frequency and timing varied across cohorts, cohort locations were not geographically balanced, and Z-scores did not adjust for gestational age (Benjamin-Chung et. al 2019).<sup>114</sup> Key exposures such as dietary diversity, nutrient consumption, micronutrient status, maternal and child morbidity indicators, pathogen-specific infections, and sub-clinical inflammation and intestinal dysfunction were measured in only a few cohorts, so were not included.<sup>148,149</sup> The absence of these exposures in the analysis, some of which have been found to be important within individual contributed cohorts,<sup>149,150</sup> means that results emphasize exposures that were more commonly collected, but likely excludes some additional causes of growth failure. Confounders were not measured in every cohort, so there could be residual confounding. Covariate adjustment had a minimal effect on most estimates (Figure B.9), and an unmeasured confounder would on average need to almost double the risk of both the exposure and the outcome to explain away observed significant associations (median E-value: 1.51, Figure B.10).<sup>151</sup>

This large-scale assessment of principal causes and near-term consequences of child growth failure found that maternal, prenatal, and at-birth characteristics are the strongest predictors of growth failure in LMICs. Shifting several key population exposures to their observed low-risk level would modestly improve LAZ and WLZ in target populations and could be expected to prevent 20-30% of incident stunting and wasting and improve Z-scores by 0.2-0.4 Z in the study populations (Figure 3.2, Figure B.6). These results, along with the relative importance of prenatal and maternal exposures, accord with the limited success of numerous postnatal preventive interventions in recent decades.<sup>64,120,152–154</sup> My results suggest that targeting the next generation of interventions toward reproductive age and pregnant women could be a promising path forward to prevent growth failure amongst their children.<sup>155,156</sup> The recent Women's First trial found prenatal nutrition supplements improved children's birth size, though there was no impact of giving supplements starting pre-conception compared to starting late in the first trimester.<sup>157</sup> Emerging evidence suggests that interventions beyond nutrition, such as those that address maternal infection and inflammation, may further contribute to decreasing in utero growth failure.<sup>102,157–161</sup> Nevertheless, a stronger focus on prenatal interventions should not distract from renewed efforts for postnatal prevention. Maternal anthropometric status strongly influenced birth size, but the parallel drop in postnatal Z-scores among children born to different maternal phenotypes was much larger than differences at birth, indicating that growth trajectories were not fully "programmed" at birth (Figure 3.3a-c). Wasting and stunting incidence was highest before age 6 months, but the dangerous concurrence of wasting and stunting peaked at age 18 months (Chapter 2 of this dissertation), and large, seasonally driven declines in WLZ were observed across all ages (Chapter 2 of this dissertation). Targeting postnatal interventions by season or by population subgroups defined by sex, socioeconomic status, maternal, and child birth characteristics identified herein should help focus preventive interventions to reduce the substantial, persistent burden of postnatal growth failure.

## **Conclusion**

Early child wasting was associated with decreases in later linear growth velocity. Child size at birth was strongly affected by maternal anthropometry but linear growth trajectories did not vary greatly by birth size. Markers of higher socio-economic status was primarily associated with a reduction in growth failure after 6 months. Increasing maternal size and child birth size would have the largest population-level impact in reducing measures of growth failure, regardless of region. Taken together, these findings suggest that interventions to prevent child growth failure may have the largest impact if targeted to improve maternal weight prior to birth as well as to improve child birthweight and length.

## 4 Consequences of child growth failure in low resource settings

### Preface

The work presented in this chapter is the product of myself and a large number of collaborators, and we plan on publishing a subset of this research and the research presented in chapter 3 as a single manuscript after submission of this dissertation, with myself as first author. The contributions of collaborators are the same as listed in the preface in Chapter 3.

### 4.1 Background

Growth failure in the form of stunting, a marker of chronic malnutrition, and wasting, a marker of acute malnutrition, is common among young children in low-resource settings. Child growth failure is associated with a higher risk of illness and mortality,<sup>115</sup> which contributed to the United Nations Sustainable Development Goal 2.2 to end malnutrition by 2030. Death from stunting and wasting sometimes occurs through direct starvation but, more commonly, undernutrition is the upstream cause of more proximate causes of death, such as infectious diseases, as undernutrition leads to a weakened immune system.<sup>9</sup> Current estimates attribute > 250,000 deaths annually to stunting and > 1 million deaths annually to wasting.<sup>116</sup> Stunted or wasted children also experience worse cognitive development<sup>1,67,68,162–164</sup> and adult economic outcomes.<sup>141</sup> Current prenatal and postnatal interventions, such as nutritional supplementation, have been insufficient to eliminate growth failure in low resource settings —motivating a search for key age windows and population subgroups in which to focus future preventive efforts. Quantifying the effect of early growth failure on severe outcomes, including child mortality, is important to assess burden and longer-term impacts on the child.

Previous research has found consistent increased risk of childhood mortality with growth failure, regardless of the specific anthropometric measure used.<sup>148,165</sup> A meta-analysis of 10 cohorts found that moderately underweight children (weight-for-age Z-score < -2) had 2.63 times the hazard of death compared to children without growth failure, while moderately stunted children (height-for-age Z-score < -2) had 2.28 times the hazard of death and moderately wasted children (weight-for-height Z-score < -2) had 3.38 times the hazard of death.<sup>10</sup> Risk of death is much higher among children who have multiple measures of growth failure; a prior meta-analysis found that children who were stunted, wasted, and underweight were 12.3 times more likely to die before age 5 years than children without any measure of growth failure.<sup>115</sup>

Chapter 2 of this dissertation highlighted the high incidence of wasting in children under 6 months old, but there is little research on the association between growth failure in young children and subsequent mortality, or on which measure of growth failure is most associated with the risk of death. Children under 6 months of age are often excluded from nutrition programs and surveys because they are assumed to receive adequate nutrition through breastfeeding.<sup>166</sup> However, although limited research has shown a high risk of mortality in children with early growth failure,<sup>167,168</sup> it is unclear which measure of growth failure are the best diagnostic indicators for heightened risk of death. For example, wasting defined by middle-upper arm circumference (MUAC) predicts mortality better than wasting defined by WLZ in older children,<sup>169</sup> and WAZ has been found to be a stronger predictor of mortality

than WLZ or LAZ in children under 6 months.<sup>170</sup>

There may be also regional differences in associations between growth failure and mortality. On an ecological level, South Asia has a higher burden of growth failure compared to Africa, with an estimated 40.7% prevalence of stunting in children under 5 and a 10.3% prevalence of severely wasting, compared to 40.1% stunted and 3.9% severely wasted in Africa.<sup>148</sup> In South Asia, 12.4% of infants are born with intrauterine growth restriction and low birthweight, compared to 8.9% of infants in Africa.<sup>148</sup> However, the under-5 mortality rate is far higher in sub-Saharan Africa (78 per 1000 live births) than in South Asia (42 per 1000 live births).<sup>171</sup> Evidence from south Asia indicates a lower case-fatality rate (1.2-5.2%) of severe wasting (WLZ < -3) than expected from WHO estimates based primarily on sub-Saharan African data (10%–20%).<sup>172–175</sup> There is also evidence that wasting is less of a marker of acute malnutrition in South Asia children compared to sub-Saharan children, as in one study over 40% of children with MAM and SAM remained acutely malnourished for over 3 months,<sup>172</sup> and children in South Asia are more likely to be persistently wasted based on Chapter 2 of this dissertation. An exploration of regional differences in the consequences of wasting and other measures of growth failure will help inform the generalizability of child malnutrition research across regions.

Among children who do not die, growth failure can lead to lifelong morbidity, as early-life growth restriction may be a key developmental origin of adult disease.<sup>3,4</sup> Children who become stunted tend to remain stunted as adults,<sup>67</sup> and wasting episodes often precede the linear growth faltering that leads to stunting.<sup>69</sup> Regardless of adult size, childhood wasting and stunting results in poorer cognitive development and worse educational attainment compared to peers who were never wasted and stunted.<sup>3,4</sup> Stunting is also estimated to reduce adult economic attainment and is associated with increased risk of adult metabolic syndrome.<sup>176,177</sup> Previous research, as well as Chapter 3 of this dissertation, indicate that children born to smaller mothers are more likely to become stunted or wasted.<sup>178</sup> Childhood growth failure may directly cause the long-term morbidity, or be a key marker of poor conditions in early life that lead to both growth failure and other adverse outcomes.<sup>179</sup>

Even if growth failure is not a direct cause of long-term morbidity, it is an important population level indicator of poor nutritional conditions and individual marker of health risk. An understanding of the relative associations between early-life measures of growth failure and later serious growth failure or morbidity will help target growth monitoring programs to the most sensitive indicators of growth failure. Here, I report a pooled analysis of the consequences of early-life measures of growth failure using 35 longitudinal cohorts or randomized control trials with child anthropometry measurements, including eight studies with data on child mortality. Studies were from 15 low- and middle-income countries in South Asia, Sub-Saharan Africa, Latin America, and Eastern Europe, and were initiated between 1969 and 2014.

## 4.2 Objectives

- 1) Estimate if early growth failure impairs subsequent linear growth velocity and if early growth failure is associated with serious growth failure later in childhood.
- 2) Estimate which measures of growth failure have the strongest association with mortality to identify which measures would have the greatest utility to growth monitoring programs as a screening tool for high-risk children.

- 3) Estimate if the risk of severe consequences from early-life growth failure vary by region.

### 4.3 Methods

My objective was to estimate relationships between measures of early life growth failure and subsequent severe growth failure or mortality by age 24 months. Cohorts were assembled as part of the Knowledge Integration (KI) initiative of the Bill & Melinda Gates Foundation which includes a database of millions of participants from studies on childbirth, growth and development. Details on the study selection process and the definitions of different measures of growth failure outcomes are included in Chapter 2 and 3 of this dissertation. The same set of 31 cohorts used in Chapter 3 are used for all analyses of non-mortality outcomes. Four additional studies with high-quality mortality information that measured children at least every 6 months were included in the mortality analyses (The Burkina Faso Zinc trial, The Vitamin-A trial in India, and the iLiNS-DOSE and iLiNS-DYAD-M trials in Malawi, Table C.1).

I estimated unadjusted and adjusted relative risks of severe outcomes, including mortality, between children who experienced measures of early life growth failure and children who did not.

#### **Outcomes:**

All-cause mortality included all deaths among children ages 1-day to 24 months who had anthropometry measurements prior to death.

Concurrent wasting and stunting prevalence at age 18 months was estimated using the anthropometry measurement taken closest to age 18 months, and within 17-19 months of age, and a child was considered concurrently stunted and wasted if they had both a LAZ  $< -2$  and a WLZ  $< -2$ .

Persistent wasting was estimated from child measurements between 6 and 24 months of age, and a child was considered persistently wasted if  $\geq 50\%$  of measurements of WLZ were  $< -2$ .

**Exposures:** I estimated associations between the outcomes listed in above and the cumulative incidence of wasting, stunting, underweight, and concurrent wasting and stunting from birth to 6 months of age, and persistent wasting before 6 months of age. In the analysis of child mortality, I also estimated associations between mortality and the cumulative incidence of wasting, stunting, underweight, and concurrent wasting and stunting from birth to 24 months of age, and persistent wasting before 24 months of age.

**Parameters of interest:** I estimated cumulative incidence ratios, unadjusted and adjusting for all pre-birth exposures, including parental education and age, household wealth, family size, and WASH conditions, as well as sex, birth month, month of measurement, and treatment arm of randomized trials.

I was also interested in determining if low weight-for-length preceded slower linear growth velocity or the onset of stunting. I estimated the difference in linear growth over three-month periods across quartiles of mean WLZ in the prior three-month period and the relative risk

of stunting cumulative incidence over three-month periods across quartiles of mean WLZ in the prior three-month period. I calculated linear growth velocity as the change in length in centimeters within 3-month age intervals, including measurements within a two-week window around each age in months to account for variation in the age of each length measurement.

For missing covariate observations, I imputed missing measurements as the median (continuous variables) or mode (categorical variables) among all children within each cohort, and analyses included an indicator variable for missingness in the adjustment set. When calculating the median for imputation, I used children as independent units rather than measurements so that children with more frequent measurements were not over-represented. To flexibly adjust for potential confounders and reduce the risk of model misspecification, I estimated adjusted PRs, CIRs, and mean differences using targeted maximum likelihood estimation (TMLE), described in more detail in Chapter 3.<sup>60,133</sup> I pooled adjusted estimates from individual cohorts using random effects models, fit using restricted maximum likelihood estimation. The pooling methods are detailed in Chapter 2. Analyses were conducted in R version 3.6.1.

#### 4.4 Results

Thirty-one cohorts from 15 countries met the inclusion criteria with quarterly measurements, and 94,019 children and 645,869 total measurements were included in the analyses of all outcomes other than mortality. Child mortality was rare and not reported in many of the KI datasets, so I relaxed inclusion criteria for studies used in the mortality analysis to include studies that measured children at least twice a year. Four additional studies were included in mortality analyses, with 14,317 children and 70,700 additional measurements (108,336 total children and 716,573 total observations, Table C.1). Cohorts were distributed throughout South Asia, Africa, and Latin America, with a single European cohort from Belarus. I dropped 1,332 (0.2%) unrealistic measurements of LAZ ( $> 6$  or  $< -6$  Z), 1,493 (0.2%) measurements of WAZ ( $> 6$  or  $< -5$  Z), and 1,834 (0.3%) measurements of WLZ ( $> 5$  or  $< -5$  Z), consistent with WHO recommendations.<sup>88</sup> I defined stunting as LAZ  $< -2$ , severe stunting as LAZ  $< -3$ , underweight as WAZ  $< -2$ , severe underweight as WAZ  $< -3$ , wasting as WLZ  $< -2$ , severe wasting as WLZ  $< -3$ , concurrent stunting and wasting as LAZ  $< -2$  and WLZ  $< -2$ , and persistent wasting as  $> 50\%$  measurements of WLZ  $< -2$  during an age period with at least 4 measurements (e.g., birth to 24 months). There were 2,418 deaths before 2 years of age in 57,903 children from 8 cohorts. Figure 4.1 shows the timing of these deaths and the preceding measures of anthropometry.

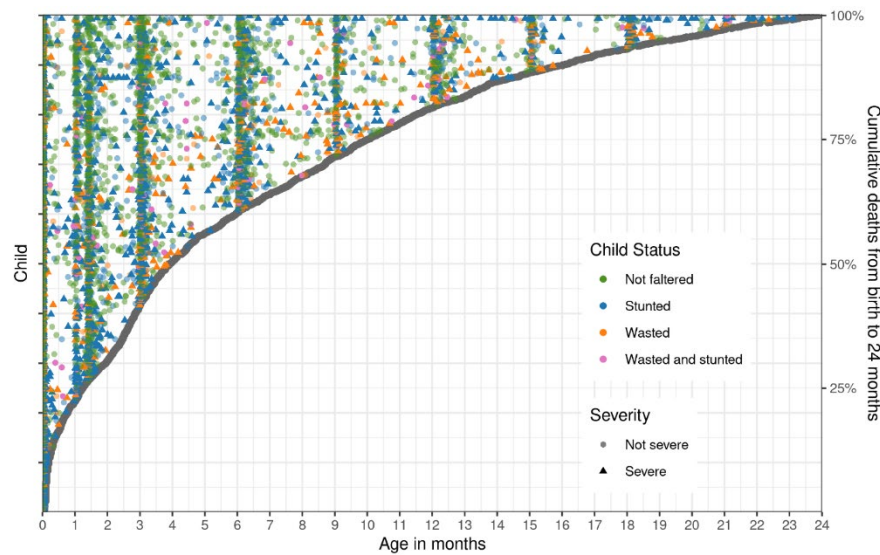


**Figure 4.1 | Age at death and preceding anthropometry measurements among recorded deaths within the included KI cohorts**

Each row is a child, with the furthest right grey dots marking the age at death, and the colored points mark anthropometry measurements prior to death. Children are sorted to show the cumulative distribution of mortality by age 24 months (marked on the right Y-axis), illustrating that more than half of child deaths occurred before 6 months. The figure is based on 2,247 recorded child deaths in children under the age of 2 years from 8 cohorts. Deaths were not plotted for 206 children with missing ages of death.

In Chapter 2 of this dissertation, I documented high incidence rates of wasting and stunting from birth to age 6 months. Individual studies have suggested that early wasting could predispose children to later linear growth failure.<sup>63,69,70,74</sup> I hypothesized that early wasting could contribute to subsequent linear growth restriction, and early growth failure could be consequential for persistent growth failure and mortality during the first 24 months of life. Among cohorts with monthly measurements, I examined age-stratified linear growth velocity by quartiles of WLZ at previous ages. I found a consistent, exposure-response relationship between higher mean WLZ and faster linear growth velocity in the following 3 months (Figure 4.2), with a corresponding inverse relationship between WLZ and incident stunting at older ages (Figure C.1). Persistent wasting from birth to 6 months (defined as > 50% of measurements with WLZ < -2 ) was the wasting measure most strongly associated with incident stunting at older ages (Figure 4.3).

I next examined the relationship between measures of growth failure in the first 6 months and serious growth-related outcomes: persistent wasting from 6-24 months and concurrent wasting and stunting at 18 months of age, both of which put children at high risk of mortality.<sup>3,9,63,75</sup> Concurrent wasting and stunting was measured at 18 months because stunting prevalence peaked at 18 months (Benjamin-Chung et. al 2019) and the largest number of children were measured at 18 months across cohorts. All measures of early growth failure were significantly associated with later, more serious growth failure, with measures of ponderal growth failure amongst the strongest predictors (Figure 4.4-5). Persistent wasting under 6 months

had the strongest association with persistent wasting after 6 months, followed by being moderately underweight (Figure 4.4). Children who were moderately underweight had the highest risk of being concurrently wasted and stunted at 18 months, followed by children who were severely underweight. Unexpectedly, moderate measures of growth failure generally had stronger associations with serious growth-related outcomes than severe measures of the same type of growth failure.

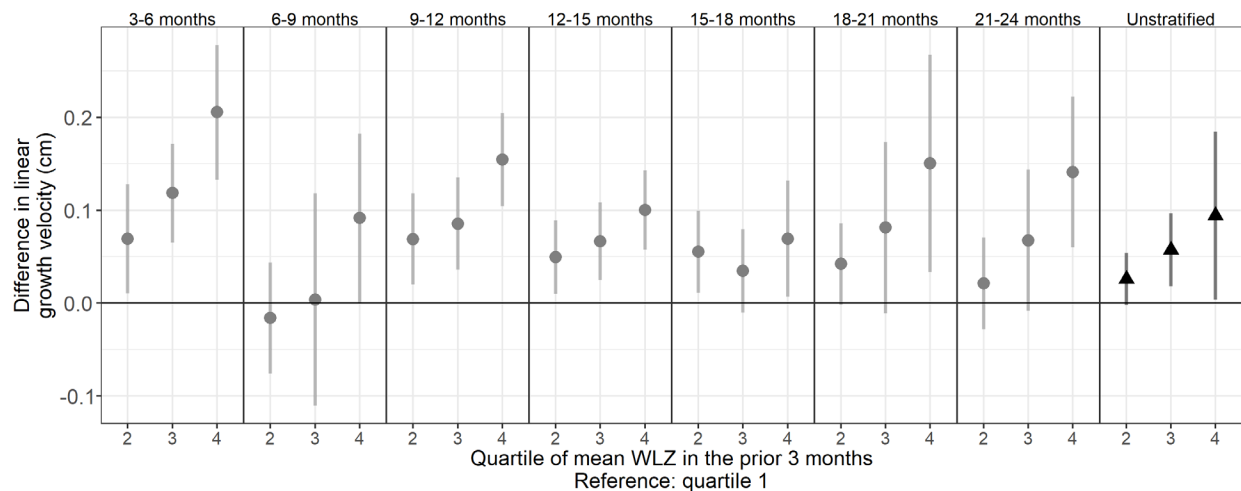


**Figure 4.2 | Adjusted differences in linear growth velocity (in centimeters) across 3-month age bands, by quartile of weight-for-length z-score (WLZ) in the preceding three months.**

The reference group is children in the first quartile of WLZ in the previous age period. At all ages, children with higher mean WLZ over a three-month period had faster linear growth velocity in the next three-month period. The panel with black points on the far right shows the pooled estimates, unstratified by child age. Linear growth velocity was calculated as the change in length in centimeters from the start of an age period to the end of an age period, using the closest anthropometry measurements within 14 days of the start and end of the age period.

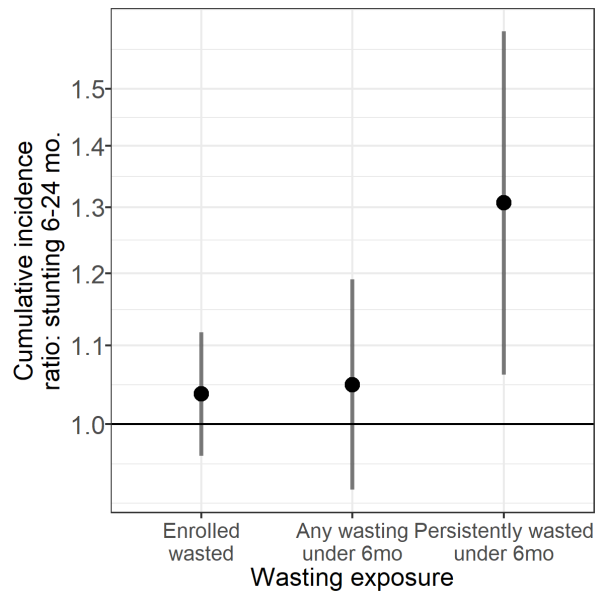


**Figure 4.3 | Relative risk of stunting onset after age 6 months between children who experienced measures of early wasting compared to children who did not experience early wasting.**

The three wasting measures were defined as: child experienced any measure of wasting ( $WLZ < -2$ ) under 6 months, child was enrolled into the study wasted, and child was persistently wasted under 6 months ( $\geq 50\%$  of measurements of  $WLZ < -2$ , and at least 4 measurements prior to age 6 months).

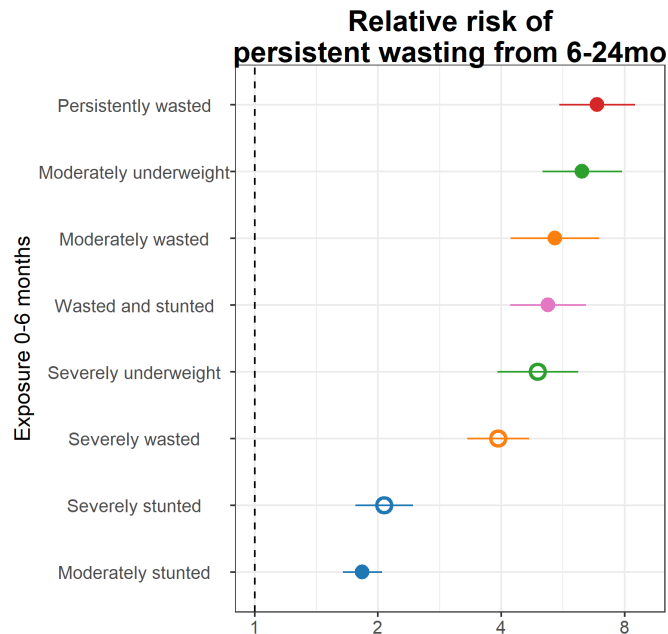
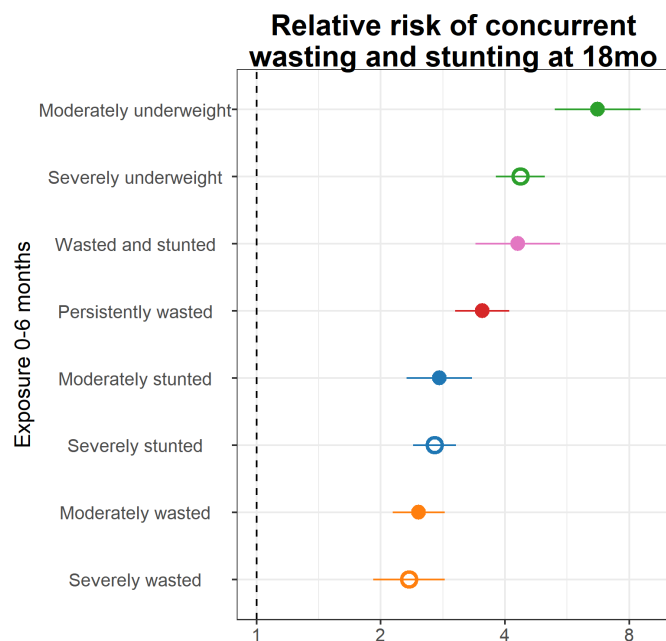


**Figure 4.4 | Risks of persistent wasting due to early life growth failure.**

Relative risk of persistent wasting from ages 6-24 months by measure of growth failure before age 6 months. The estimates include up to 24 cohorts, 7,043 cases, and 76,347 children. Estimates compare risk among children with each measure of growth failure before age 6 months with risk among children who did not experience the specific measure of growth failure. Points are colored by the anthropometry measure, and are open circles if marking a severe measure of growth failure (Z-score < -3) and are closed circles if marking a moderate measure (Z-score < -2).



#### **Figure 4.5 | Risks of concurrent wasting and stunting due to early life growth failure.**

Relative risk of co-occurrent wasting and stunting at age 18 months by measure of growth failure before age 6 months. Estimates included up to 21 cohorts, 1,635 cases, and 21,681 children. Estimates compare risk among children with each measure of growth failure before age 6 months with risk among children who did not experience the specific measure of growth failure. Points are colored by the anthropometry measure, and are open circles if marking a severe measure of growth failure (Z-score < -3) and are closed circles if marking a moderate measure (Z-score < -2).

Finally, I estimated the relative risk of mortality across measures of growth failure in the first 6 months within eight cohorts that reported mortality endpoints, including 2,583 child deaths by age 24 months (4% of children in the eight cohorts). Analyses used all-cause mortality occurring before children turned two years old. Crude case fatality rate among children with early growth failure was high: 8.6% for severe stunting, 7.4% for severe wasting, 10.9% for severe underweight, 8.9% for concurrent wasting and stunting, and 8.1% for persistent wasting. All measures of early growth failure were significantly associated with higher risk of death by age 24 months, and those most strongly associated with death were severely underweight before age 6 months (RR=4.8, 95% CI: 4.1, 5.6), concurrent wasting and stunting (RR=4.8, 95% CI: 3.9, 5.9), and persistent wasting under 6 months (RR=3.4, 95% CI: 3.0, 3.8) (Figure 4.6). Unadjusted estimates were not meaningfully different from adjusted estimates (Figure C.2). Estimates using any measure of growth failure before 24 months were lower than comparable estimates using measures of growth failure before 6 months, except for the measure of persistent wasting between birth and 24 months (Figure C.3). Middle-upper-arm circumference (MUAC), an alternative anthropometric indicator for wasting, has been found to be a better predictor of mortality risk in young children.<sup>169</sup> Within the three included studies that measures MUAC, I compared estimated risk of death between children ever wasted in the first 6 months of life to children not wasted in the first 6 months of life, with wasting either defined using WLZ or MUAC (Figure 4.7). Children with MUAC-defined wasting had a higher risk of death than children with WLZ-defined wasting, with a larger difference in risk between indicators when used to define severe wasting.

Risk of death, persistent wasting, or concurrent wasting and stunting did not consistently vary by region (Figure 4.8). Severe wasting and severe stunting were more associated with mortality in South Asian studies compared to sub-Saharan African studies. However, case fatality rate was much higher in sub-Saharan African studies. Case fatality rate of severe stunting was 12.6% in sub-Saharan Africa versus 7.3% in South Asia, case fatality rate of severe wasting was 11.5% in sub-Saharan Africa versus 4.6% in South Asia, and case fatality rate of severe underweight was 26.1% in sub-Saharan Africa versus 7.5% in South Asia. Early stunting or severe stunting had a stronger association with persistent wasting in Latin America compared to South Asia or sub-Saharan Africa. There were no strong regional differences in associations between measures of early growth failure and concurrent wasting and stunting. Among the three studies with MUAC measurements, there were no clear differences in estimates from the two Jivita trials from Bangladesh and the Kenaba cohort from The Gambia (Figure 4.7).

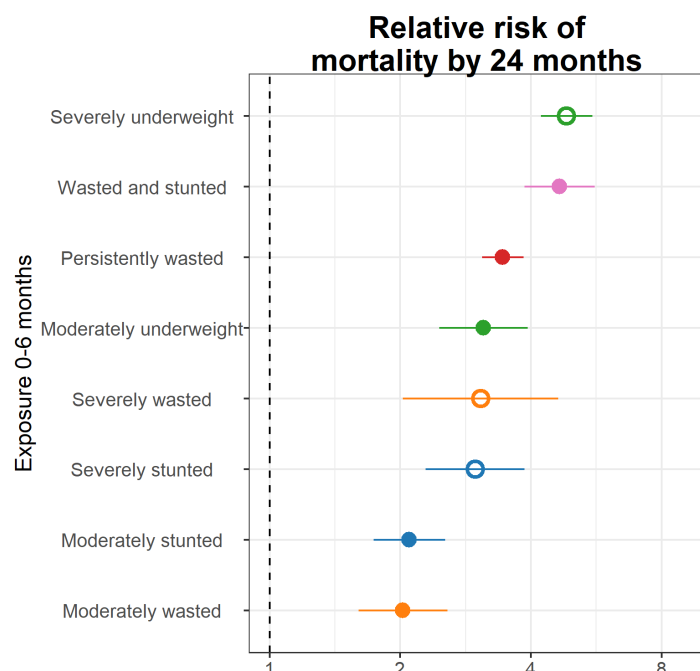
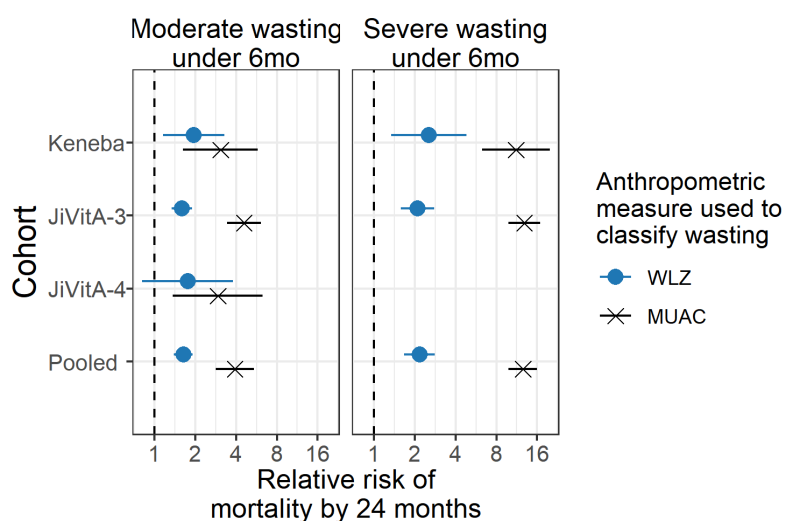


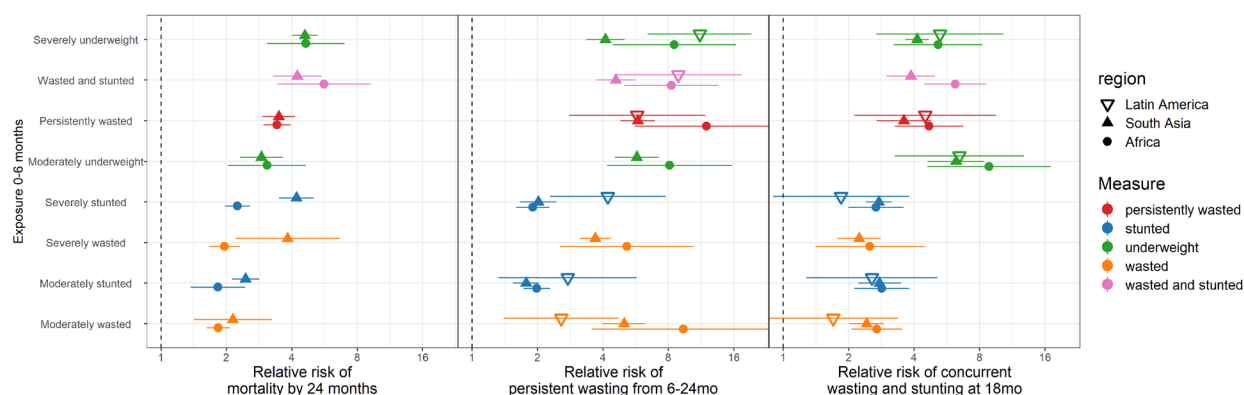
**Figure 4.6 | Risks of mortality due to early life growth failure.**

Relative risk of all-cause mortality by measure of growth failure before age 6 months. Estimates included up to 8 cohorts, 2,418 deaths, and 57,903 children. Estimates compare risk among children with each measure of growth failure before age 6 months with risk among children who did not experience the specific measure of growth failure. Points are colored by the anthropometry measure, and are open circles if marking a severe measure of growth failure (Z-score < -3) and are closed circles if marking a moderate measure (Z-score < -2).



### Figure 4.7 | Risks of mortality by WLZ- or MUAC-defined wasting episodes

Comparison of the estimated adjusted risk of mortality before 24 months among studies that measures both weight-for-length Z-scores (WLZ) and middle-upper-arm circumference (MUAC). Kenaba is a cohort from The Gambia and the two Jivita studies are trials from Bangladesh.



### Figure 4.8 | Risks of mortality and serious growth failure due to early-life growth failure, stratified by region.

Relative risk of all-cause mortality by measure of growth failure before age 6 months, stratified by region. Estimates are from the same set of cohorts as in Figures 4-6. Estimates compare risk among children with each measure of growth failure before age 6 months with risk among children who did not experience the specific measure of growth failure. Points shapes denote the geographic region of the pooled studies, and points are colored by the anthropometry measure, and are open if marking a severe measure of growth failure (Z-score < -3) and closed if marking a moderate measure (Z-score < -2).

## 4.5 Discussion

There were several limitations to this analysis. Included cohorts were highly monitored, so mortality rates were likely lower than in the general population. Without detailed medical histories, in some children growth failure prior to death may have been a sequela of an underlying condition like malaria, diarrheal disease, or severe respiratory infection that caused the death, rather than the cause itself. I also did not analyze when growth failure occurred in relationship to age of death, or whether children recovered from growth failure prior to death due to the disparity in measurement times across studies. Some children may have been born small but fully recovered for two years before death, or some children may have experienced growth failure immediately preceding death, caused by the underlying cause of death. Because mortality was all cause, it may have included deaths like accidents not plausibly influenced by nutritional

status. There was also no data on treatment referral guidelines, so wasted children may have received treatment or nutritional supplementation not available to the local population. Confounders were not measured in every cohort, so there could be residual confounding in the estimates. However, unadjusted estimates of mortality did not vary greatly from adjusted estimates (Figure C.3).

Associations between early-life wasting and mortality may underestimate the risk of death among children born wasted. The 2006 WHO growth standards used to calculate WLZ do not have Z-scores for children with lengths <45cm, so children born with a length of <45 cm were missing WLZ measures until they grew longer than 45cm. Nine and a half percent of included children who were measured in the first week of life had a length measurement below 45cm, and children who were missing WLZ at birth due to low length measurements had 4.26 (95% CI: 3.81, 4.55) times the risk of death before age 2 years than children with WLZ recorded at birth. Although most of these children had follow-up measurements before 6 months, only 12% were wasted in subsequent measurements before 6 months. Additionally, associations between measures of growth failure other than wasting and mortality did not vary meaningfully after dropping children without measurements in the first month of life, but the relative risks of mortality estimate for wasting and severe wasting decreased (Figure C.4). This indicates that children missing birth measures of wasting were both more likely to be wasted after one month of age than children with birth measures and were more likely to die before 24 months of age. It is possible, therefore, that the associations between early wasting and mortality are underestimates due to limitations in WHO growth standards for infants, and are biased downwards compared to associations between stunting or underweight and mortality, as the range of LAZ and WAZ standards are not limited at birth.

The associations between growth failure and mortality estimated in this chapter are consistent with a prior IPD meta-analysis of 10 cohorts.<sup>165</sup> Both found significant increased risk of death for all measures of growth failure, with risk increasing with severity (lower Z-scores). I found that the risk of death doubled in moderately stunted children compared to non-stunted children, similar to their estimate of 2.28 times the hazard. However, their study compared children under 5 with Z-scores between -2 and -3 to children under 5 with Z-scores > -1, while I compared children under 6 months with Z-scores < -2 to children under 6 months with Z-scores > -2, so the estimates are not directly comparable. They also estimated that wasting had the strongest association with mortality of anthropometric indicators examined (HR: 3.38), while I found the underweight had the strongest association (RR: 3.0). These results support prior findings that underweight may be a better indicator of risk of death in children under 6 months, while wasting is a better indicator in older children.<sup>172,180</sup> Wasting defined by MUAC may be more predictive indicator of mortality than underweight, though MUAC was only measured in 3 studies and so further research is needed in young children. WAZ may be a more pragmatic screening tool than MUAC as many existing growth monitoring programs only measure child weight.<sup>181</sup>

I found that lower WLZ preceded slower linear growth velocity, and wasting preceded stunting, which indicates that stunting may be preventable by avoiding periods of acute malnutrition caused by hunger or malnutrition.<sup>3,69</sup> Measures of early growth failure were all significantly associated with persistent wasting and concurrent wasting and stunting, two indicators of serious growth failure with strong associations with mortality. This analysis was

conducted in a much larger set of studies than the mortality analysis (35 studies versus eight studies), and it supports the results of the mortality analysis in a broader range of settings. The primary difference from the results of the mortality analysis was that the RRs associated with moderate stunting were larger than the RRs associated with severe stunting, and the RRs associated with moderate wasting were larger than the RRs associated with severe wasting. Further research could explore if severe cases are more likely to be misclassified due to measurement error, attenuating estimates towards the null. Alternatively, severe cases are more likely to be acute and reversible episodes, while a subset of moderate cases capture intrinsically small children without the growth potential to catch up to mean growth standards, and therefore they are more likely to remain wasted or both wasted and stunted.

Prior research has highlighted the lower-case fatality rate of severe acute malnutrition in South Asia and higher prevalence of malnutrition indicators than sub-Saharan Africa.<sup>172,172</sup> I also found a higher case-fatality of wasting in sub-Saharan Africa compared to South Asia, and the case-fatalities observed in the KI studies are consistent with prior research. Case fatality rate of severe wasting was 11.5% in sub-Saharan African studies, comparable to the 10-20% World Health Organization estimate derived primarily from African studies, versus 4.6% in South Asia, comparable to the 4% estimates from 4 Indian cohorts. Therefore, although a case of wasting is of more absolute concern in sub-Saharan Africa, preventing early wasting in South Asia should still be a key target of public health programs due to the relative increase in risk of death and because of the prevalence of wasting in young children in South Asia (Chapter 2).

In conclusion, children who experienced early ponderal or linear growth failure were at much higher risk of persistent growth failure, later concurrent wasting and stunting, and were 2.0 to 4.8 times more likely to die by age 24 months. Screening based on weight could help identify children at highest risk of death before age 24 months, as weight-for-age Z-scores had strongest association with mortality and only require one measurement of weight, which is often the only anthropometry recorded in growth monitoring programs.<sup>181</sup> The high risk of severe consequences for children who experienced early growth failure indicates that pre-conception and pregnancy are key windows of opportunities to intervene with new preventive interventions. Prevention of early growth failure before age 6 months should be a high priority for global health programs because I observed that all measures of early life growth failure puts children at substantially higher risk of death by age 24 months, regardless of location, and preventing the mortality and long-term morbidity associated with child growth failure is one of the most impactful targets to prevent loss of life and disability-adjusted life-years.<sup>66</sup>

## **Conclusion**

Prevention of growth failure during the first 6 months of life should be a high priority for global health programs because of its demonstrated, strong relationship with mortality during the birth to 24-month period. Child underweight, as measured by weight-for-age Z-scores had the strongest association with mortality among the analyzed measures of growth failure, and a stronger association with indicators of severe growth failure than wasting or stunting. WAZ should be the anthropometry measurement used to screen for high-risk children under 6 months. There were no strong regional differences in associations, indicating that early growth failure is of concern regardless of location.

# 5 Conclusion

## 5.1 Key Findings

Findings of *Chapter 2: Longitudinal analysis of child wasting and co-occurrence with stunting in low-resource settings*

### **Objective 1: Patterns in wasting over time**

Acute malnutrition accounts for an immense disease burden and is implicated as a key, underlying cause of child mortality in low resource settings. Child wasting is a leading indicator to measure the Sustainable Development Goal target to end malnutrition by 2030. Prevailing methods to measure wasting rely on cross-sectional surveys that are unable to measure onset, recovery, and persistence — key features of wasting epidemiology that could inform preventive interventions and disease burden estimates.<sup>148</sup> Of particular interest in my results are the estimates of age-specific incidence, the duration of episodes, how often children spontaneously recover from wasting without intervention, and strongly seasonal patterns influence population mean WLZ. I was able to leverage the unprecedented library of individual participant data across the KI longitudinal studies to better describe the course of wasting episodes and compare patterns across countries and across child ages.

In Chapter 2, I showed through an analysis of 18 longitudinal cohorts that child wasting is a highly dynamic process of incident onset and recovery, and that peak incidence is between birth and 3 months — far earlier than peak prevalence at 12-15 months. By age 24 months the proportion of children who had ever experienced a wasting episode (33%) was more than 5-fold higher than prevalence (6%), suggesting that the wasting burden is likely far higher than cross-sectional surveys suggest. Treatment programs have traditionally focused on children from 6-59 months. These results, however, motivate a new focus on extending preventive interventions for child wasting to pregnant and lactating mothers, and for preventive and therapeutic interventions to include children below age 6 months in addition to current targets of ages 6-59 months.

### **Objective 2: Concurrence of wasting and stunting**

Comorbid stunting and wasting is associated with a much higher risk of mortality but is too rare to examine in most existing studies. Children who are both stunted and wasted have 12.3 times the risk of death compared to children who are neither wasted nor stunted.<sup>9</sup> The WHO does not estimate the prevalence of co-occurrent wasting and stunting due to data limitations, but I was able to examine both concurrent wasting and stunting, and the longitudinal patterns in wasting and stunting onset prior to co-occurrence to identify if the acute loss of nutrients and energy from a wasting episode diverts energy from growth and causes the onset of stunting.<sup>6</sup> Although co-occurring stunting and wasting is rare, it is an understudied subgroup of children with severe morbidity who may receive the most benefit from targeted intervention. I found that children who were wasted very young were also far more likely to experience co-occurrent wasting and stunting and to become persistently wasted during their first 24 months. Although

wasting and stunting have often been addressed in isolation, the majority of wasted children older than one year were also stunted, so wasting treatment should be a key component of stunting prevention efforts.

### **Objective 3: Seasonality of wasting**

I found high seasonality in weight-for-length Z-scores, and population mean Z-scores varied more in countries with high amount of seasonality, and seasonal changes were closely tied to rainfall across diverse settings. Child weight may fluctuate over the course of the year due to seasonal food scarcity or increases in food prices,<sup>182,183</sup> or seasonal patterns in infection.<sup>184,185</sup> I found that child birthweight varied across season as well, indicating that maternal nutrition was impacted by season as well. Maternal infection or food insecurity can both influence birthweight,<sup>186</sup> but I did not have the data to explore the relative contribution of infection or food scarcity on either birthweight or seasonal trends in WLZ. Interventions should be delivered prior to seasonal peaks in wasting to prevent the increase, possibly through monitoring and treating infections, providing nutritional supplements, or providing financial assistance to offset increases in food prices.

## *Findings of Chapter 3: Causes of child growth failure in low resource settings*

### **Objective 1: Rank order causes of child growth failure**

Numerous studies have examined risk factors for stunting<sup>73,110,187–193</sup> and wasting.<sup>46,71,73,110,190,191,194–197</sup> A key limitation of most large-scale risk factor studies is that they have drawn predominantly on cross-sectional surveys to elucidate relationships. Cross-sectional studies fail to capture the dynamic process of acute malnutrition, cannot guarantee that an exposure precedes an outcome, and by virtue of design cannot measure more severe outcomes of public health importance such as persistent wasting (Chapter 2). Careful studies of longitudinal cohorts have summarized key characteristics associated with growth failure,<sup>27</sup> including diarrheal disease, wasting, season, and child sex, but none pooled multiple cohorts, and to my knowledge most did not estimate and compare the reduction in growth failure that would occur if a risk factor was eliminated from a population (i.e., the population-attributable risk).<sup>46</sup> This measure is arguably most relevant for targeting preventive public health efforts because they combine both the strength of a potential cause with its prevalence in the population.<sup>198</sup> To my knowledge, there have not been other large-scale efforts to harmonize exposure and outcome information across numerous longitudinal cohorts to generate pooled estimates of effect to inform intervention development and deployment.

The primary proximate cause of stunting and wasting is inadequate nutrition, but upstream causes are complex. Broadly, previous research suggests stunting and wasting share the same set of risk factors,<sup>71–73</sup> with generally stronger associations with stunting than with wasting, and with the highest burden of both conditions in South Asia,<sup>61</sup> although there is substantial within-region variation.<sup>82,191,199</sup> The regional variation is driven by socio-economic (SES), cultural differences,<sup>200–204</sup> the urban/rural divide,<sup>110,187,190</sup> and climactic differences.<sup>190,191,195</sup> Within populations, a large set of biological and SES risk factors are associated with stunting. Strong biological determinants of post-natal stunting are inter-uterine growth restriction or preterm

birth.<sup>46,72,188,190,193,205</sup>, small parental anthropometry<sup>188</sup>, being male,<sup>187,190,195,206–208</sup> mother's age,<sup>188,190</sup> short birth intervals,<sup>188</sup> and birth order<sup>190</sup>. Post-birth risk factors found in prior research include cooking fuel/air pollution,<sup>110,196,207</sup> illness,<sup>190,191,193,195,208–210</sup> and environmental enteric disorder,<sup>124</sup> household wealth,<sup>71,110,190,207,208,211</sup> parental education<sup>73,110,187,190,197,207,208,212</sup> and employment,<sup>73,73,187</sup> household size,<sup>73,187,195,207</sup> and water, sanitation, and hygiene (WASH) characteristics.<sup>73,110,187–193</sup> Lastly, as stunting is a measure of chronic malnutrition, recommended breastfeeding practices,<sup>110,190,213</sup> dietary diversity,<sup>110,193,208</sup> and food security<sup>191,195,214</sup> have also been linked to lower risks of stunting, and prior wasting may increase the risk of stunting.<sup>63,69,70</sup>

The results from Chapter 3 are consistent with this previously published literature. I found that parental anthropometry, sex, and birth size are the strongest determinants of later growth failure. Sex can't be modified, and parental height can't be modified during pregnancy, but these factors would be useful to screen children at the highest risk of faltering for monitoring or intervention. Parental anthropometry influences the risk of growth failure as a direct consequence of birth size, but the strong association is also related to other social and environmental factors as the associations between mother's height and weight and the cumulative incidence of stunting and wasting were still significant after excluding children stunted or wasted at birth. Maternal anthropometry had a stronger association than paternal anthropometry, which was expected as maternal stature effects gestational weight gain and breastmilk quality, as well as being an indicator for family SES.<sup>140,144,215</sup>

Child sex cannot be modified, but a better understanding of what increases boys' risks of failure may help development of sex-specific interventions. In particular, boys have been found to have a higher severity of most infectious diseases.<sup>216</sup> However, there were not strong associations between longitudinal prevalence of diarrhea and stunting or wasting, though this analysis was limited to a few studies with frequent monitoring and only captures all-cause diarrhea.

Both clean cooking fuel and improved sanitation had strong associations with many growth outcomes, consistent with prior findings,<sup>73,110,187–193,196,207</sup> and both may increase growth failure by increasing infections. However, both cooking fuel usage and WASH characteristics may just be markers for wealth rather than a strong, direct cause of growth failure, as a systematic review<sup>217</sup> and several large randomized control trials found limited to no effect of WASH on reducing stunting or wasting.<sup>80,81,125,218</sup>

The post-natal characteristics of exclusive or predominant breastfeeding and household food security were significantly associated with stunting and wasting, but the strengths of associations were small. Breastfeeding promotion and nutritional supplements are common interventions to reduce wasting and stunting that have had disappointing results,<sup>7</sup> but targeting these interventions to the children at the highest risk of failure may increase the treatment effects.

The population attributable difference estimates from the optimal intervention analyses did not vary greatly from the equivalent static population attributable difference estimates, implying that the low-risk level of exposures did not vary across different children (Figure 3.2). For some variables, like birthweight or diarrheal disease, it would have been hard to interpret if the low-risk level varied across children, but for other variables like birth order I did not have an a-priori hypothesis of which level was lowest risk, and it would not have been surprising if the low-risk level of birth order varied by family education level or household food security. However, because most child, parental, and household characteristics had a clear, universal low risk level,

and due to the additional computing time needed to conduct the optimal intervention analysis, the optimal intervention analysis did not improve the rank ordering of causes of growth failure.

### **Objective 2: Estimating variation in associations by region, child age, measure of growth failure, or severity of growth failure.**

Predominantly, the direction of associations did not vary by region, child age, or measure of growth, but the magnitude did. Male sex was less associated with low LAZ in South Asia compared to other regions. The smaller gender differences in LAZ in south Asia may be due to cultural preferences for boys; prior research found that girls, but not boys, in south Asia were more likely to become stunted the more brothers they had.<sup>219</sup> Possibly, nutritional interventions in south Asia would be more effective if supplemented with interventions grounded in behavioral science designed to shift cultural norms around female child feeding practices.

While most exposures associated with growth failure were associated with growth failure at all ages (Figures B.3, B.4), several measures of socioeconomic status, including household wealth, household size, household food security, and maternal education, were more associated with wasting or stunting in older children (Figure 3.3d, B.3, B.4). These age-specific findings likely reflect the protection infants have from poor nutritional, WASH, and household condition while exclusively breastfeeding and prior to crawling. Later child birth order was protective of wasting and stunting before 6 months, but a risk factor for wasting and stunting after 6 months. More research is needed to confirm and explore the drivers of this finding, but possibly mothers are more experienced in infant care with non-firstborn infants, while there is insufficient food for older non-firstborn children or increased infectious disease transmission from their siblings. Potentially, in low-resource settings, resource-constrained public health organizations can increase impact by targeting new, first-time mothers with educational support, and by targeting older, non-firstborn children with nutritional support. There were no strong differences in risk factors for wasting and stunting, except that the three month period of highest rainfall was strongly associated with wasting, but not stunting, in accordance with the acute nature of wasting and chronic nature of stunting.

### **Objective 3: The relative importance of child size at birth compared to postnatal growth**

Pooling spline curves of child growth trajectories across cohorts showed a consistent relationship with maternal size (Figure 3.3a-c). Maternal height, weight, and BMI affects child birth WLZ and LAZ, and the slopes of trajectories do not vary greatly regardless of birth size. Children who have smaller mothers tend to be born smaller, and on average they did not catch up to their peers by two years of age. Mean LAZ decreased consistently from birth in each category of maternal size, while WLZ increased from birth to 3-4 months of age before children's WLZ trajectories diverged slightly. A pessimistic interpretation of these patterns is that parental anthropometry determines children's birth size, and catch-up growth does not occur after birth. However, the mean drop in Z-scores over the first two years of life was much larger than the differences at birth, implying that the local conditions that cause growth faltering outweighed the impact of maternal size. If these children were raised in a high-income country, or if

interventions are developed that prevent growth faltering, the growth trajectories might be flat from birth (for LAZ) or from the 3-month peak (for WLZ), or catch-up growth may occur. The challenge will be developing the right combination of nutritional, WASH, educational, and economic interventions that prevents the post-birth growth failure even in children born small.

#### Findings of *Chapter 4: Consequences of child growth failure in low resource settings*

##### **Objective 1: Associations between early growth failure and subsequent linear growth velocity and serious growth failure later in childhood.**

While wasting is commonly considered a measure of acute malnutrition, and stunting is considered a measure of chronic malnutrition, having a lower WLZ was associated with subsequent slower linear growth velocity and later stunting. Possibly, periods of acute malnutrition from seasonal food scarcity or infection are sufficient to suppress linear growth enough that catch-up growth does not occur or is insufficient to prevent later stunting. Preventing seasonal wasting through targeted delivery of food supplementation (and antibiotics if wasting is complicated by infection) might be an effective public-health intervention as it may both reduce mortality from acute malnutrition as well as reduce long-term growth faltering.

The all measures of early growth failure were significantly associated with mortality, concurrent wasting and stunting at 18 months, and persistent wasting after 6 months. Due to the strong association between early concurrent wasting and mortality and early persistent wasting and mortality, as well as the difficulty in collecting mortality information in low-resource settings, these measures could be used as proxies for later severe health outcomes. As the incidence of these outcomes is higher than mortality, and can be measured in all children at a standardized time, these outcomes could be used in randomized trials as surrogate outcomes for mortality, reducing trial sizes, durations, and costs. However, more research is needed into why moderate measures of early growth failure had stronger associations with these surrogate outcomes than severe measures.

Measures of early growth failure were all significantly associated with persistent wasting and concurrent wasting and stunting, two indicators of serious growth failure with strong associations with mortality. This analysis was conducted in a much larger set of studies than the mortality analysis (35 studies versus eight studies), and it supports the results of the mortality analysis in a broader range of settings. The primary difference from the results of the mortality analysis was that the RRs associated with moderate stunting were larger than the RRs associated with severe stunting, and the RRs associated with moderate wasting were larger than the RRs associated with severe wasting.. Further research could explore if severe cases are more likely to be misclassified due to measurement error, attenuating estimates towards the null. Alternatively, severe cases are more likely to be acute and reversible episodes, while a subset of moderate cases capture intrinsically small children without the growth potential to catch up to mean growth standards, and therefore they are more likely to remain wasted or both wasted and stunted.

##### **Objective 2: Associations between growth failure and mortality.**

All measures of growth failure had strong associations with mortality, reinforcing growth failure as a key public health concern, and underweight was a stronger predictor than stunting or wasting, so WAZ should be used to screen children at high risk, especially in children under 6 months. Early growth failure had a strong association with mortality, and early wasting preceded the onset of stunting, but most exposures examined in chapter 3 were not associated with wasting in the first 6 months of life. Maternal weight was a strong predictor of early wasting, but a better understanding of the causes of early wasting is needed both for screening and prevention. Wasting defined by middle-upper-arm circumference had a stronger association with mortality than WLZ in the three studies that measured both MUAC and mortality. These findings support, with a larger and more diverse set of children, previous research from a Burkina Faso birth cohort which found that low MUAC, low weight, and in calculable WLZ at birth were associated with mortality.<sup>220</sup>

Growth monitoring programs should therefore use WAZ or MUAC, rather than WLZ, to screen children under 6 months for those at high risk of death or later growth failure. World Health Organization growth standards should be updated to avoid missing WLZ estimates when length is below 45cm. Otherwise, the burden of wasting in young children will be underestimated when measured by WLZ in populations with a high burden.

### **Objective 3: Regional differences in the risk of severe consequences from early-life growth failure.**

Previous research has shown a lower fatality of wasting in south Asia compared to similar studies in sub-Saharan Africa, possibly due to different average body composition between regions influencing what a healthy WLZ is in that region, but my results did not fully support that.<sup>107,108,221</sup> While case-fatality of wasting was much higher in sub-Saharan African studies than south Asian studies, the background child mortality in the sub-Saharan African studies was also higher, and I found similar relative risks between measures of early growth failure and mortality in south Asian cohorts compared to sub-Saharan African cohorts. Within the three studies measuring mortality and MUAC, the magnitude in the differences in associations between mortality and wasting defined with WLZ and wasting defined MUAC were roughly the same, even though one study was from The Gambia and two were from Bangladesh. As MUAC and WLZ measure different facets of body composition, this provides evidence that average body composition did not differ between the two countries in a manner that affects the risk of death from wasting. Also, case-fatality from growth failure before 6 months of age was higher in the data I analyzed than in estimates in older children from other studies, and I also found a higher risk of death from measures of growth failure before 6 months compared to the same measures before 24 months, both of which indicate earlier growth failure increases risk of mortality.<sup>221</sup> Though more progress needs to be made in reducing child mortality in Africa due to the higher rate, as early wasting, stunting, and underweight incidence was highest in south Asian cohorts, preventing early growth failure in south Asia through prenatal or post-natal interventions is also imperative to reduce child mortality.

## 5.2 Scientific Conclusions

The dissertation responds to the urgent need to ensure that all children start their life course healthy and on equal footing. There are wide disparities between high income countries and LMIC in wealth, health, and educational attainment. These arise partly from structural differences in government and economies, but also the disadvantage incurred by stunting and wasting at such young age, including missed schooling, increased infection, and poor cognitive development.<sup>11</sup> Reducing stunting may allow children born in LMIC a more equal start to life and a better chance to thrive. Additionally, preventing the direct mortality and morbidity caused by wasting and stunting is one of the most impactful targets to prevent loss of life and disability-adjusted life-years (DALYs).<sup>66</sup>

The analyses presented in this dissertation use, to my knowledge, the largest longitudinal dataset of stunting and wasting ever assembled, with standardized exposure and outcome definitions to reduce bias and enough data to analyze rare subgroups. The exploration of the timing and location of wasting episodes helps inform when and where intervention efforts can have the most impact, and the systematic analysis of conditions that precede episode onset will help target preventative interventions to the children at highest risk.

I found that wasting onset occurs earlier than prior research had indicated, and children wasted early have a high risk of linear growth faltering and becoming both wasted and stunted. Children were more likely to be wasted at birth in south Asia, but all measures of early growth failure were strongly associated with mortality in all regions. My findings indicate an intergenerational trap of growth failure; children of smaller mothers are more likely to be born small, and children born small are more likely to die or become permanently stunted. Increasing maternal size and child birth size would have the largest population-level impact in reducing measures of growth failure among risk factors considered, so interventions to prevent child growth may have the largest impact if delivered earlier, including in utero. Additionally, interventions targeted to reduce rainy-season peaks in wasting may both decrease the direct mortality caused by wasting as well as decrease long-term linear growth failure.

## References

1. Black, R. E. *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* **382**, 427–451 (2013).
2. Humphrey, J. H. & Prendergast, A. J. Population-level linear growth faltering in low-income and middle-income countries. *Lancet Glob. Health* **5**, e1168–e1169 (2017).
3. Briend, A., Khara, T. & Dolan, C. Wasting and stunting--similarities and differences: policy and programmatic implications. *Food Nutr. Bull.* **36**, S15-23 (2015).
4. Chang, C. Y. *et al.* Children successfully treated for moderate acute malnutrition remain at risk for malnutrition and death in the subsequent year after recovery. *J. Nutr.* **143**, 215–220 (2013).
5. Briend, A. & Prinzo, Z. W. Dietary management of moderate malnutrition: time for a change. *Food Nutr. Bull.* **30**, S265-266 (2009).

6. WHO | UNICEF-WHO-The World Bank: Joint child malnutrition estimates - Levels and trends. *WHO* <http://www.who.int/nutgrowthdb/estimates/en/>.
7. Bhutta, Z. A. *et al.* What works? Interventions for maternal and child undernutrition and survival. *Lancet* **371**, 417–440 (2008).
8. KI Project. *HBGDki* <http://kiglobalhealth.org/data/>.
9. McDonald, C. M. *et al.* The effect of multiple anthropometric deficits on child mortality: meta-analysis of individual data in 10 prospective studies from developing countries. *Am. J. Clin. Nutr.* **97**, 896–901 (2013).
10. Wamani, H., Åström, A. N., Peterson, S., Tumwine, J. K. & Tylleskär, T. Boys are more stunted than girls in Sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys. *BMC Pediatr.* **7**, 17 (2007).
11. Black, R. E. *et al.* Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet* **371**, 243–260 (2008).
12. Imdad, A. & Bhutta, Z. A. Effect of preventive zinc supplementation on linear growth in children under 5 years of age in developing countries: a meta-analysis of studies for input to the lives saved tool. *BMC Public Health* **11**, S22 (2011).
13. Lenters, L. M., Wazny, K., Webb, P., Ahmed, T. & Bhutta, Z. A. Treatment of severe and moderate acute malnutrition in low- and middle-income settings: a systematic review, meta-analysis and Delphi process. *BMC Public Health* **13**, S23 (2013).
14. Smith, E. R. *et al.* Modifiers of the effect of maternal multiple micronutrient supplementation on stillbirth, birth outcomes, and infant mortality: a meta-analysis of individual patient data from 17 randomised trials in low-income and middle-income countries. *Lancet Glob. Health* **5**, e1090–e1100 (2017).
15. Simmonds, M., Stewart, G. & Stewart, L. A decade of individual participant data meta-analyses: A review of current practice. *Contemp. Clin. Trials* **45**, 76–83 (2015).
16. Gewin, V. Data sharing: An open mind on open data. *Nature* **529**, 117–119 (2016).
17. Riley, R. D., Lambert, P. C. & Abo-Zaid, G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* **340**, c221 (2010).
18. Huang, Y. *et al.* Distribution and epidemiological characteristics of published individual patient data meta-analyses. *PloS One* **9**, e100151 (2014).
19. Ohmann, C. *et al.* Sharing and reuse of individual participant data from clinical trials: principles and recommendations. *BMJ Open* **7**, e018647 (2017).
20. Breastfeeding and HIV International Transmission Study Group *et al.* Late postnatal transmission of HIV-1 in breast-fed children: an individual patient data meta-analysis. *J. Infect. Dis.* **189**, 2154–2166 (2004).
21. Morrison, C. S. *et al.* Hormonal Contraception and the Risk of HIV Acquisition: An Individual Participant Data Meta-analysis. *PLOS Med.* **12**, e1001778 (2015).

22. Becquet, R. *et al.* Children who acquire HIV infection perinatally are at higher risk of early death than those acquiring infection through breastmilk: a meta-analysis. *PloS One* **7**, e28510 (2012).
23. International Perinatal HIV Group *et al.* The mode of delivery and the risk of vertical transmission of human immunodeficiency virus type 1--a meta-analysis of 15 prospective cohort studies. *N. Engl. J. Med.* **340**, 977–987 (1999).
24. Kosack, C. S. *et al.* Towards more accurate HIV testing in sub-Saharan Africa: a multi-site evaluation of HIV RDTs and risk factors for false positives. *J. Int. AIDS Soc.* **20**, 21345 (2017).
25. International Perinatal HIV Group. Duration of ruptured membranes and vertical transmission of HIV-1: a meta-analysis from 15 prospective cohort studies. *AIDS Lond. Engl.* **15**, 357–368 (2001).
26. Kovalchik, S. A. Mother's CD4+ Count Moderates the Risk Associated with Higher Parity for Late Postnatal HIV-Free Survival of Breastfed Children: An Individual Patient Data Meta-Analysis of Randomized Controlled Trials. *AIDS Behav.* **16**, 79–85 (2012).
27. Li, J. Z. *et al.* Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. *JAMA* **305**, 1327–1335 (2011).
28. Low, N. *et al.* Intravaginal practices, bacterial vaginosis, and HIV infection in women: individual participant data meta-analysis. *PLoS Med.* **8**, e1000416 (2011).
29. Raboud, J. M. *et al.* A Meta-Analysis of Six Placebo-Controlled Trials of Thiazolidinedione Therapy for HIV Lipodystrophy. *HIV Clin. Trials* **11**, 39–50 (2010).
30. Raju, R., Peters, B. S. & Breen, R. A. M. Lung infections in the HIV-infected adult. *Curr. Opin. Pulm. Med.* **18**, 253 (2012).
31. Torrone, E. A. *et al.* Prevalence of sexually transmitted infections and bacterial vaginosis among women in sub-Saharan Africa: An individual participant data meta-analysis of 18 HIV prevention studies. *PLOS Med.* **15**, e1002511 (2018).
32. Zwahlen, M. & Egger, M. Progression and mortality of untreated HIV-positive individuals living in resource-limited settings: update of literature review and evidence synthesis. **17**.
33. Getahun, H. *et al.* Development of a standardized screening rule for tuberculosis in people living with HIV in resource-constrained settings: individual participant data meta-analysis of observational studies. *PLoS Med.* **8**, e1000391 (2011).
34. Patra, J., Jha, P., Rehm, J. & Suraweera, W. Tobacco Smoking, Alcohol Drinking, Diabetes, Low Body Mass Index and the Risk of Self-Reported Symptoms of Active Tuberculosis: Individual Participant Data (IPD) Meta-Analyses of 72,684 Individuals in 14 High Tuberculosis Burden Countries. *PLOS ONE* **9**, e96433 (2014).
35. Abidi, S. k. *et al.* An Individual Patient Data Meta-Analysis to Estimate the Diagnostic Accuracy of a Machine Learning-Based Software for Analyzing Chest X-Rays of Persons with

Symptoms of Pulmonary Tuberculosis: Preliminary Findings. in C55. *TUBERCULOSIS DIAGNOSIS* A5167–A5167 (American Thoracic Society, 2019). doi:10.1164/ajrccm-conference.2019.199.1\_MeetingAbstracts.A5167.

36. Boyles, T. *et al.* The diagnosis of tuberculous meningitis in adults and adolescents: protocol for a systematic review and individual patient data meta-analysis to inform a multivariable prediction model. *Wellcome Open Res.* **4**, 19 (2019).
37. Artemether-Quinine Meta-analysis Study Group. A meta-analysis using individual patient data of trials comparing artemether with quinine in the treatment of severe falciparum malaria. *Trans. R. Soc. Trop. Med. Hyg.* **95**, 637–650 (2001).
38. Makanga, M. *et al.* Efficacy and safety of the six-dose regimen of artemether-lumefantrine in pediatrics with uncomplicated Plasmodium falciparum malaria: a pooled analysis of individual patient data. *Am. J. Trop. Med. Hyg.* **74**, 991–998 (2006).
39. Whengang Youdom, S., Samson, A., Basco, L. K. & Thalabard, J.-C. Multiple treatment comparisons in a series of anti-malarial trials with an ordinal primary outcome and repeated treatment evaluations. *Malar. J.* **11**, 147 (2012).
40. WWARN Artemisinin based Combination Therapy (ACT) Africa Baseline Study Group. Clinical determinants of early parasitological response to ACTs in African patients with uncomplicated falciparum malaria: a literature review and meta-analysis of individual patient data. *BMC Med.* **13**, 212 (2015).
41. Julé, A. M., Vaillant, M., Lang, T. A., Guérin, P. J. & Olliaro, P. L. The Schistosomiasis Clinical Trials Landscape: A Systematic Review of Antischistosomal Treatment Efficacy Studies and a Case for Sharing Individual Participant-Level Data (IPD). *PLoS Negl. Trop. Dis.* **10**, e0004784 (2016).
42. Stojkovic, M. *et al.* Treatment Response of Cystic Echinococcosis to Benzimidazoles: A Systematic Review. *PLoS Negl. Trop. Dis.* **3**, e524 (2009).
43. Bhutta, Z. A. *et al.* Therapeutic effects of oral zinc in acute and persistent diarrhea in children in developing countries: pooled analysis of randomized controlled trials. *Am. J. Clin. Nutr.* **72**, 1516–1522 (2000).
44. Beyerlein, A. *et al.* Infant formula supplementation with long-chain polyunsaturated fatty acids has no effect on Bayley developmental scores at 18 months of age--IPD meta-analysis of 4 large clinical trials. *J. Pediatr. Gastroenterol. Nutr.* **50**, 79–84 (2010).
45. Rogne, T. *et al.* Associations of Maternal Vitamin B12 Concentration in Pregnancy With the Risks of Preterm Birth and Low Birth Weight: A Systematic Review and Meta-Analysis of Individual Participant Data. *Am. J. Epidemiol.* **185**, 212–223 (2017).
46. Christian, P. *et al.* Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low- and middle-income countries. *Int J Epidemiol* **42**, 1340–1355 (2013).
47. Tudur Smith, C. *et al.* Individual participant data meta-analyses compared with meta-analyses based on aggregate data. *Cochrane Database Syst. Rev.* **9**, MR000007 (2016).

48. Lambert, P. C., Sutton, A. J., Abrams, K. R. & Jones, D. R. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J. Clin. Epidemiol.* **55**, 86–94 (2002).
49. Stewart, L. A. & Tierney, J. F. To IPD or not to IPD?: Advantages and Disadvantages of Systematic Reviews Using Individual Patient Data. *Eval. Health Prof.* **25**, 76–97 (2002).
50. Dias, S., Sutton, A. J., Welton, N. J. & Ades, A. E. Heterogeneity: Subgroups, Meta-Regression, Bias And Bias-Adjustment. (2012).
51. Berlin, J. A. *et al.* Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat. Med.* **21**, 371–387 (2002).
52. Ferguson, A. R., Nielson, J. L., Cragin, M. H., Bandrowski, A. E. & Martone, M. E. Big data from small data: data-sharing in the ‘long tail’ of neuroscience. *Nat. Neurosci.* **17**, 1442–1447 (2014).
53. Hernán, M. A. & Robins, J. M. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am. J. Epidemiol.* **183**, 758–764 (2016).
54. Tierney, J. F. & Stewart, L. A. Investigating patient exclusion bias in meta-analysis. *Int. J. Epidemiol.* **34**, 79–87 (2005).
55. Koopman, L., van der Heijden, G. J. M. G., Grobbee, D. E. & Rovers, M. M. Comparison of methods of handling missing data in individual patient data meta-analyses: an empirical example on antibiotics in children with acute otitis media. *Am. J. Epidemiol.* **167**, 540–545 (2008).
56. Carrig, M. M., Manrique-Vallier, D., Ranby, K. W., Reiter, J. P. & Hoyle, R. H. A Nonparametric, Multiple Imputation-Based Method for the Retrospective Integration of Data Sets. *Multivar. Behav. Res.* **50**, 383–397 (2015).
57. WHO Multicentre Growth Reference Study Group. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. *Geneva World Health Organ.* 312 pages (2006).
58. Kelley, G. A. & Kelley, K. S. Retrieval of Individual Participant Data for Exercise Meta-Analyses May Not Be Worth the Time and Effort. *BioMed Res. Int.* **2016**, 5059041 (2016).
59. Bria, E., Gralla, R. J., Raftopoulos, H. & Giannarelli, D. Comparing two methods of meta-analysis in clinical research - individual patient data-based (IPD) and literature-based abstracted data (AD) methods: Analyzing five oncology issues involving more than 10,000 patients in randomized clinical trials (RCTs). *J. Clin. Oncol.* **25**, 6512–6512 (2007).
60. Laan, M. J. van der & Rose, S. *Targeted Learning: Causal Inference for Observational and Experimental Data.* (Springer-Verlag, 2011).
61. WHO | Joint child malnutrition estimates - Levels and trends (2019 edition). *WHO* <http://www.who.int/nutgrowthdb/estimates2018/en/>.

62. Malnutrition and Infection: Complex Mechanisms and Global Impacts. <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0040115>.
63. Myatt, M. *et al.* Children who are both wasted and stunted are also underweight and have a high risk of death: a descriptive epidemiology of multiple anthropometric deficits using data from 51 countries. *Arch. Public Health* **76**, (2018).
64. Bhutta, Z. A. *et al.* Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? *The Lancet* **382**, 452–477 (2013).
65. World Health Organization & Nutrition for Health and Development. *Guideline*. (2013).
66. Horton, R. Maternal and child undernutrition: an urgent opportunity. *The Lancet* **371**, 179 (2008).
67. Grantham-McGregor, S. *et al.* Developmental potential in the first 5 years for children in developing countries. *The Lancet* **369**, 60–70 (2007).
68. Martins, V. J. B. *et al.* Long-Lasting Effects of Undernutrition. *Int. J. Environ. Res. Public Health* **8**, 1817–1846 (2011).
69. Richard, S. A. *et al.* Wasting Is Associated with Stunting in Early Childhood. *J. Nutr.* **142**, 1291–1296 (2012).
70. Stobaugh, H. C. *et al.* Children with Poor Linear Growth Are at Risk for Repeated Relapse to Wasting after Recovery from Moderate Acute Malnutrition. *J. Nutr.* **148**, 974–979 (2018).
71. Martorell, R. & Young, M. F. Patterns of Stunting and Wasting: Potential Explanatory Factors. *Adv. Nutr.* **3**, 227–233 (2012).
72. Saaka, M. & Galaa, S. Z. Relationships between Wasting and Stunting and Their Concurrent Occurrence in Ghanaian Preschool Children. *J. Nutr. Metab.* **2016**, (2016).
73. Hoffman, D., Cacciola, T., Barrios, P. & Simon, J. Temporal changes and determinants of childhood nutritional status in Kenya and Zambia. *J. Health Popul. Nutr.* **36**, 27 (2017).
74. Isanaka, S., Hitchings, M. D. T., Berthé, F., Briend, A. & Grais, R. F. Linear growth faltering and the role of weight attainment: Prospective analysis of young children recovering from severe wasting in Niger. *Matern. Child. Nutr.* **0**, e12817.
75. C, A., T, K., C, D., Ja, B. & Group, W. T. I. Research priorities on the relationship between wasting and stunting. *Field Exch.* **53** 28 (2016).
76. Lenters, L., Wazny, K. & Bhutta, Z. A. Management of Severe and Moderate Acute Malnutrition in Children. in *Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities, Third Edition (Volume 2)* (eds. Black, R. E., Laxminarayan, R., Temmerman, M. & Walker, N.) (The International Bank for Reconstruction and Development / The World Bank, 2016).
77. Das, J. K., Salam, R. A., Imdad, A. & Bhutta, Z. A. Infant and Young Child Growth. in *Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities, Third Edition*

(Volume 2) (eds. Black, R. E., Laxminarayan, R., Temmerman, M. & Walker, N.) (The International Bank for Reconstruction and Development / The World Bank, 2016).

78. Young, H. & Marshak, A. A discussion paper on the scope of the problem, its drivers, and strategies for moving forward for policy, practice, and research. 55.

79. Nabwera, H. M., Fulford, A. J., Moore, S. E. & Prentice, A. M. Growth faltering in rural Gambian children after four decades of interventions: a retrospective cohort study. *Lancet Glob. Health* **5**, e208–e216 (2017).

80. Luby, S. P. *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* **6**, e302–e315 (2018).

81. Null, C. *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* **6**, e316–e329 (2018).

82. Osgood-Zimmerman, A. *et al.* Mapping child growth failure in Africa between 2000 and 2015. *Nature* **555**, 41–47 (2018).

83. Bulti, A. *et al.* Improving estimates of the burden of severe acute malnutrition and predictions of caseload for programs treating severe acute malnutrition: experiences from Nigeria. *Arch. Public Health* **75**, (2017).

84. Brown, K. H., Black, R. E. & Becker, S. Seasonal changes in nutritional status and the prevalence of malnutrition in a longitudinal study of young children in rural Bangladesh. *Am. J. Clin. Nutr.* **36**, 303–313 (1982).

85. Egata, G., Berhane, Y. & Worku, A. Seasonal variation in the prevalence of acute undernutrition among children under five years of age in east rural Ethiopia: a longitudinal study. *BMC Public Health* **13**, (2013).

86. Kinyoki, D. K. *et al.* Space-time mapping of wasting among children under the age of five years in Somalia from 2007 to 2010. *Spat. Spatio-Temporal Epidemiol.* **16**, 77–87 (2016).

87. Schoenbuchner, S. M. *et al.* The relationship between wasting and stunting: a retrospective cohort analysis of longitudinal data in Gambian children from 1976 to 2016. *Am. J. Clin. Nutr.* doi:10.1093/ajcn/nqy326.

88. Organization, W. H. & Fund (UNICEF), U. N. C. *Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old.* (World Health Organization, 2019).

89. Viechtbauer, W. Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *J. Educ. Behav. Stat.* **30**, 261–293 (2005).

90. Raudenbush, S. W. Analyzing Effect Sizes: Random-Effects Models. in *The Handbook of Research Synthesis and Meta-Analysis* 295–315 (Russell Sage Foundation, 2009).

91. Wood, S. N., Pya, N. & Säfken, B. Smoothing Parameter and Model Selection for General Smooth Models. *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).

92. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
93. Walsh, R. P. D. & Lawler, D. M. Rainfall Seasonality: Description, Spatial Patterns and Change Through Time. *Weather* **36**, 201–208 (1981).
94. Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. *J. Stat. Softw.* **36**, 1–48 (2010).
95. Papageorgiou, A. T. *et al.* The INTERGROWTH-21st fetal growth standards: toward the global integration of pregnancy and pediatric care. *Am. J. Obstet. Gynecol.* **218**, S630–S640 (2018).
96. Jaganath, D. *et al.* First Detected *Helicobacter pylori* Infection in Infancy Modifies the Association Between Diarrheal Disease and Childhood Growth in Peru. *Helicobacter* **19**, 272–279 (2014).
97. Hennig, B. J. *et al.* Cohort Profile: The Kiang West Longitudinal Population Study (KWLPs)—a platform for integrated research and health care provision in rural Gambia. *Int. J. Epidemiol.* **46**, e13 (2017).
98. de Onis, M. *et al.* Comparison of the World Health Organization (WHO) Child Growth Standards and the National Center for Health Statistics/WHO international growth reference: implications for child health programmes. *Public Health Nutr.* **9**, 942–947 (2006).
99. Garenne, M., Myatt, M., Khara, T., Dolan, C. & Briend, A. Concurrent wasting and stunting among under-five children in Niakhar, Senegal. *Matern. Child. Nutr.* **15**, (2019).
100. Payandeh, A., Saki, A., Safarian, M., Tabesh, H. & Siadat, Z. Prevalence of Malnutrition among Preschool Children in Northeast of Iran, A Result of a Population Based Study. *Glob. J. Health Sci.* **5**, 208–212 (2013).
101. Blencowe, H. *et al.* National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *Lancet Glob. Health* **0**, (2019).
102. Ashorn, P. *et al.* Co-causation of reduced newborn size by maternal undernutrition, infections, and inflammation. *Matern. Child. Nutr.* **14**, e12585 (2018).
103. Victora, C. G., Onis, M. de, Hallal, P. C., Blössner, M. & Shrimpton, R. Worldwide Timing of Growth Faltering: Revisiting Implications for Interventions. *Pediatrics* **125**, e473–e480 (2010).
104. Hemalatha, R., Radhakrishna, K. & Kumar, Bn. Undernutrition in children & critical windows of opportunity in Indian context. *Indian J. Med. Res.* **148**, 612 (2018).
105. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals - ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S0140673616315938>.
106. Efficacy of three feeding regimens for home-based management of children with uncomplicated severe acute malnutrition: a randomised trial in India | BMJ Global Health. <https://gh.bmj.com/content/1/4/e000144>.

107. Victora, C. G. The Association between Wasting and Stunting: An International Perspective. *J. Nutr.* **122**, 1105–1110 (1992).
108. Post, C. L. & Victora, C. G. The low prevalence of weight-for-height deficits in Brazilian children is related to body proportions. *J. Nutr.* **131**, 1290–1296 (2001).
109. Aguayo, V. M., Aneja, S., Badgaiyan, N. & Singh, K. Mid upper-arm circumference is an effective tool to identify infants and young children with severe acute malnutrition in India. *Public Health Nutr.* **18**, 3244–3248 (2015).
110. Kikafunda, J. K., Walker, A. F., Collett, D. & Tumwine, J. K. Risk factors for early childhood malnutrition in Uganda. *Pediatrics* **102**, E45 (1998).
111. Berkley, J. *et al.* Assessment of severe malnutrition among hospitalized children in rural Kenya: comparison of weight for height and mid upper arm circumference. *JAMA* **294**, 591–597 (2005).
112. Manary, M. J. & Sandige, H. L. Management of acute moderate and severe childhood malnutrition. *BMJ* **337**, a2180 (2008).
113. Collins, S. *et al.* Management of severe acute malnutrition in children. *The Lancet* **368**, 1992–2000 (2006).
114. Perumal, N. *et al.* Effect of correcting for gestational age at birth on population prevalence of early childhood undernutrition. *Emerg. Themes Epidemiol.* **15**, 3 (2018).
115. McDonald, C. M. *et al.* The effect of multiple anthropometric deficits on child mortality: meta-analysis of individual data in 10 prospective studies from developing countries. *Am. J. Clin. Nutr.* **97**, 896–901 (2013).
116. Forouzanfar, M. H. *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* **388**, 1659–1724 (2016).
117. Luby, S. P. *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* **6**, e302–e315 (2018).
118. Null, C. *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* **6**, e316–e329 (2018).
119. Panjwani, A. & Heidkamp, R. Complementary Feeding Interventions Have a Small but Significant Impact on Linear and Ponderal Growth of Children in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis. *J. Nutr.* **147**, 2169S–2178S (2017).
120. Das, J. K. *et al.* Preventive lipid-based nutrient supplements given with complementary foods to infants and young children 6 to 23 months of age for health, nutrition, and developmental outcomes. *Cochrane Database Syst. Rev.* **5**, CD012611 (2019).

121. Kramer, M. S. *et al.* Breastfeeding and infant growth: biology or bias? *Pediatrics* **110**, 343–347 (2002).
122. Christian, P. *et al.* Effect of fortified complementary food supplementation on child growth in rural Bangladesh: a cluster-randomized trial. *Int. J. Epidemiol.* **44**, 1862–1876 (2015).
123. Korpe, P. S. & Petri, W. A. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol. Med.* **18**, 328–336 (2012).
124. Budge, S., Parker, A. H., Hutchings, P. T. & Garbutt, C. Environmental enteric dysfunction and child stunting. *Nutr. Rev.* **77**, 240–253 (2019).
125. Arnold, B. F., Null, C., Luby, S. P. & Colford, J. M. Implications of WASH Benefits trials for water and sanitation – Authors’ reply. *Lancet Glob. Health* **6**, e616–e617 (2018).
126. Humphrey, J. H. *et al.* Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: a cluster-randomised trial. *Lancet Glob. Health* **7**, e132–e147 (2019).
127. Hubbard, A. E. & Laan, M. J. V. D. Population intervention models in causal inference. *Biometrika* **95**, 35–47 (2008).
128. Westreich, D. From Patients to Policy: Population Intervention Effects in Epidemiology. *Epidemiol. Camb. Mass* **28**, 525–528 (2017).
129. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr. Oslo Nor. 1992 Suppl.* **450**, 76–85 (2006).
130. Gasparrini, A., Armstrong, B. & Kenward, M. G. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat. Med.* **31**, 3821–3839 (2012).
131. Greenland, S., Robins, J. M. & Pearl, J. Confounding and Collapsibility in Causal Inference. *Stat. Sci.* **14**, 29–46 (1999).
132. VanderWeele, T. J. & Shpitser, I. A new criterion for confounder selection. *Biometrics* **67**, 1406–1413 (2011).
133. Gruber, S. & Laan, M. van der. Targeted Maximum Likelihood Estimation: A Gentle Introduction. *UC Berkeley Div. Biostat. Work. Pap. Ser.* (2009).
134. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).
135. Hastie, T. & Tibshirani, R. Generalized Additive Models. *Stat. Sci.* **1**, 297–310 (1986).
136. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. - KDD 16* 785–794 (2016)  
doi:10.1145/2939672.2939785.
137. McNutt, L.-A., Wu, C., Xue, X. & Hafner, J. P. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am. J. Epidemiol.* **157**, 940–943 (2003).
138. Population attributable risk (PAR). in *Encyclopedia of Public Health* (ed. Kirch, W.) 1117–1118 (Springer Netherlands, 2008). doi:10.1007/978-1-4020-5614-7\_2685.

139. WHO | The WHO Child Growth Standards. *WHO* <http://www.who.int/childgrowth/en/>.
140. Varela-Silva, M. I., Azcorra, H., Dickinson, F., Bogin, B. & Frisancho, A. R. Influence of maternal stature, pregnancy age, and infant birth weight on growth during childhood in Yucatan, Mexico: A test of the intergenerational effects hypothesis. *Am. J. Hum. Biol.* **21**, 657–663 (2009).
141. McGovern, M. E., Krishna, A., Aguayo, V. M. & Subramanian, S. V. A review of the evidence linking child stunting to economic outcomes. *Int. J. Epidemiol.* **46**, 1171–1191 (2017).
142. Haddad, L., Alderman, H., Appleton, S., Song, L. & Yohannes, Y. Reducing Child Malnutrition: How Far Does Income Growth Take Us? *World Bank Econ. Rev.* **17**, 107–131 (2003).
143. Young, M. F. *et al.* Role of maternal preconception nutrition on offspring growth and risk of stunting across the first 1000 days in Vietnam: A prospective cohort study. *PloS One* **13**, e0203201 (2018).
144. Addo, O. Y. *et al.* Maternal Height and Child Growth Patterns. *J. Pediatr.* **163**, 549–554.e1 (2013).
145. Martorell, R. & Zongrone, A. Intergenerational Influences on Child Growth and Undernutrition. *Paediatr. Perinat. Epidemiol.* **26**, 302–314 (2012).
146. Bzikowska-Jura, A. *et al.* Maternal Nutrition and Body Composition During Breastfeeding: Association with Human Milk Composition. *Nutrients* **10**, (2018).
147. Garg, A. & Morduch, J. Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *J. Popul. Econ.* **11**, 471–493 (1998).
148. Black, R. E. *et al.* Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* **371**, 243–260 (2008).
149. Kosek, M. *et al.* Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *Am. J. Trop. Med. Hyg.* **88**, 390–396 (2013).
150. Investigators, M.-E. N. Relationship between growth and illness, enteropathogens and dietary intakes in the first 2 years of life: findings from the MAL-ED birth cohort study. *BMJ Glob. Health* **2**, (2017).
151. VanderWeele, T. J. & Ding, P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann. Intern. Med.* **167**, 268 (2017).
152. Bhutta, Z. A. *et al.* What works? Interventions for maternal and child undernutrition and survival. *The Lancet* **371**, 417–440 (2008).
153. Piwoz, E., Sundberg, S. & Rooke, J. Promoting Healthy Growth: What Are the Priorities for Research and Action? *Adv. Nutr.* **3**, 234–241 (2012).
154. Kristjansson, E. *et al.* Food supplementation for improving the physical and psychosocial health of socio-economically disadvantaged children aged three months to five years. *Cochrane Database Syst. Rev.* CD009924 (2015) doi:10.1002/14651858.CD009924.pub2.

155. WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience. (World Health Organization, 2016).
156. Ota, E., Hori, H., Mori, R., Tobe-Gai, R. & Farrar, D. Antenatal dietary education and supplementation to increase energy and protein intake. *Cochrane Database Syst. Rev.* (2015) doi:10.1002/14651858.CD000032.pub3.
157. Hambidge, K. M. *et al.* A multicountry randomized controlled trial of comprehensive maternal nutrition supplementation initiated before conception: the Women First trial. *Am. J. Clin. Nutr.* **109**, 457–469 (2019).
158. Unger, H. W. *et al.* Azithromycin-containing intermittent preventive treatment in pregnancy affects gestational weight gain, an important predictor of birthweight in Papua New Guinea – an exploratory analysis. *Matern. Child. Nutr.* **12**, 699–712 (2016).
159. Hallamaa, L. *et al.* The impact of maternal antenatal treatment with two doses of azithromycin and monthly sulphadoxine-pyrimethamine on child weight, mid-upper arm circumference and head circumference: A randomized controlled trial. *PLOS ONE* **14**, e0216536 (2019).
160. Hallamaa, L. *et al.* Child Health Outcomes After Presumptive Infection Treatment in Pregnant Women: A Randomized Trial. *Pediatrics* **141**, (2018).
161. Beeson, J. G. & Simpson, J. A. The potential benefit of scaling up malaria prevention to reduce low birth weight in Africa. *PLOS Med.* **14**, e1002244 (2017).
162. Sudfeld, C. R. *et al.* Malnutrition and Its Determinants Are Associated with Suboptimal Cognitive, Communication, and Motor Development in Tanzanian Children. *J. Nutr.* **145**, 2705–2714 (2015).
163. Crookston, B. T. *et al.* Impact of early and concurrent stunting on cognition. *Matern. Child. Nutr.* **7**, 397–409 (2011).
164. Reinhardt, K. & Fanzo, J. Addressing Chronic Malnutrition through Multi-Sectoral, Sustainable Approaches: A Review of the Causes and Consequences. *Front. Nutr.* **1**, 13 (2014).
165. Olofin, I. *et al.* Associations of Suboptimal Growth with All-Cause and Cause-Specific Mortality in Children under Five Years: A Pooled Analysis of Ten Prospective Studies. *PLoS ONE* **8**, (2013).
166. Lopriore, C., Dop, M.-C., Solal-Céligny, A. & Lagnado, G. Excluding infants under 6 months of age from surveys: impact on prevalence of pre-school undernutrition. *Public Health Nutr.* **10**, 79–87 (2007).
167. Vygen, S. B., Roberfroid, D., Captier, V. & Kolsteren, P. Treatment of severe acute malnutrition in infants aged <6 months in Niger. *J. Pediatr.* **162**, 515-521.e3 (2013).
168. Mwangome, M. K., Fegan, G., Fulford, T., Prentice, A. M. & Berkley, J. A. Mid-upper arm circumference at age of routine infant vaccination to identify infants at elevated risk of death: a retrospective cohort study in the Gambia. *Bull. World Health Organ.* **90**, 887–894 (2012).

169. Myatt, M., Khara, T. & Collins, S. A review of methods to detect cases of severely malnourished children in the community for their admission into community-based therapeutic care programs. *Food Nutr. Bull.* **27**, S7-23 (2006).
170. Vesel, L. *et al.* Use of new World Health Organization child growth standards to assess how infant malnutrition relates to breastfeeding and mortality. *Bull. World Health Organ.* **88**, 39–48 (2010).
171. CME Info - Child Mortality Estimates. <https://childmortality.org/>.
172. Prost, A. *et al.* Mortality and recovery following moderate and severe acute malnutrition in children aged 6–18 months in rural Jharkhand and Odisha, eastern India: A cohort study. *PLOS Med.* **16**, e1002934 (2019).
173. Singh, K. *et al.* Management of children with severe acute malnutrition: Experience of Nutrition Rehabilitation Centers in Uttar Pradesh, India. *Indian Pediatr.* **51**, 21–25 (2014).
174. Pelletier, D. L. The relationship between child anthropometry and mortality in developing countries: implications for policy, programs and future research. *J. Nutr.* **124**, 2047S-2081S (1994).
175. Sachdev, H. S., Sinha, S., Sareen, N., Pandey, R. M. & Kapil, U. Survival and recovery in severely wasted under-five children without community management of acute malnutrition programme. *Indian Pediatr.* **54**, 817–824 (2017).
176. Hoddinott, J., Alderman, H., Behrman, J. R., Haddad, L. & Horton, S. The economic rationale for investing in stunting reduction. *Matern. Child. Nutr.* **9**, 69–82 (2013).
177. De Lucia Rolfe, E. *et al.* Associations of stunting in early childhood with cardiometabolic risk factors in adulthood. *PLoS ONE* **13**, (2018).
178. Victora, C. G. *et al.* Maternal and child undernutrition: consequences for adult health and human capital. *Lancet* **371**, 340–357 (2008).
179. Dewey, K. G. & Begum, K. Long-term consequences of stunting in early life. *Matern. Child. Nutr.* **7**, 5–18 (2011).
180. Mwangome, M. *et al.* Diagnostic criteria for severe acute malnutrition among infants aged under 6 mo. *Am. J. Clin. Nutr.* **105**, 1415–1423 (2017).
181. de Onis, M., Wijnhoven, T. M. A. & Onyango, A. W. Worldwide practices in child growth monitoring. *J. Pediatr.* **144**, 461–465 (2004).
182. Hillbruner, C. & Egan, R. Seasonality, household food security, and nutritional status in Dinajpur, Bangladesh. *Food Nutr. Bull.* **29**, 221–231 (2008).
183. Gilbert, C. L., Christiaensen, L. & Kaminski, J. Food price seasonality in Africa: Measurement and extent. *Food Policy* **67**, 119–132 (2017).
184. Chao, D. L., Roose, A., Roh, M., Kotloff, K. L. & Proctor, J. L. The seasonality of diarrheal pathogens: A retrospective study of seven sites over three years. *PLoS Negl. Trop. Dis.* **13**, e0007211 (2019).

185. Fisman, D. Seasonality of viral infections: mechanisms and unknowns. *Clin. Microbiol. Infect.* **18**, 946–954 (2012).
186. Roy, M. Maternal infection, malnutrition, and low birth weight. *J. Postgrad. Med.* **62**, 270–271 (2016).
187. Keino, S., Plasqui, G., Ettyang, G. & van den Borne, B. Determinants of stunting and overweight among young children and adolescents in sub-Saharan Africa. *Food Nutr. Bull.* **35**, 167–178 (2014).
188. Danaei, G. *et al.* Risk Factors for Childhood Stunting in 137 Developing Countries: A Comparative Risk Assessment Analysis at Global, Regional, and Country Levels. *PLoS Med.* **13**, (2016).
189. Cumming, O. & Cairncross, S. Can water, sanitation and hygiene help eliminate stunting? Current evidence and policy implications. *Matern. Child. Nutr.* **12 Suppl 1**, 91–105 (2016).
190. Akombi, B. J. *et al.* Stunting, Wasting and Underweight in Sub-Saharan Africa: A Systematic Review. *Int. J. Environ. Res. Public. Health* **14**, (2017).
191. Sinha, R. K., Dua, R., Bijalwan, V., Rohatgi, S. & Kumar, P. Determinants of Stunting, Wasting, and Underweight in Five High-Burden Pockets of Four Indian States. *Indian J. Community Med. Off. Publ. Indian Assoc. Prev. Soc. Med.* **43**, 279–283 (2018).
192. Partap, U., Young, E. H., Allotey, P., Sandhu, M. S. & Reidpath, D. D. Characterisation and correlates of stunting among Malaysian children and adolescents aged 6-19 years. *Glob. Health Epidemiol. Genomics* **4**, (2019).
193. Prado, E. L. *et al.* Path analyses of risk factors for linear growth faltering in four prospective cohorts of young children in Ghana, Malawi and Burkina Faso. *BMJ Glob. Health* **4**, e001155 (2019).
194. Harding, K. L., Aguayo, V. M. & Webb, P. Factors associated with wasting among children under five years old in South Asia: Implications for action. *PloS One* **13**, e0198749 (2018).
195. Kinyoki, D. K., Berkley, J. A., Moloney, G. M., Kandala, N.-B. & Noor, A. M. Predictors of the risk of malnutrition among children under the age of 5 years in Somalia. *Public Health Nutr.* **18**, 3125–3133 (2015).
196. Goyal, N. & Canning, D. Exposure to Ambient Fine Particulate Air Pollution in Utero as a Risk Factor for Child Stunting in Bangladesh. *Int. J. Environ. Res. Public. Health* **15**, (2017).
197. Fadare, O., Amare, M., Mavrotas, G., Akerele, D. & Ogunniyi, A. Mother's nutrition-related knowledge and child nutrition outcomes: Empirical evidence from Nigeria. *PloS One* **14**, e0212775 (2019).
198. Northridge, M. E. Public health methods--attributable risk as a link between causality and public health action. *Am. J. Public Health* **85**, 1202–1204 (1995).

199. Green, M. A., Corsi, D. J., Mejía-Guevara, I. & Subramanian, S. V. Distinct clusters of stunted children in India: An observational study. *Matern. Child. Nutr.* **14**, e12592 (2018).
200. Flax, V. L. 'It was caused by the carelessness of the parents': cultural models of child malnutrition in southern Malawi. *Matern. Child. Nutr.* **11**, 104–118 (2015).
201. Larrea, C. & Freire, W. Social inequality and child malnutrition in four Andean countries. *Rev. Panam. Salud Pública* **11**, 356–364 (2002).
202. Richards, E. *et al.* Going beyond the surface: Gendered intra-household bargaining as a social determinant of child health and nutrition in low and middle income countries. *Soc. Sci. Med.* **95**, 24–33 (2013).
203. Pande, R. P. Selective gender differences in childhood nutrition and immunization in rural India: The role of siblings. *Demography* **40**, 395–418 (2003).
204. Morales, R., Aguilar, A. M. & Calzadilla, A. Geography and culture matter for malnutrition in Bolivia. *Econ. Hum. Biol.* **2**, 373–389 (2004).
205. Sanin, K. I. *et al.* Micronutrient adequacy is poor, but not associated with stunting between 12–24 months of age: A cohort study findings from a slum area of Bangladesh. *PloS One* **13**, e0195072 (2018).
206. Wamani, H., Åström, A. N., Peterson, S., Tumwine, J. K. & Tylleskär, T. Boys are more stunted than girls in Sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys. *BMC Pediatr.* **7**, 17 (2007).
207. Mushtaq, M. U. *et al.* Prevalence and socio-demographic correlates of stunting and thinness among Pakistani primary school children. *BMC Public Health* **11**, 790 (2011).
208. Abdulahi, A., Shab-Bidar, S., Rezaei, S. & Djafarian, K. Nutritional Status of Under Five Children in Ethiopia: A Systematic Review and Meta-Analysis. *Ethiop. J. Health Sci.* **27**, 175–188 (2017).
209. Schnee, A. E. *et al.* Identification of Etiology-Specific Diarrhea Associated With Linear Growth Faltering in Bangladeshi Infants. *Am. J. Epidemiol.* **187**, 2210–2218 (2018).
210. Checkley, W. *et al.* Multi-country analysis of the effects of diarrhoea on childhood stunting. *Int. J. Epidemiol.* **37**, 816–830 (2008).
211. Krishna, A. *et al.* Short-term and long-term associations between household wealth and physical growth: a cross-comparative analysis of children from four low- and middle-income countries. *Glob. Health Action* **8**, 26523 (2015).
212. Kang, Y. *et al.* Nutritional status and risk factors for stunting in preschool children in Bhutan. *Matern. Child. Nutr.* **14 Suppl 4**, e12653 (2018).
213. Jones, A. D. *et al.* World Health Organization infant and young child feeding indicators and their associations with child anthropometry: a synthesis of recent findings. *Matern. Child. Nutr.* **10**, 1–17 (2014).
214. Sreeramareddy, C. T., Ramakrishnareddy, N. & Subramaniam, M. Association between household food access insecurity and nutritional status indicators among children aged <5 years

- in Nepal: results from a national, cross-sectional household survey. *Public Health Nutr.* **18**, 2906–2914 (2015).
215. Widen, E. M. *et al.* Maternal Weight Loss during Exclusive Breastfeeding Is Associated with Reduced Weight and Length Gain in Daughters of HIV-Infected Malawian Women. *J. Nutr.* **143**, 1168–1175 (2013).
  216. Muenchhoff, M. & Goulder, P. J. R. Sex Differences in Pediatric Infectious Diseases. *J. Infect. Dis.* **209**, S120–S126 (2014).
  217. Gera, T., Shah, D. & Sachdev, H. S. Impact of Water, Sanitation and Hygiene Interventions on Growth, Non-diarrheal Morbidity and Mortality in Children Residing in Low- and Middle-income Countries: A Systematic Review. *Indian Pediatr.* **55**, 381–393 (2018).
  218. Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial Team *et al.* The Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial: Rationale, Design, and Methods. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **61 Suppl 7**, S685–702 (2015).
  219. Raj, A., McDougal, L. P. & Silverman, J. G. Gendered Effects of Siblings on Child Malnutrition in South Asia: Cross-sectional analysis of Demographic and Health Surveys from Bangladesh, India, and Nepal. *Matern. Child Health J.* **19**, 217–226 (2015).
  220. Mwangome, M. *et al.* Anthropometry at birth and at age of routine vaccination to predict mortality in the first year of life: A birth cohort study in BukinaFaso. *PLOS ONE* **14**, e0213523 (2019).
  221. Prost, A. *et al.* Mortality and recovery following moderate and severe acute malnutrition in children aged 6-18 months in rural Jharkhand and Odisha, eastern India: A cohort study. *PLoS Med.* **16**, e1002934 (2019).
  222. Ballard, T., Coates, J., Swindale, A. & Deitchler, M. Household Hunger Scale: Indicator Definition and Measurement Guide. 23.
  223. Coates, J. Measuring Food Insecurity: Going Beyond Indicators of Income and Anthropometry. 106 (1825).
  224. Coates, J., Swindale, A. & Bilinsky, P. Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide: Version 3: (576842013-001). (2007) doi:10.1037/e576842013-001.
  225. Griffith, L. E. *et al.* Statistical approaches to harmonize data on cognitive measures in systematic reviews are rarely reported. *J. Clin. Epidemiol.* **68**, 154–162 (2015).
  226. Stewart, L. A. *et al.* Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. *JAMA* **313**, 1657–1665 (2015).
  227. Griffith, L. *et al.* *Harmonization of Cognitive Measures in Individual Participant Data and Aggregate Data Meta-Analysis.* (Agency for Healthcare Research and Quality (US), 2013).

228. Zeng, X. *et al.* The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *J. Evid.-Based Med.* **8**, 2–10 (2015).
229. Geeta, A. *et al.* Reliability, technical error of measurements and validity of instruments for nutritional status assessment of adults in Malaysia. *Singapore Med. J.* **50**, 1013–1018 (2009).
230. WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP). *Progress on drinking water, sanitation and hygiene : 2017 update and SDG baselines*. (World Health Organization (WHO) and UNICEF, 2017).
231. Howe, L. D., Hargreaves, J. R. & Huttly, S. R. Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerg. Themes Epidemiol.* **5**, 3 (2008).
232. Vyas, S. & Kumaranayake, L. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy Plan.* **21**, 459–468 (2006).
233. Nevitt, S. J. *et al.* Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. *BMJ* **357**, j1390 (2017).
234. Bertagnolli, M. M. *et al.* Advantages of a Truly Open-Access Data-Sharing Model. *N. Engl. J. Med.* **376**, 1178–1181 (2017).
235. Banzi, R. *et al.* Evaluation of repositories for sharing individual-participant data from clinical studies. *Trials* **20**, 169 (2019).
236. Veroniki, A. A. *et al.* Retrieval of individual patient data depended on study characteristics: a randomized controlled trial. *J. Clin. Epidemiol.* **113**, 176–188 (2019).
237. Committee on Strategies for Responsible Sharing of Clinical Trial Data, Board on Health Sciences Policy & Institute of Medicine. *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. (National Academies Press (US), 2015).
238. Loder, E. & Groves, T. The BMJ requires data sharing on request for all trials. *BMJ* **350**, h2373 (2015).
239. Lavery, J. V. & IJsselmuiden, C. The Research Fairness Initiative: Filling a critical gap in global research ethics. *Gates Open Res.* **2**, (2018).
240. Serwadda, D., Ndebele, P., Grabowski, M. K., Bajunirwe, F. & Wanyenze, R. K. Open data sharing and the Global South—Who benefits? *Science* **359**, 642–643 (2018).
241. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nat. News* **533**, 452 (2016).
242. Mathew, T. & Nordström, K. Comparison of one-step and two-step meta-analysis models using individual patient data. *Biom. J. Biom. Z.* **52**, 271–287 (2010).
243. Riley, R. D. *et al.* Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat. Med.* **27**, 1870–1893 (2008).

244. Olkin, I. & Sampson, A. Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics* **54**, 317–322 (1998).
245. Mathew, T. & Nordström, K. On the equivalence of meta-analysis using literature and using individual patient data. *Biometrics* **55**, 1221–1223 (1999).
246. Bowden, J., Tierney, J. F., Simmonds, M., Copas, A. J. & Higgins, J. P. Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random effects model. *Res. Synth. Methods* **2**, 150–162 (2011).
247. Riley, R. D., Simmonds, M. C. & Look, M. P. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *J. Clin. Epidemiol.* **60**, 431–439 (2007).
248. Houwelingen, H. C. van, Arends, L. R. & Stijnen, T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat. Med.* **21**, 589–624 (2002).
249. Riley, R. D. Commentary: Like it and lump it? Meta-analysis using individual participant data. *Int. J. Epidemiol.* **39**, 1359–1361 (2010).
250. Debray, T. P. A. *et al.* Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res. Synth. Methods* **6**, 293–309 (2015).
251. Burke, D. L., Ensor, J. & Riley, R. D. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat. Med.* **36**, 855–875 (2017).
252. Munafò, M. R. *et al.* A manifesto for reproducible science. *Nat. Hum. Behav.* **1**, 1–9 (2017).
253. Jones, M. P. Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression. *J. Am. Stat. Assoc.* **91**, 222 (1996).
254. Witten, I. H. & Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition.* (Morgan Kaufmann, 2005).
255. Hubbard, A. *et al.* Time-Dependent Prediction and Evaluation of Variable Importance Using SuperLearning in High Dimensional Clinical Data. *J. Trauma Acute Care Surg.* **75**, S53–S60 (2013).
256. Burgess, S., White, I. R., Resche-Rigon, M. & Wood, A. M. Combining multiple imputation and meta-analysis with individual participant data. *Stat. Med.* **32**, 4499–4514 (2013).
257. Diagnosing and responding to violations in the positivity assumption - Maya L Petersen, Kristin E Porter, Susan Gruber, Yue Wang, Mark J van der Laan, 2012.

## Appendix A: Supplementary material for Chapter 2

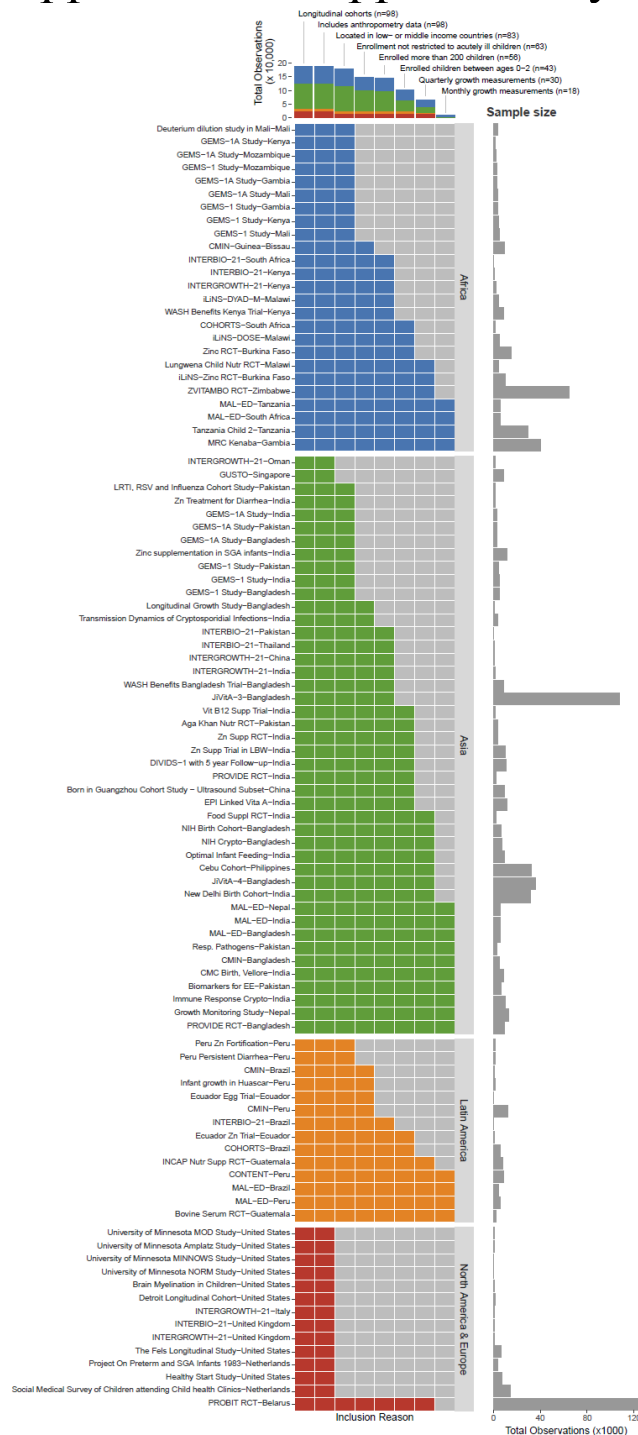


**Figure A.1| KI cohort selection.**

Analyses focused on longitudinal cohorts to enable the estimation of prospective incidence rates and growth velocity. On July 15, 2018, there were 86 longitudinal studies on GHAP. From this set, I applied five inclusion criteria to select cohorts for analysis. My rationale for each criterion follows. (1) *Studies were conducted in lower*

*income or middle-income countries.* My target of inference for analyses was children in LMICs, which remains a key target population for preventive interventions. (2) *Studies measured length and weight between birth and age 24 months.* I was principally interested in growth failure during the first two years of life including at birth, thought to be the key window for linear growth failure <sup>7</sup>. (3) *Studies did not restrict enrollment to acutely ill children.* My focus on descriptive analyses led me to target, to the extent possible, the general population. I thus excluded some studies that exclusively enrolled acutely ill children, such as children who presented to hospital with acute diarrhea or who were severely malnourished. (4) *Studies enrolled at least 200 children.* Age-stratified incident episodes of stunting and wasting were sufficiently rare that I wanted to ensure each cohort would have enough information to estimate rates before contributing to pooled estimates. (5) *Studies collected anthropometry measurements at least every 3 months.* I limited studies to those with higher temporal resolution to ensure that I adequately captured incident episodes and recovery. I further restricted analyses of wasting incidence and recovery to cohorts with monthly measurements because of high temporal variation in WHZ within individuals.

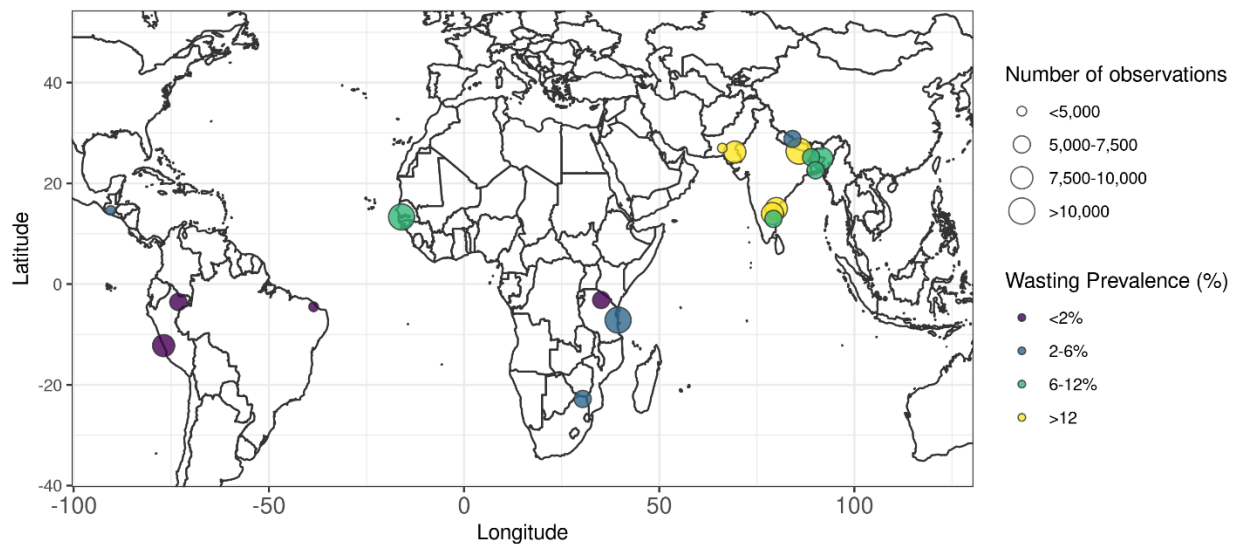
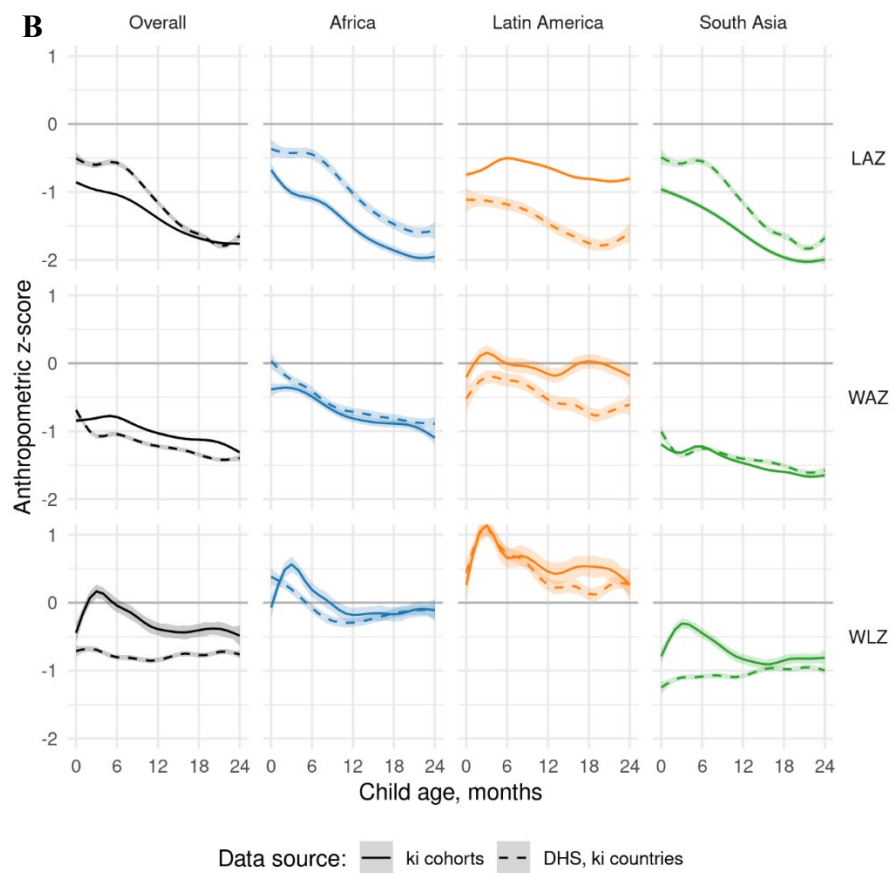
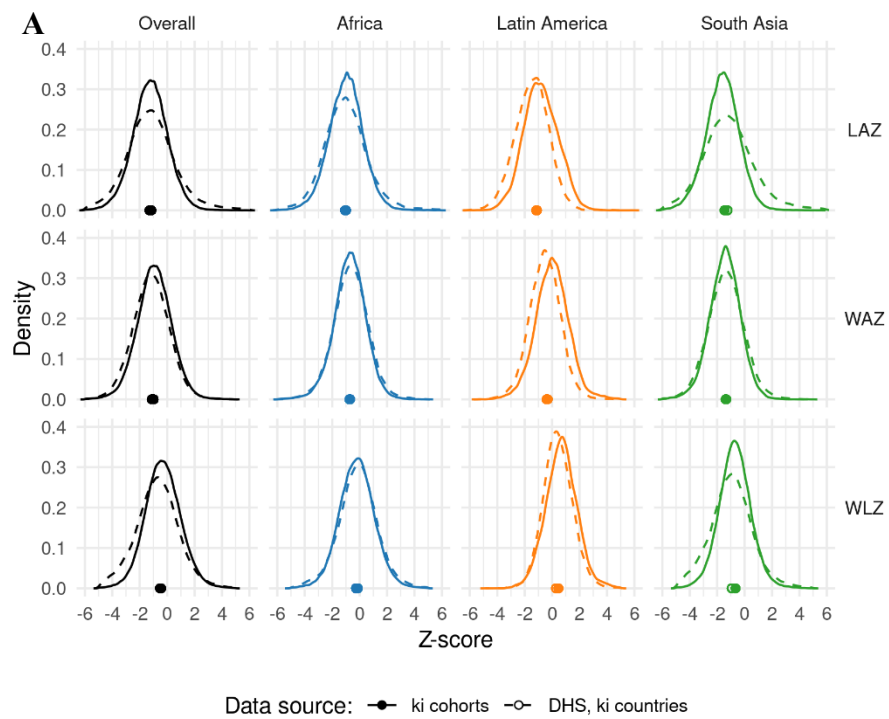


**Figure A.2 | Geographic location of monthly-measured KI cohorts.**

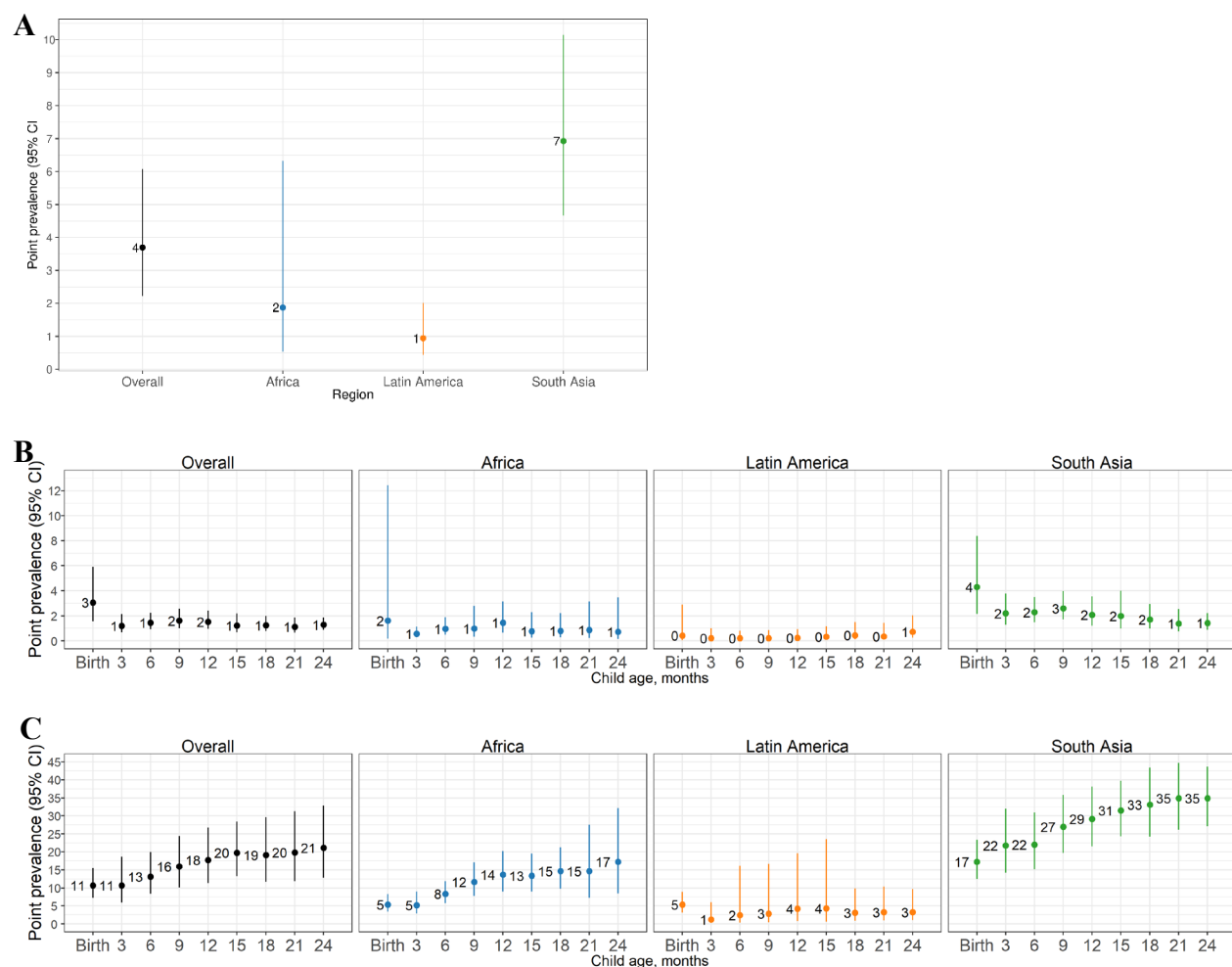
Locations are approximate and jittered slightly for display.



### Figure A.3 | Comparison of cohort anthropometry to population-based samples

(A) Kernel density distributions of length-for-age (LAZ), weight-for-age (WAZ), and weight-for-length Z-scores (WLZ) from measurements among children under 24 months old in 18 KI longitudinal cohorts (solid line) and among children measured in population-based, Demographic and Health Surveys (dashed). Estimates stratified by WHO region: AFRO (Africa), SEARO (South East Asia), PAHO (Latin America). DHS estimates include the most recent survey from all countries measured in each WHO region. Overall estimates pool across the three regions.

(B) Mean LAZ, WAZ, and WLZ by age among 18 KI longitudinal cohorts (solid line) and in Demographic and Health Surveys (dashed). Means estimated with cubic splines and shaded regions show approximate, simultaneous 95% credible intervals.

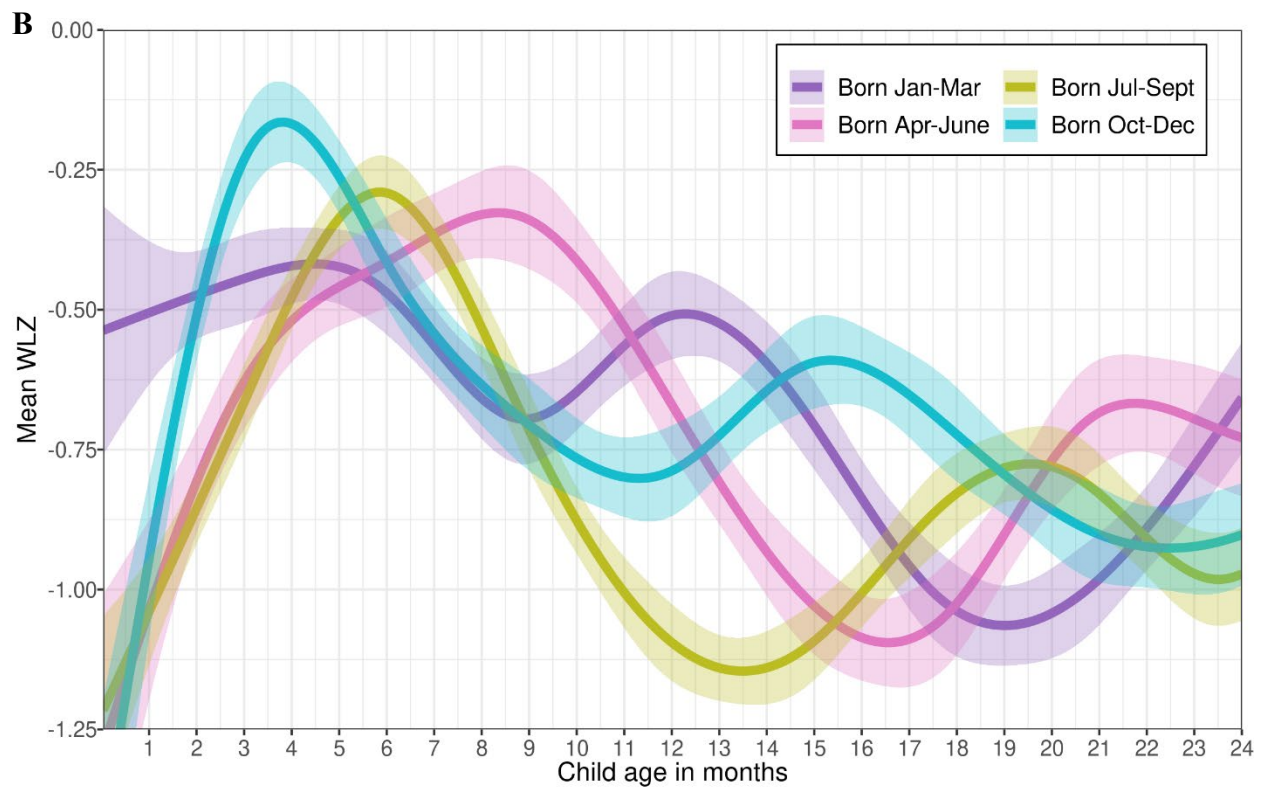
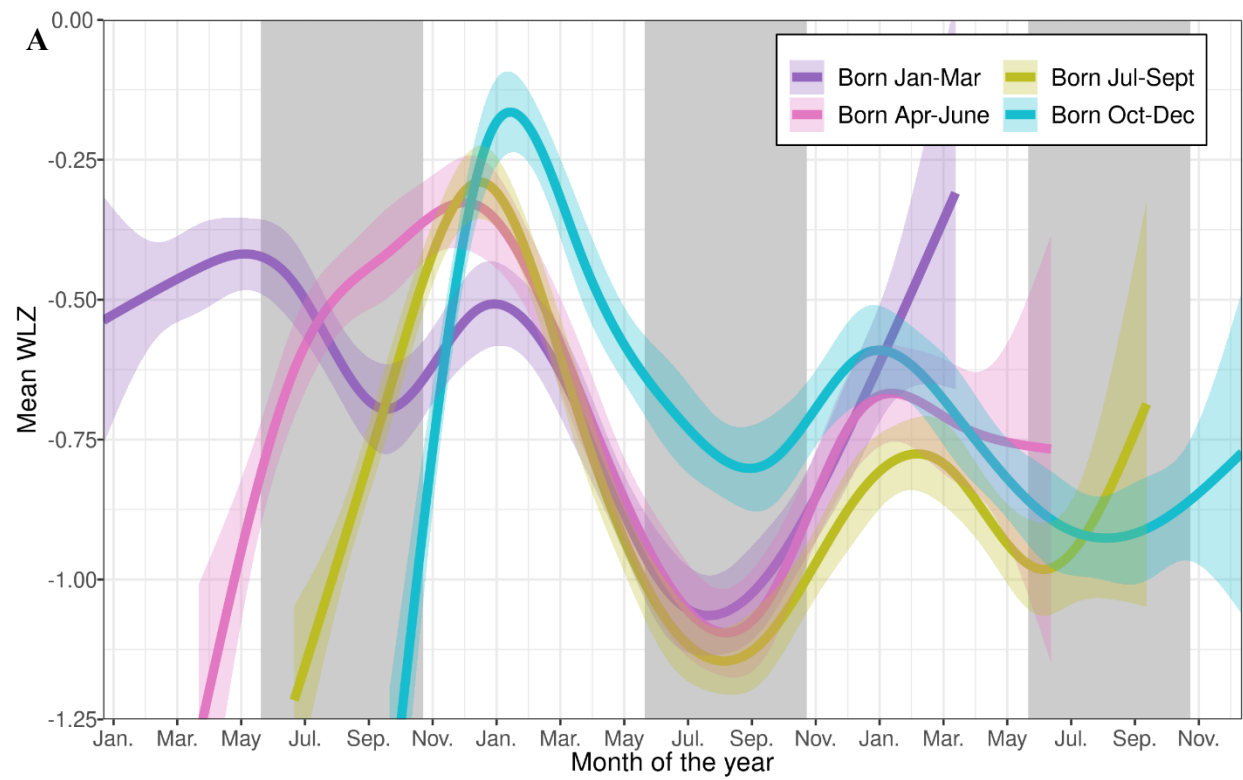


### Figure A.4 | Prevalence of persistent wasting, severe wasting, and underweight by region.

(A) Proportion of children persistently wasted ( $\geq 50\%$  of measurements from birth to 24 months of age), overall and stratified by region.

(B) Prevalence of severe wasting ( $WLZ < -3$ ) by age and region.

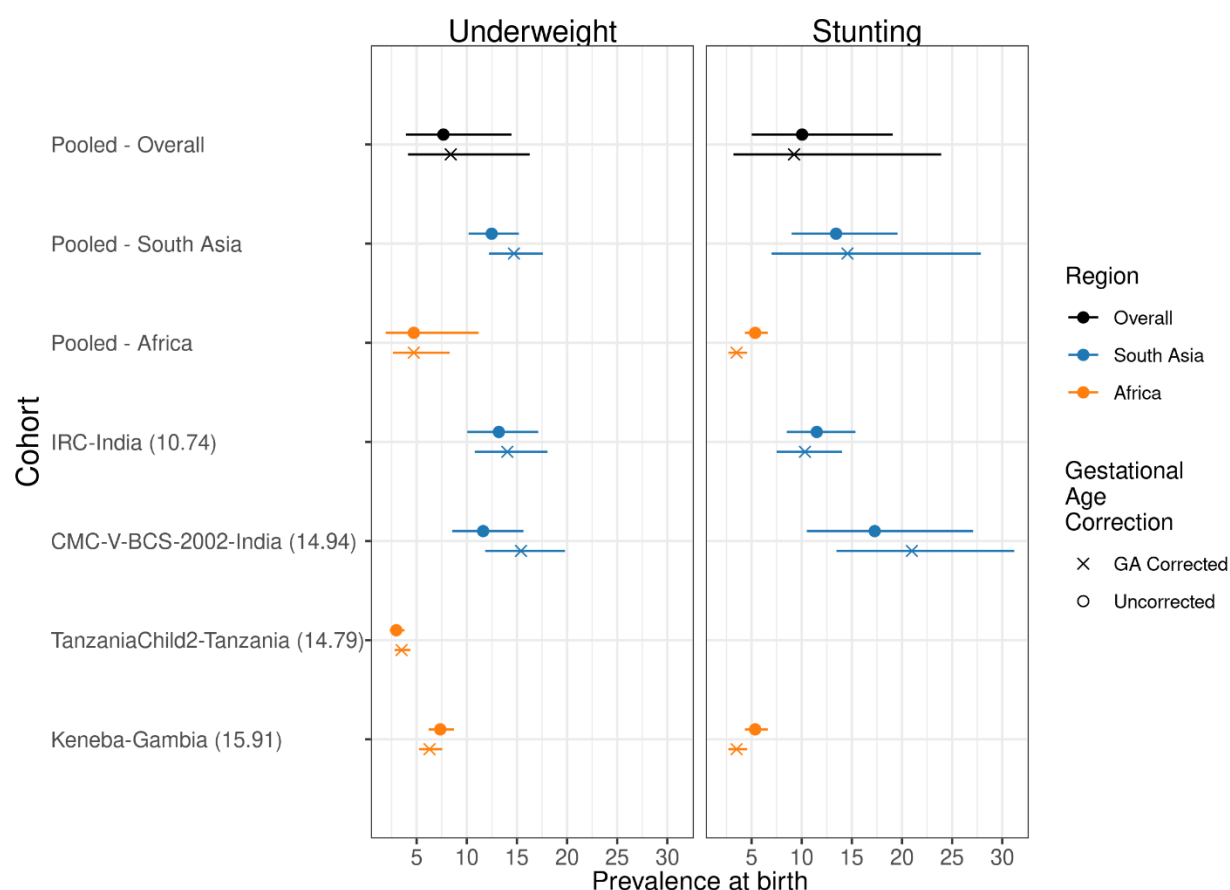
(C) Prevalence of underweight (weight-for-age Z-score  $< -2$ ) by age and region.



## Figure A.5 | Month of birth affects seasonal patterns in WLZ

(a) Mean WLZ by calendar month among South Asian cohorts, with children stratified by birth month. The X-axis begins at January 1<sup>st</sup> of the year of the first January birthday in each cohort. Grey backgrounds indicate the approximate timing of seasonal monsoons in South Asia (June-September). Shaded regions around spline fits indicate 95% simultaneous confidence intervals. Ten cohorts, 3,880 children, and 75,371 measurements were used to estimate the splines. South Asian children born in July-September had the lowest mean WLZ overall and children born April-September had larger seasonal declines in WLZ during their second year of life than children born October-March.

(b) Mean WLZ from birth to age 24 months among children from South Asian cohorts stratified by birth month. Shaded regions around spline fits indicate 95% simultaneous confidence intervals. Data used is the same as in panel (A).



## Figure A.6 | Comparison of underweight and stunting prevalence at birth with and without gestational age correction

This figure includes the results from correcting at-birth Z-scores in the KI cohorts that measured gestational age (GA). The corrections are using the Intergrowth standards and are implemented using the R *growthstandards* package (<https://ki-tools.github.io/growthstandards/>). Overall, the prevalence at birth decreased slightly after correcting for gestational age, but the cohort-specific results are inconsistent. Observations with GA outside of the Intergrowth standards range (<168 or > 300 days) were dropped for both the corrected and uncorrected data. Prevalence increased after GA correction in some cohorts due to high rates of late-term births based on reported GA. There were no length measurements at birth in the EE or Tanzania Child 2 cohorts, so they do not have stunting estimates. There were 4,822 measurements used in the underweight analysis and 1,931 measurements used in the stunting analysis. Gestational age was estimates based on mother's recall of the last menstrual period in the IRC, CMC-V-BCS-2002, and Tanzania Child cohorts, and was based on the Dubowitz method (newborn exam) in the MRC Keneba cohort.

Table A.1| Summary of KI cohorts used in Chapter 2

| Region, Study ID          | Country  | Study Years    | Design             | Children Enrolled * | Anthropometry measurement ages (months) | Total measurements* | Primary References                          |
|---------------------------|----------|----------------|--------------------|---------------------|-----------------------------------------|---------------------|---------------------------------------------|
| South Asia                |          |                |                    |                     |                                         |                     |                                             |
| Biomarkers for EE         | Pakistan | 2013-2015      | Prospective cohort | 380                 | Birth, 1, 2, ..., 18                    | 8428                | Iqbal et al 2018 Nature Scientific Reports  |
| Resp. Pathogens           | Pakistan | 2011 - 2014    | Prospective cohort | 284                 | Birth, 1, 2, ..., 17                    | 3164                | Ali et al 2016- Journal of Medical Virology |
| Growth Monitoring Study   | Nepal    | 2012 - Ongoing | Prospective cohort | 686                 | Birth, 1, 2, ..., 24                    | 13340               | Not yet published                           |
| MAL-ED                    | Nepal    | 2010 - 2014    | Prospective cohort | 240                 | Birth, 1, 2, ..., 24                    | 5695                | Shrestha et al 2014 Clin Infect Dis NPL     |
| CMC Birth Cohort, Vellore | India    | 2002 - 2006    | Prospective cohort | 373                 | Birth, 0.5, 1, 1.5, ..., 24             | 8697                | Gladstone et al. 2011 NEJM                  |
| MAL-ED                    | India    | 2010 - 2012    | Prospective cohort | 251                 | Birth, 1, 2, ..., 24                    | 5697                | John et al 2014 Clin Infect Dis IND         |
| Vellore Crypto Study      | India    | 2008 - 2011    | Prospective cohort | 410                 | Birth, 1, 2, ..., 24                    | 9729                | Kattula et al. 2014 BMJ Open                |

|                                                                                                                                                                             |              |             |                    |      |                                                          |       |                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-------------|--------------------|------|----------------------------------------------------------|-------|---------------------------------------------|
| CMIN                                                                                                                                                                        | Bangladesh   | 1993 - 1996 | Prospective Cohort | 272  | Birth, 3, 6, ..., 24                                     | 5372  | Pathela et al 2006<br>Acta Paediatrica      |
| MAL-ED                                                                                                                                                                      | Bangladesh   | 2010 - 2014 | Prospective cohort | 263  | Birth, 1, 2, ..., 24                                     | 5592  | Ahmed et al 2014<br>Clin Infect Dis         |
| PROVIDE RCT                                                                                                                                                                 | Bangladesh   | 2011 - 2014 | Individual RCT     | 700  | Birth, 6, 10, 12, 14, 17, 18, 24, 39, 40, 52, 53 (weeks) | 9202  | Kirkpatrick et al 2015<br>Am J Trop Med Hyg |
| Africa                                                                                                                                                                      |              |             |                    |      |                                                          |       |                                             |
| MAL-ED                                                                                                                                                                      | Tanzania     | 2009 - 2014 | Prospective cohort | 261  | Birth, 1, 2, ..., 24                                     | 5698  | Mduma et al 2014<br>Clin Infect Dis TZN     |
| Tanzania Child 2                                                                                                                                                            | Tanzania     | 2007 - 2011 | Individual RCT     | 2396 | 1, 2, ..., 20                                            | 29518 | Locks et al Am J Clin Nutr 2016             |
| MAL-ED                                                                                                                                                                      | South Africa | 2009 - 2014 | Prospective cohort | 312  | Birth, 1, 2, ..., 24                                     | 6151  | Bessong et al 2014<br>Clin Infect Dis SA    |
| MRC Keneba                                                                                                                                                                  | Gambia       | 1987 - 1997 | Cohort             | 2920 | Birth, 1, 2, ..., 24                                     | 40117 | Schoenbuchner et al. 2019, AJCN             |
| Latin America                                                                                                                                                               |              |             |                    |      |                                                          |       |                                             |
| MAL-ED                                                                                                                                                                      | Peru         | 2009 - 2014 | Prospective cohort | 302  | Birth, 1, 2, ..., 24                                     | 6127  | Yori et al 2014 Clin Infect Dis PER         |
| CONTENT                                                                                                                                                                     | Peru         | 2007 - 2011 | Prospective cohort | 215  | Birth, 1, 2, ..., 24                                     | 8339  | Jaganath et al 2014<br>Helicobacter         |
| Bovine Serum RCT                                                                                                                                                            | Guatemala    | 1997 - 1998 | Individual RCT     | 315  | Baseline, 1, 2, ..., 8                                   | 2545  | Begin et al. 2008, EJC                      |
| MAL-ED                                                                                                                                                                      | Brazil       | 2010 - 2014 | Prospective cohort | 233  | Birth, 1, 2, ..., 24                                     | 4838  | Lima et al 2014 Clin Infect Dis             |
| *Children enrolled is for children with measurements under 2 years of age. Total measurements are number of measurements of anthropometry on children under 2 years of age. |              |             |                    |      |                                                          |       |                                             |

## Appendix B: Supplementary material for Chapter 3

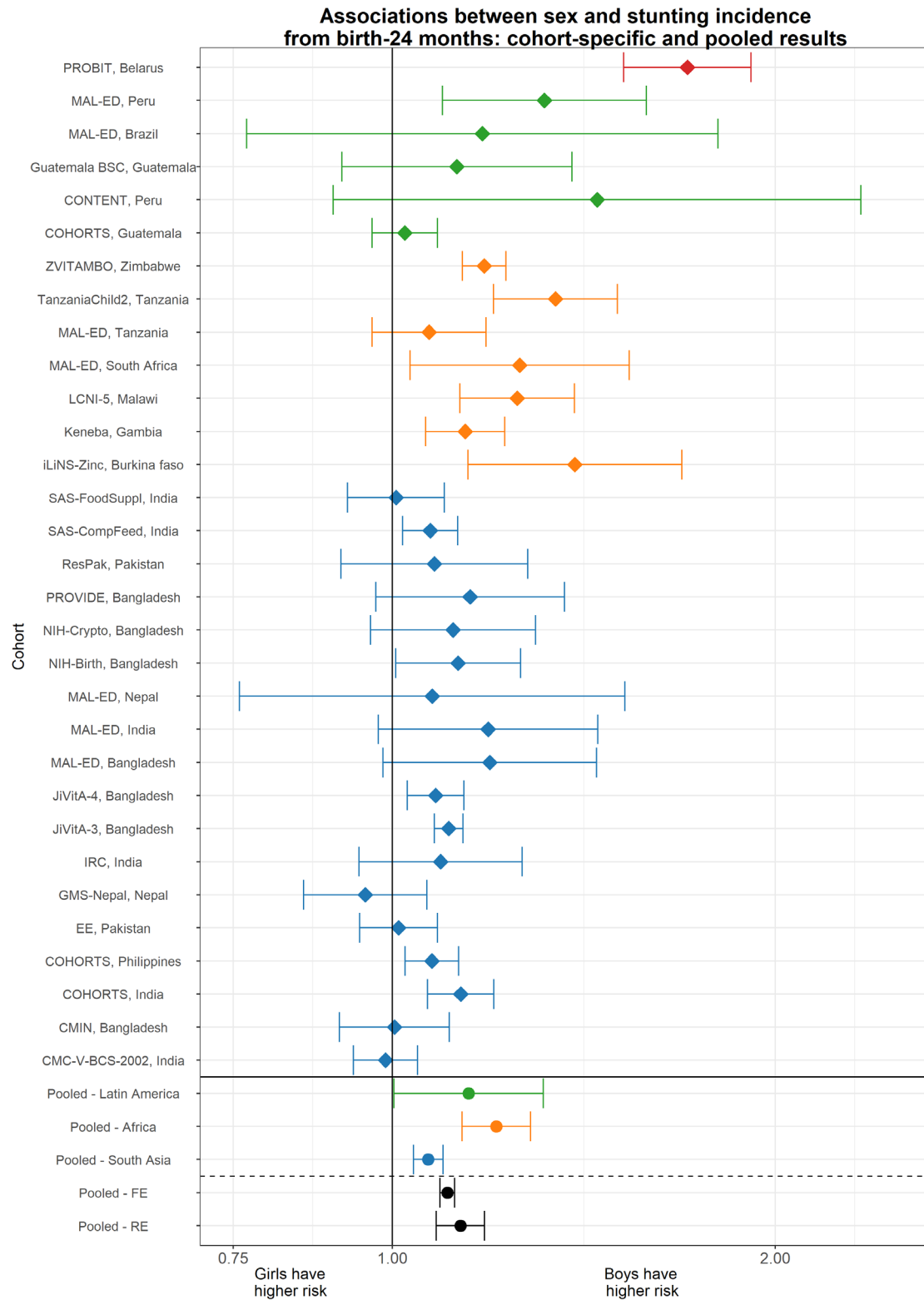


**Figure B.1 | Example forest plot of cohort-specific and pooled parameter estimates**

Cohort-specific estimates of the cumulative incidence ratio of stunting are plotted on each row, comparing the risk of any stunting from birth to 24 months among boys compared to a reference level of girls. Below the solid horizontal line are region-specific pooled measures of association, pooled using random-effects models. Below the dashed line are overall pooled measures of association, comparing pooling using random or fixed effects models. The primary results reported throughout the manuscript are overall (not region stratified) estimates pooled using random effects models.



**Figure B.2 | Heatmap of significance and direction across exposure-outcome combinations.**

The heatmap shows the significance and direction of estimates through the cell colors, separated across primary outcomes by child age. Red and orange cells are harmful exposures, while blue and green cells are protective exposures. The outcomes are labeled at the top of the columns, with each set of three columns the set of three ages analyzed for that outcome. Each row is a level of an exposure variable, with reference levels excluded. Rows are sorted top to bottom by increasing average p-value. Grey cells denote comparisons that were not estimated or could not be estimated because of data sparsity in the exposure-outcome combination.

## Attributable difference - LAZ, stratified by age

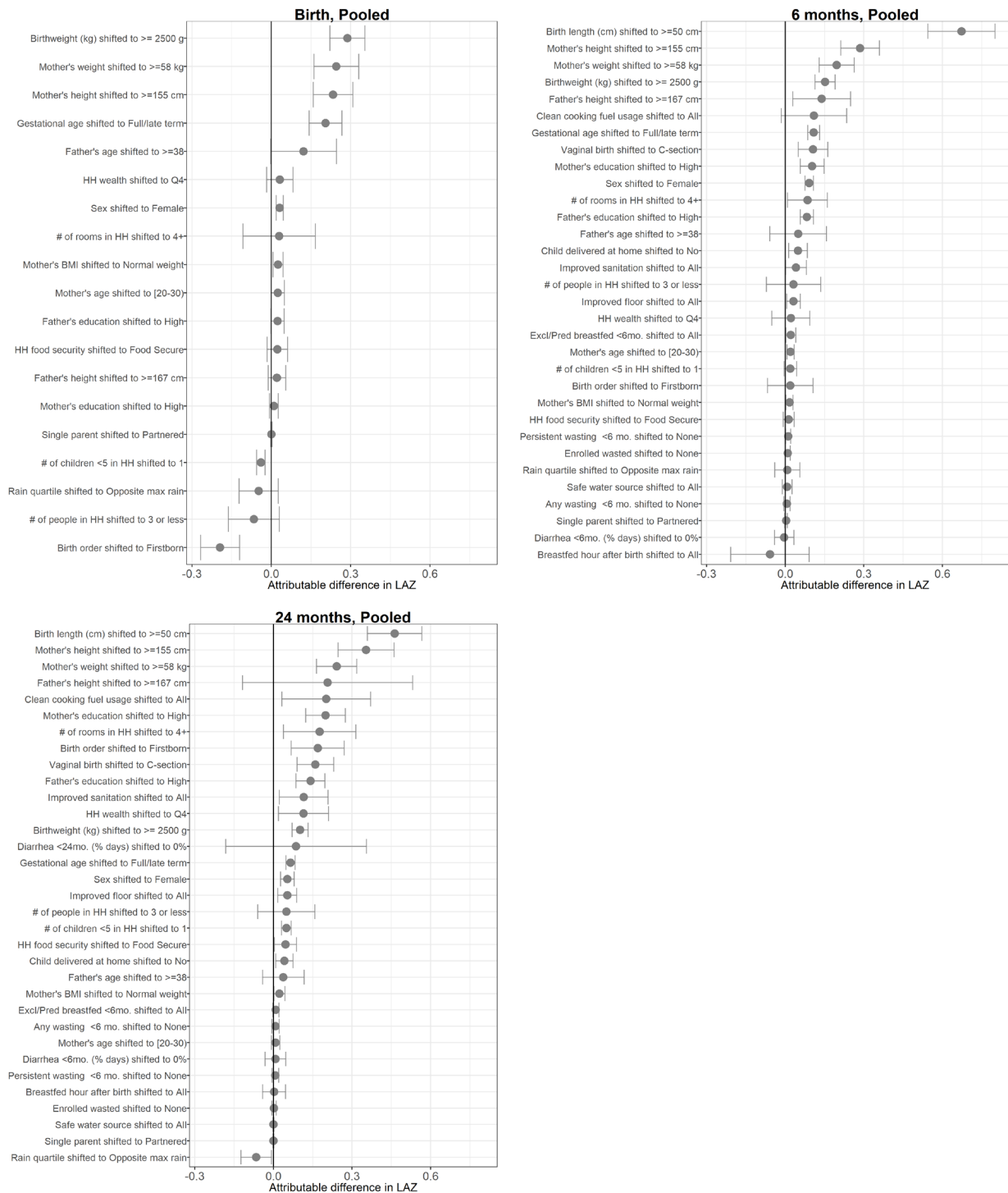
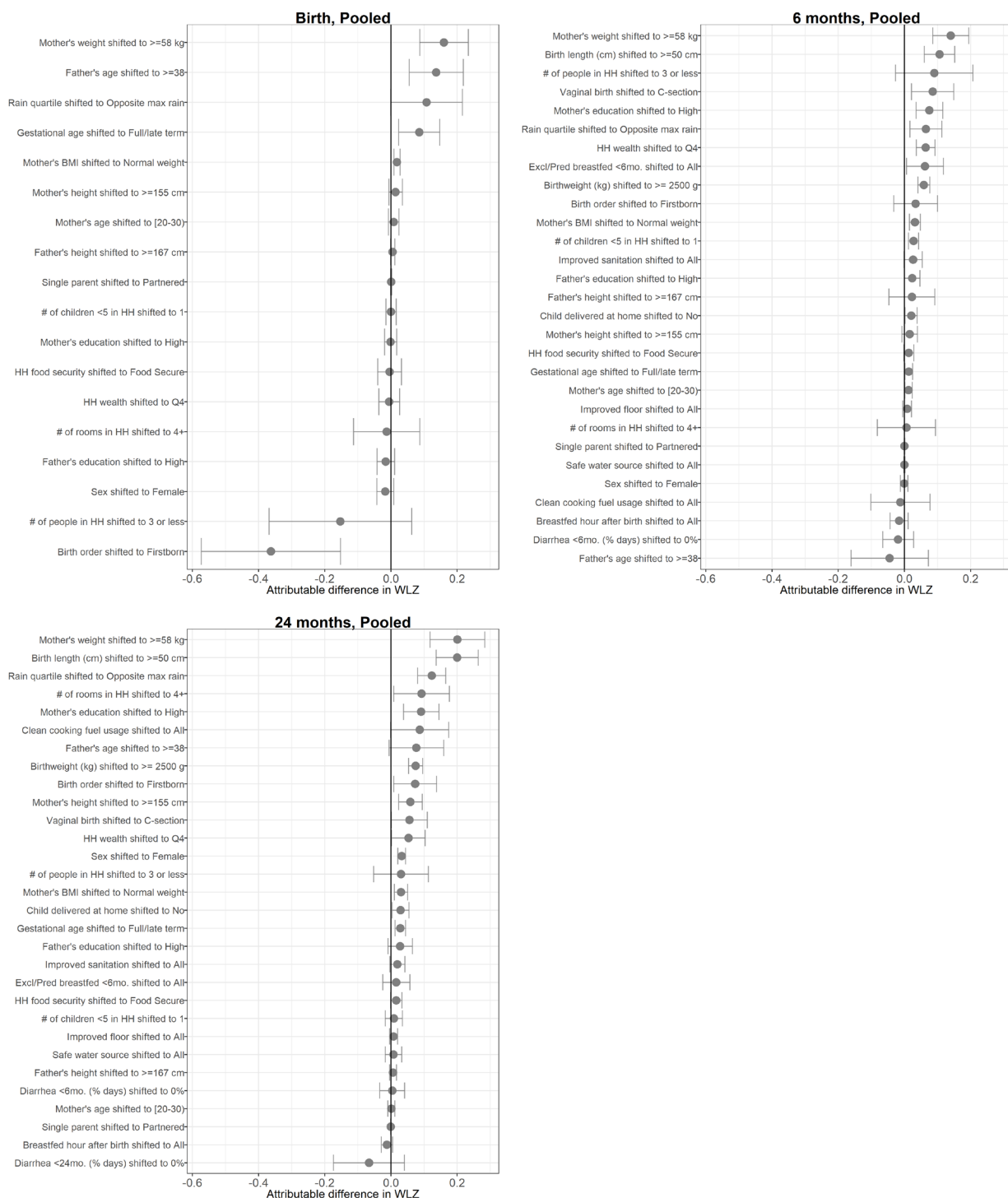


**Figure B.3 | Age-stratified population attributable differences in length-for-age Z-scores.**

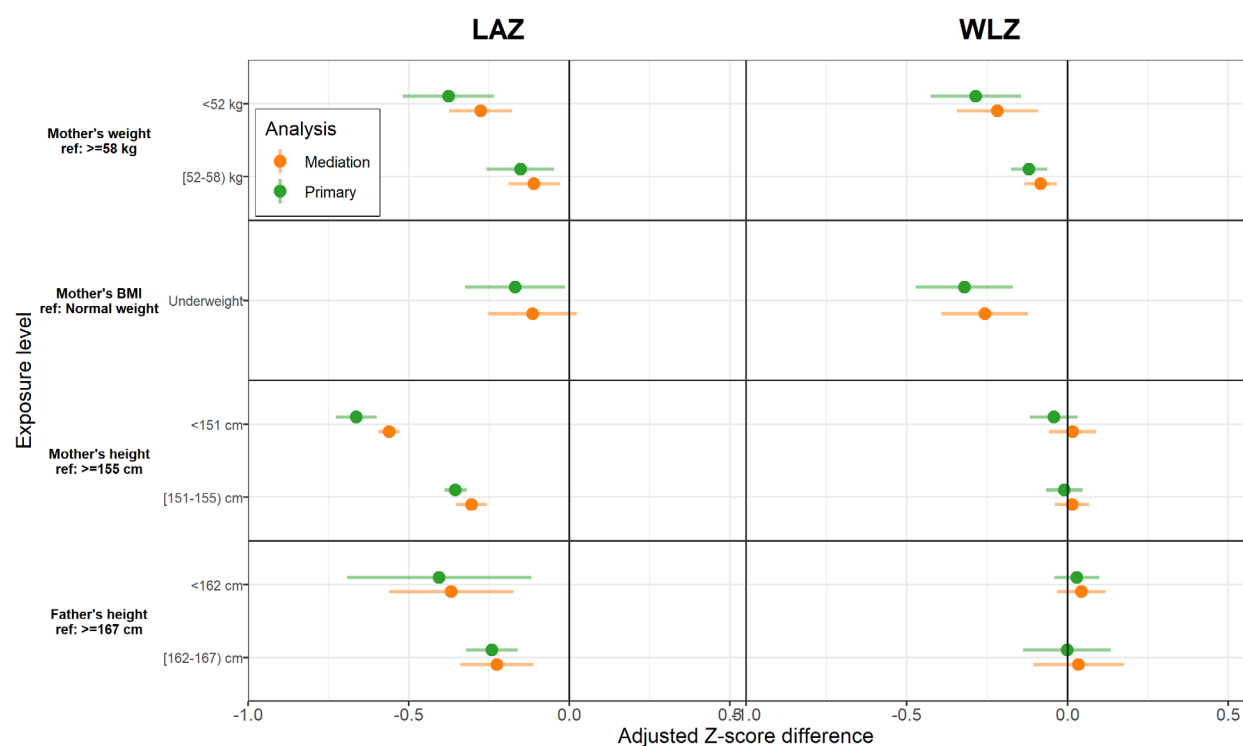
Exposures, rank ordered by population attributable difference on child LAZ at 24 months, stratified by the age of the child at the time of anthropometry measurement. The population attributable difference is the expected difference in population mean Z-score if all children had the reference level of the exposure rather than the observed distribution. For all plots, reference levels are printed next to the name of the exposure. Estimates are adjusted for all other measured exposures not on the causal pathway.

## Attributable difference - WLZ, stratified by age



## Figure B.4 | Age-stratified population attributable differences in weight-for-length Z-scores.

Exposures, rank ordered by population attributable difference on child WLZ at 24 months, stratified by the age of the child at the time of anthropometry measurement. The population attributable difference is the expected difference in population mean Z-score if all children had the reference level of the exposure rather than the observed distribution. For all plots, reference levels are printed next to the name of the exposure. Estimates are adjusted for all other measured exposures not on the causal pathway.



## Figure B.5 | Mediation by birth characteristics of associations between parental anthropometry and child Z-scores at 24 months.

The plot panels show the mediating effect of adjusting for child birth anthropometry and at-birth characteristics on the estimated Z-score differences between levels of parental anthropometry. There is little to partial mediation of the association between parental anthropometry and child LAZ and WLZ at 24 months by at-birth characteristics, implying that the causal pathway between parental stature and child growth failure operates through both affecting birth size as well as other pathways. Primary estimates were adjusted for all other measured exposures not on the causal pathway, while the mediation analysis estimates are adjusted for the same set of measured exposures, plus birth weight, birth length, gestational age at birth, birth order, vaginal birth vs. C-section, and home vs. hospital delivery. Only estimates from cohorts measuring at least 4 of the 6 at-birth characteristics are used to estimate the pooled Z-score differences.

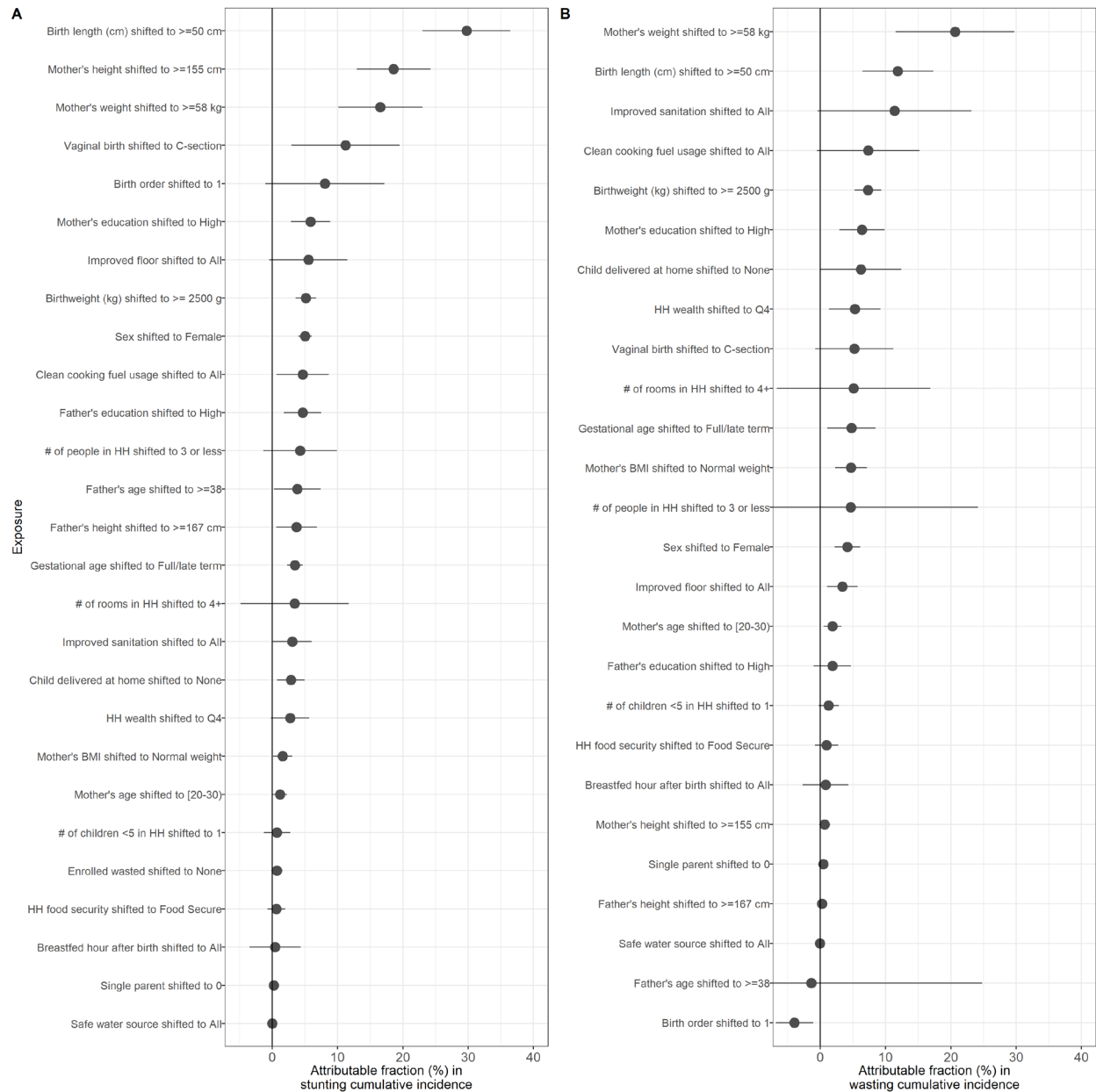


**Figure B.6 | Rank-ordered associations between child, parental, and household characteristics and population attributable fractions of stunting and wasting.**

(a) Exposures, rank ordered by population attributable difference on the cumulative incidence of child stunting between birth and 24 months. The population attributable fraction is the estimated proportion of the observed outcome in the whole population attributable to the exposure. For postnatal and at-birth exposures, including birth characteristics, breastfeeding practice, food security, and diarrheal disease, the cumulative incidence of stunting from 6-24 months is used.

(b) Exposures, rank ordered by population attributable difference on the cumulative incidence of child wasting between birth and 24 months. The population attributable fraction is the estimated proportion of the observed outcome in the

whole population attributable to the exposure. For postnatal and at-birth exposures, including birth characteristics, breastfeeding practice, food security, and diarrheal disease, the cumulative incidence of wasting from 6-24 months is used.

For all plots, reference levels are printed next to the name of the exposure. Estimates are adjusted for all other measured exposures not on the causal pathway.

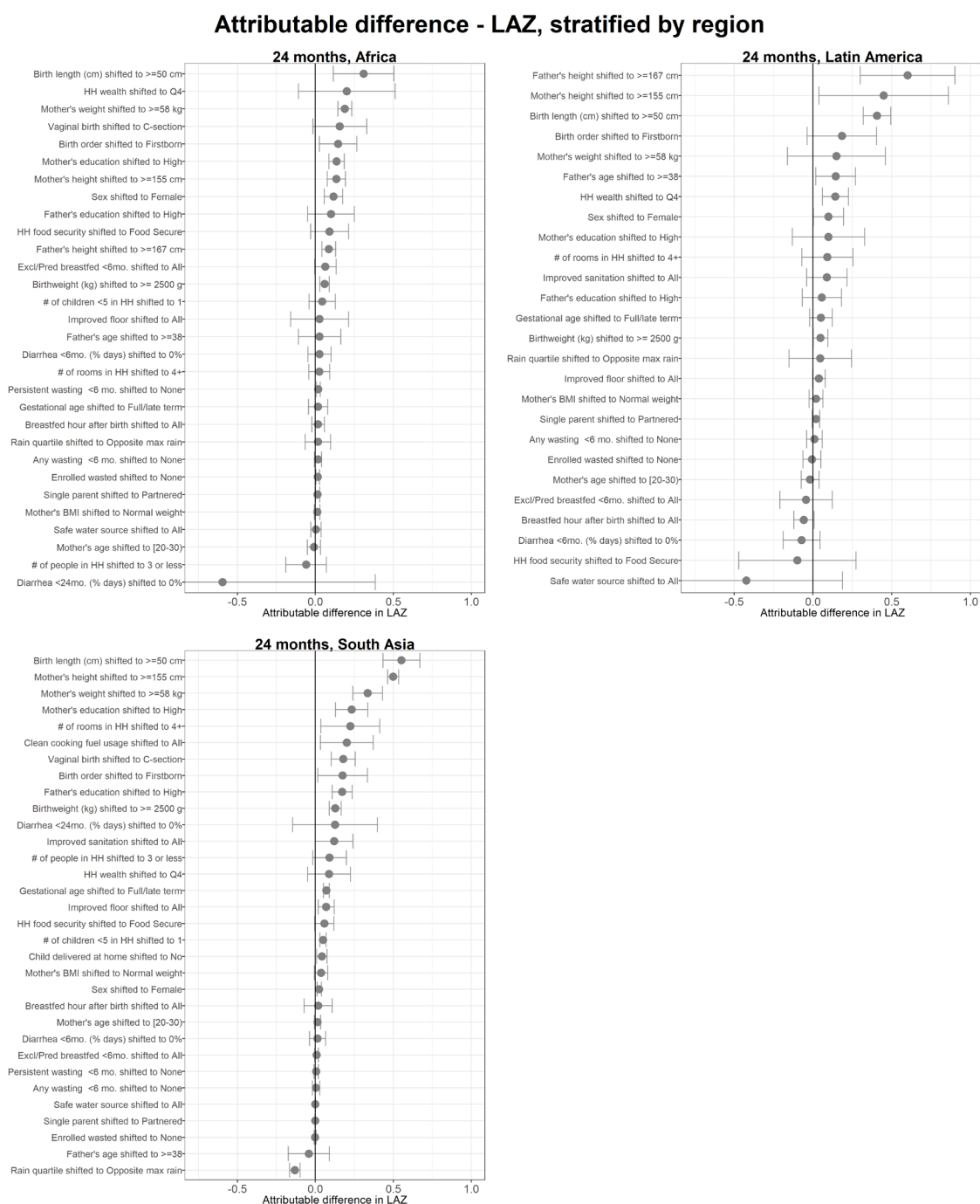


**Figure B.7 | Regionally-stratified population attributable differences in length-for-age Z-scores.**

Exposures, rank ordered by population attributable difference on child LAZ at 24 months, stratified by region. The population attributable difference is the expected difference in population mean Z-score if all children had the reference level of the exposure rather than the observed distribution. For all plots, reference levels are printed next to the name of the exposure. Estimates were adjusted for all other measured exposures not on the causal pathway. The confidence intervals extend beyond the scale for diarrhea under 24 months in Africa and for safe water source in Latin America.

## Attributable difference - WLZ, stratified by region

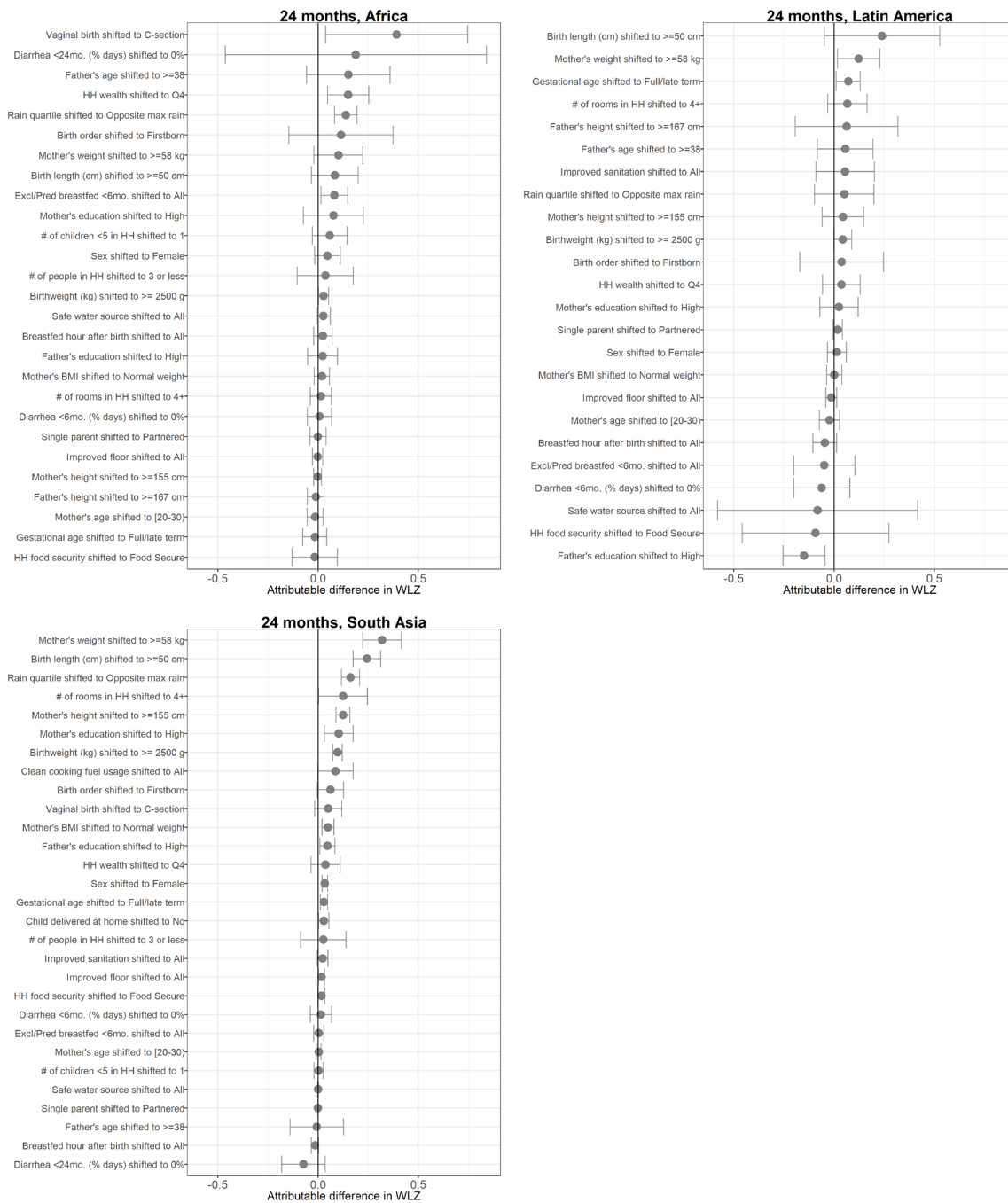


**Figure B.8 | Regionally-stratified population attributable differences in weight-for-length Z-scores.**

Exposures, rank ordered by population attributable difference on child WLZ at 24 months, stratified by region. The population attributable difference is the expected difference in population mean Z-score if all children had the reference level of the exposure rather than the observed distribution. For all plots, reference levels are printed next to the name of the exposure. Estimates were adjusted for all other measured exposures not on the causal pathway.

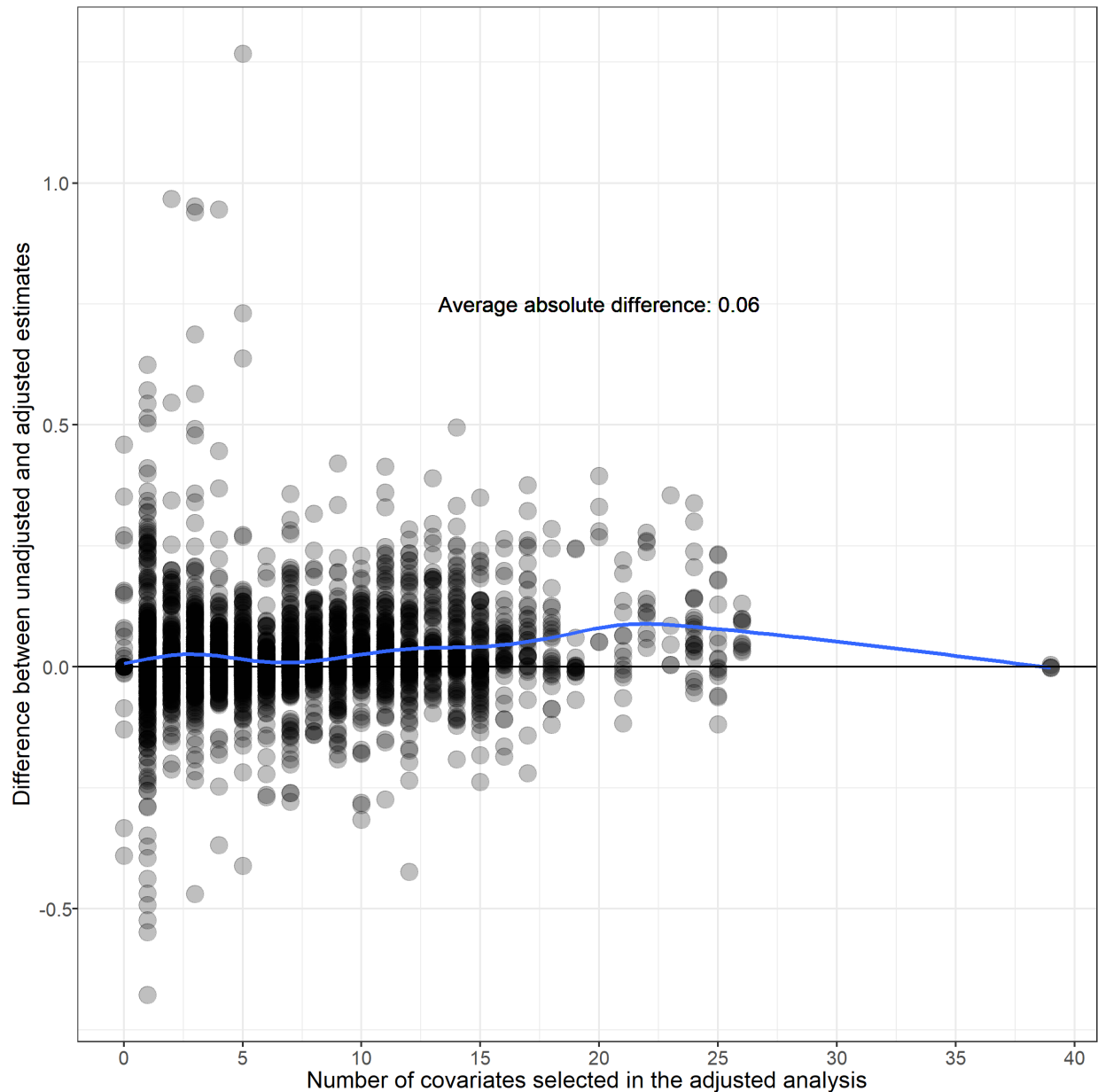
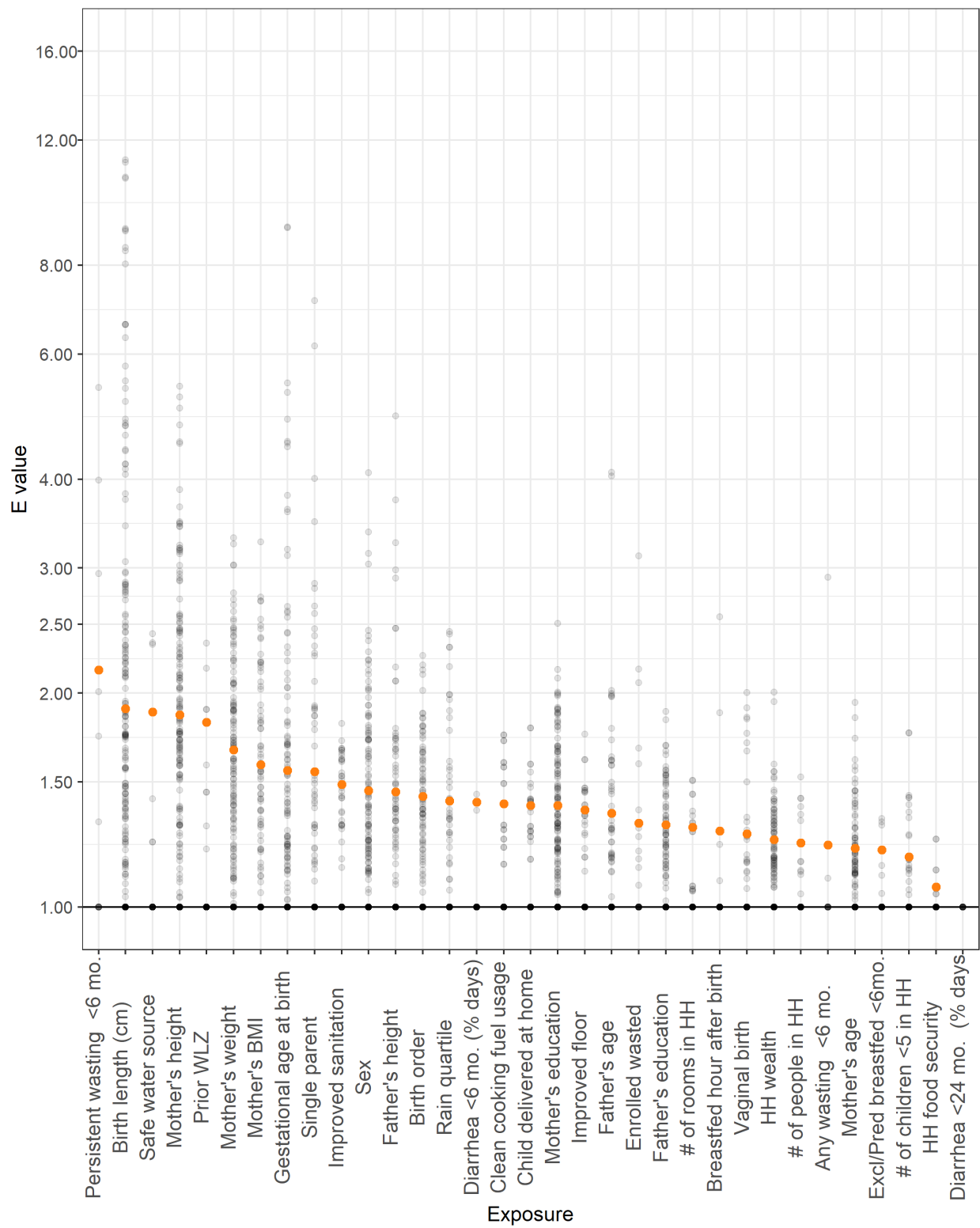


**Figure B.9 | Difference between adjusted and unadjusted Z-score effects by number of selected adjustment variables.**

Points mark the difference in estimates unadjusted and adjusted estimates of the difference in average Z-scores between exposed and unexposed children across 31 cohorts, 33 exposures and length-for-age and weight-for-length Z-score outcomes included in the analysis. Different cohorts measured different sets of exposures, and a different number of adjustment covariates were chosen for each cohort-specific estimate based on outcome sparsity, so cohort-specific estimates adjust for different covariates and numbers of covariates. The plot shows no systematic bias between unadjusted and adjusted estimates based on number of covariates chosen. The blue line shows the average difference between adjusted estimates from unadjusted estimates, fitted using a cubic spline.



## Figure B.10 | Assessing sensitivity of estimates to unmeasured confounding using E-values

An E-value is the minimum strength of association in terms of relative risk that an unmeasured confounder would need to have with both the exposure and the outcome to explain away an estimated exposure–outcome association.<sup>151</sup> Orange points mark the E-values for the pooled estimates of relative risk for each exposure. Grey points are cohort-specific E-values for each exposure-outcome relationship. Non-significant pooled estimates have points plotted at 1.0. Orange points are median E-values among statistically significant estimates for each exposure.

*Table B.1| Summary of KI cohorts used in Chapter 3*

| Region, Study ID          | Country    | Study Years    | Design             | Children Enrolled* | Anthropometry measurement ages (months)                  | Total measurements* | Primary References                         |
|---------------------------|------------|----------------|--------------------|--------------------|----------------------------------------------------------|---------------------|--------------------------------------------|
| South Asia                |            |                |                    |                    |                                                          |                     |                                            |
| Biomarkers for EE         | Pakistan   | 2013-2015      | Prospective cohort | 380                | Birth, 1, 2, ..., 18                                     | 8918                | Iqbal et al 2018 Nature Scientific Reports |
| Resp. Pathogens           | Pakistan   | 2011 - 2014    | Prospective cohort | 284                | Birth, 1, 2, ..., 17                                     | 3177                | Ali et al-2016-Journal_of_Medical_Virology |
| Growth Monitoring Study   | Nepal      | 2012 - Ongoing | Prospective cohort | 698                | Birth, 1, 2, ..., 24                                     | 13487               | Not yet published                          |
| MAL-ED                    | Nepal      | 2010 - 2014    | Prospective cohort | 240                | Birth, 1, 2, ..., 24                                     | 5936                | Shrestha et al 2014 Clin Infect Dis NPL    |
| CMC Birth Cohort, Vellore | India      | 2002 - 2006    | Prospective cohort | 373                | Birth, 0.5, 1, 1.5, ..., 24                              | 9131                | Gladstone et al. 2011 NEJM                 |
| MAL-ED                    | India      | 2010 - 2012    | Prospective cohort | 251                | Birth, 1, 2, ..., 24                                     | 5947                | John et al 2014 Clin Infect Dis IND        |
| Vellore Crypto Study      | India      | 2008 - 2011    | Prospective cohort | 410                | Birth, 1, 2, ..., 24                                     | 9825                | Kattula et al. 2014 BMJ Open               |
| CMIN                      | Bangladesh | 1993 - 1996    | Prospective Cohort | 280                | Birth, 3, 6, ..., 24                                     | 5399                | Pathela et al 2006 Acta Paediatrica        |
| MAL-ED                    | Bangladesh | 2010 - 2014    | Prospective cohort | 265                | Birth, 1, 2, ..., 24                                     | 5816                | Ahmed et al 2014 Clin Infect Dis           |
| PROVIDE RCT               | Bangladesh | 2011 - 2014    | Individual RCT     | 700                | Birth, 6, 10, 12, 14, 17, 18, 24, 39, 40, 52, 53 (weeks) | 12165               | Kirkpatrick et al 2015 Am J Trop Med Hyg   |

|                                                          |              |             |                    |       |                            |        |                                       |
|----------------------------------------------------------|--------------|-------------|--------------------|-------|----------------------------|--------|---------------------------------------|
| Food Suppl RCT                                           | India        | 1995 - 1996 | Individual RCT     | 418   | Baseline, 6, 9, 12         | 2242   | Bhandari et al 2001 J Nutri           |
| Optimal Infant Feeding                                   | India        | 1999 - 2001 | Cluster RCT        | 1535  | Birth, 3, 6, ..., 18       | 9539   | Bhandari et al 2004 J Nutri           |
| New Delhi Birth Cohort                                   | India        | 1969 - 2002 | Prospective cohort | 7452  | Birth, 3, 6, 9, 12, 18, 24 | 32463  | Bhargava et al 2009 NEJM              |
| NIH Birth Cohort                                         | Bangladesh   | 2008 - 2009 | Prospective Cohort | 629   | Birth, 3, 6, ..., 12       | 6216   | Korpe et al. 2016 PLOS NTD            |
| JiVitA-4 Trial                                           | Bangladesh   | 2012 - 2014 | Cluster RCT        | 5444  | 6, 9, 12, 14, 18           | 36167  | Christian et al 2015 IJE              |
| JiVitA-3 Trial                                           | Bangladesh   | 2008 - 2012 | Cluster RCT        | 27342 | Birth, 1, 3, 6, 12, 24     | 109535 | West et al JAMA 2014                  |
| NIH Cryptosporidium Study                                | Bangladesh   | 2014 - 2017 | Prospective cohort | 758   | Birth, 3, 6, ..., 24       | 9774   | Steiner et al 2018 Clin Infect Dis    |
| The Cebu Longitudinal Health and Nutrition Survey        | Philippines  | 1983 - 2005 | Prospective cohort | 3073  | Birth, 2, 4, ..., 24       | 32636  | Adair 2007 Am J Hum Biol              |
| Africa                                                   |              |             |                    |       |                            |        |                                       |
| MAL-ED                                                   | Tanzania     | 2009 - 2014 | Prospective cohort | 262   | Birth, 1, 2, ..., 24       | 5857   | Mduma et al 2014 Clin Infect Dis TZN  |
| Tanzania Child 2                                         | Tanzania     | 2007 - 2011 | Individual RCT     | 2400  | 1, 2, ..., 20              | 32198  | Locks et al Am J Clin Nutr 2016       |
| MAL-ED                                                   | South Africa | 2009 - 2014 | Prospective cohort | 314   | Birth, 1, 2, ..., 24       | 6478   | Bessong et al 2014 Clin Infect Dis SA |
| MRC Keneba                                               | Gambia       | 1987 - 1997 | Cohort             | 2931  | Birth, 1, 2, ..., 24       | 40952  | Schoenbuchner et al. 2019, AJCN       |
| ZVITAMBO Trial                                           | Zimbabwe     | 1997 - 2001 | Individual RCT     | 14104 | Birth, 6 wks, 3, 6, 9, 12  | 73651  | Malaba et al 2005 Am J Clin Nutr      |
| Lungwena Child Nutrition RCT                             | Malawi       | 2011 - 2014 | Individual RCT     | 840   | Birth, 1-6 wks, 6, 12 18   | 4346   | Mangani et al. 2015, Mat Child Nutr   |
| iLiNS-Zinc Study                                         | Burkina Faso | 2010 - 2012 | Cluster RCT        | 3266  | 9, 12, 15, 18              | 10552  | Hess et al 2015 PlosOne               |
| Latin America                                            |              |             |                    |       |                            |        |                                       |
| MAL-ED                                                   | Peru         | 2009 - 2014 | Prospective cohort | 303   | Birth, 1, 2, ..., 24       | 6442   | Yori et al 2014 Clin Infect Dis PER   |
| CONTENT                                                  | Peru         | 2007 - 2011 | Prospective cohort | 215   | Birth, 1, 2, ..., 24       | 8339   | Jaganath et al 2014 Helicobacter      |
| Bovine Serum RCT                                         | Guatemala    | 1997 - 1998 | Individual RCT     | 315   | Baseline, 1, 2, ..., 8     | 2551   | Begin et al. 2008, EJCN               |
| MAL-ED                                                   | Brazil       | 2010 - 2014 | Prospective cohort | 233   | Birth, 1, 2, ..., 24       | 5092   | Lima et al 2014 Clin Infect Dis       |
| INCAP Nutrition Supplementation Trial Longitudinal Study | Guatemala    | 1969 - 1977 | Cluster RCT        | 1406  | Birth, 3, 6, ..., 24       | 7965   | Habicht et al 1995 J Nutr             |
| Europe                                                   |              |             |                    |       |                            |        |                                       |
| PROBIT Study                                             | Belarus      | 1996 - 1997 | Cluster RCT        | 16898 | 1, 2, 3, 6, 9, 12          | 124509 | Kramer et al 2001 JAMA                |

| Mortality analysis only                                                                                                                                                     |              |           |                |      |                   |       |                                     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-----------|----------------|------|-------------------|-------|-------------------------------------|
| Burkina Faso Zinc trial                                                                                                                                                     | Burkina Faso | 2010-2011 | Cluster RCT    | 7167 | 6, 10, 14, 17, 22 | 15155 | Becquey et al 2016 J Nutr           |
| Vitamin A Trial                                                                                                                                                             | India        | 1995-1996 | Cluster RCT    | 3983 | 1, 3, 6, 9, 12    | 32570 | WHO CHD Vitamin A Group 1998 Lancet |
| iLiNS-DOSE                                                                                                                                                                  | Malawi       | 2009-2011 | Individual RCT | 1932 | 6, 9, 12, 18      | 13801 | Maleta et al. 2015 J Nutr           |
| iLiNS-DYAD-M                                                                                                                                                                | Malawi       | 2011-2015 | Individual RCT | 1235 | 1, 6, 12, 18      | 9207  | Ashorn et al 2015 J. Nutr           |
| *Children enrolled is for children with measurements under 2 years of age. Total measurements are number of measurements of anthropometry on children under 2 years of age. |              |           |                |      |                   |       |                                     |

*Table B.2| List of potential causes of growth failure analyzed in Chapter 3*

All exposures included in the analysis, as well as the categories the exposures were classified into across all cohorts, categorization rules, and the total number of children and percent of children in each category. I selected the exposures of interest based on variables present in multiple cohorts that met my inclusion criteria, were found to be important determinants of stunting and wasting in prior literature and could be harmonized across cohorts for pooled analyses.

| Exposure variable        | N children under 24 months with both measured exposure and length | Exposure levels [N (%)]<br>First listed level is reference                                   | Categorization rules                                                    |
|--------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| Sex                      | 93756                                                             | Female: 45638 (48.7%)<br>Male: 48118 (51.3%)                                                 |                                                                         |
| Gestational age at birth | 53500                                                             | Full or late term: 27079 (50.6%)<br>Early term: 18377 (34.3%)<br>Preterm: 8044 (15%)         | <260 days is preterm, [260-274) days is early term, >= 274 is full term |
| Birthweight (kg)         | 80162                                                             | Normal or high birthweight: 63631 (79.4%)<br>Low birthweight: 16531 (20.6%)                  |                                                                         |
| Birth length (cm)        | 75832                                                             | >=50 cm: 26909 (35.5%)<br>[48-50) cm: 19308 (25.5%)<br><48 cm: 29615 (39.1%)                 |                                                                         |
| Birth order              | 58044                                                             | Firstborn: 20041 (34.5%)<br><br>Secondborn: 17125 (29.5%)<br>Thirdborn or later: 20878 (36%) |                                                                         |
| Delivery location        | 17158                                                             | Hospital or clinic: 9123 (53.2%)<br>Home: 8035 (46.8%)                                       |                                                                         |
| Delivery method          | 66826                                                             | C-section: 5684 (8.5%)<br>Vaginal birth: 61142 (91.5%)                                       |                                                                         |

|                                    |       |                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------------------------------|-------|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Maternal weight                    | 63917 | >=58 kg: 20534 (32.1%)<br>[52-58) kg: 9868 (15.4%)<br><52 kg: 33515 (52.4%)                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Maternal height                    | 71666 | >=155 cm: 35237 (49.2%)<br>[151-155) cm: 12996 (18.1%)<br><151 cm: 23433 (32.7%)                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Maternal body mass index (BMI)     | 62281 | Normal weight: 49423 (79.4%)<br>Underweight: 12858 (20.6%)                                                   | WHO categories (<18.5 BMI is underweight) Excludes mothers whose only weight measurement was taken during pregnancy.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| Mother's age                       | 85002 | [20-30): 50917 (59.9%)<br><20: 19310 (22.7%)<br>>=30: 14775 (17.4%)                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Maternal education                 | 84912 | High: 26564 (31.3%)<br>Medium: 30319 (35.7%)<br>Low: 28029 (33%)                                             | Classified by splitting distribution of numbers of years of educations into thirds within each cohort, or grouping ordered categories of educational attainment into three levels.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Paternal height                    | 16268 | >=167 cm: 13569 (83.4%)<br>[162-167) cm: 1713 (10.5%)<br><162 cm: 986 (6.1%)                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Paternal age                       | 23328 | >=38: 1873 (8%)<br>[32-38): 3912 (16.8%)<br><32: 17543 (75.2%)                                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Paternal education                 | 75358 | High: 15623 (20.7%)<br>Medium: 32968 (43.7%)<br>Low: 26767 (35.5%)                                           | Classified by splitting distribution of numbers of years of educations into thirds within each cohort, or grouping ordered categories of educational attainment into three levels.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Caregiver marital status           | 50468 | Married/partnered: 48223 (95.6%)<br>Single/widow/partner overseas: 2245 (4.4%)                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Asset based household wealth index | 44263 | Wealth Q4: 11671 (26.4%)<br>Wealth Q1: 10844 (24.5%)<br>Wealth Q2: 10496 (23.7%)<br>Wealth Q3: 11252 (25.4%) | First principal component of a principal components analysis of all recorded assets owned by the household (examples: cell phone, bicycle, car).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Household food security            | 25848 | Food Secure: 13082 (50.6%)<br>Mildly Food Insecure: 8632 (33.4%)<br>Food Insecure: 4134 (16%)                | Combination of three food security scales:<br><ol style="list-style-type: none"> <li>1. The Household Hunger Scale (HHS)<sup>222</sup></li> <li>2. Food Access Survey Tool (FAST)<sup>223</sup></li> <li>3. USAID Household Food Insecurity Access Scale (HFIAS), with middle 2 categories classified as mildly food insecure.<sup>224</sup></li> </ol> <p>And one survey question from the NIH Bangladesh birth cohort and NIH Bangladesh Cryptosporidium cohort:<br/>         “In terms of household food availability, how do you classify your household?”</p> <ol style="list-style-type: none"> <li>1. Deficit in whole year</li> <li>2. Sometimes deficit</li> </ol> |

|                                                                      |       |                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----------------------------------------------------------------------|-------|------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                      |       |                                                                                                                        | <p>3. Neither deficit nor surplus</p> <p>4. Surplus</p> <p>Where the middle two categories were classified as mildly food insecure.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Improved floor                                                       | 37686 | Yes: 6477 (17.2%)<br>No: 31209 (82.8%)                                                                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Improved sanitation                                                  | 44216 | Yes: 27350 (61.9%)<br>No: 16866 (38.1%)                                                                                | WHO Joint Monitoring program definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Improved water source                                                | 36980 | Yes: 35392 (95.7%)<br>No: 1588 (4.3%)                                                                                  | WHO Joint Monitoring program definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Clean cooking fuel usage                                             | 2678  | Yes: 1338 (50%)<br>No: 1340 (50%)                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Number of children <5 in the household                               | 35453 | One: 21653 (61.1%)<br>Two or more: 13800 (38.9%)                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Number of individuals in the household                               | 3732  | 3 or less: 692 (18.5%)<br>4-5: 1609 (43.1%)<br>6-7: 887 (23.8%)<br>8+: 544 (14.6%)                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Number of rooms in household                                         | 42851 | Four or more: 2819 (6.6%)<br>Three: 4155 (9.7%)<br>Two: 11351 (26.5%)<br>One: 24526 (57.2%)                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Rain season                                                          | 7462  | Opposite max rain: 1579 (21.1%)<br>Pre max rain: 1884 (25.3%)<br>Max rain: 2170 (29.1%)<br>Post max rain: 1829 (24.5%) | Rainfall data was extracted from the University of East Anglia Climate Research Unit CRU TS dataset (version 3.23, 23 October 2015) and averaged by month over the duration of cohorts. <sup>92</sup> The three-month period opposite the three months of maximum rainfall was used as the reference level (e.g., if July-September was the period of maximum rainfall, the reference level is child mean WLZ during January-March). Due to the time-varying nature of this exposure, N's are reported for children with length measures at 24 months and measures of rain season. |
| Breastfed hour after birth                                           | 50188 | Yes: 12483 (24.9%)<br>No: 37705 (75.1%)                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Exclusive or predominant breastfeeding in the first 6 months of life | 26948 | Yes: 18471 (68.5%)<br>No: 8477 (31.5%)                                                                                 | Exclusive breastfeeding: mother reported only feeding child breastmilk on all dietary surveys<br>Predominant breastfeeding: mother reported only feeding child breastmilk, other liquids, or medicines on all dietary surveys                                                                                                                                                                                                                                                                                                                                                      |
| Cumulative percent of days with diarrhea under 6 months              | 4812  | 0%: 2196 (45.6%)<br>(0%, 5%]: 1336 (27.8%)<br>>5%: 1280 (26.6%)                                                        | Percent days defined as proportion of disease surveillance days a child had diarrhea during. Diarrhea defined by 3 or more loose stools, or bloody stool, in a 24-hour period. Only included studies with at least 100 disease surveillance measurements during age range.                                                                                                                                                                                                                                                                                                         |

|                                                          |            |                                                                                    |                                                                                                                                                                                                                                                                            |
|----------------------------------------------------------|------------|------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cumulative percent of days with diarrhea under 24 months | 13430      | 0%: 1600 (11.9%)<br>(0%, 5%]: 8705 (64.8%)<br>>5%: 3125 (23.3%)                    | Percent days defined as proportion of disease surveillance days a child had diarrhea during. Diarrhea defined by 3 or more loose stools, or bloody stool, in a 24-hour period. Only included studies with at least 100 disease surveillance measurements during age range. |
| Enrolled or born stunted                                 | 96217      | No: 79110 (82.2%)<br>Yes: 17107 (17.8%)                                            |                                                                                                                                                                                                                                                                            |
| Enrolled or born wasted                                  | 8664189263 | No: 73722 (85.1%)<br>Yes: 12919 (14.9%)<br>No: 76610 (85.8%)<br>Yes: 12653 (14.2%) |                                                                                                                                                                                                                                                                            |
| Any wasting under 6 months                               | 10328      | No: 7849 (76%)<br>Yes: 2479 (24%)                                                  |                                                                                                                                                                                                                                                                            |
| Persistently wasting under 6 months                      | 9631       | No: 9176 (95.3%)<br>Yes: 455 (4.7%)                                                |                                                                                                                                                                                                                                                                            |

## Appendix C: Supplementary material for Chapter 4

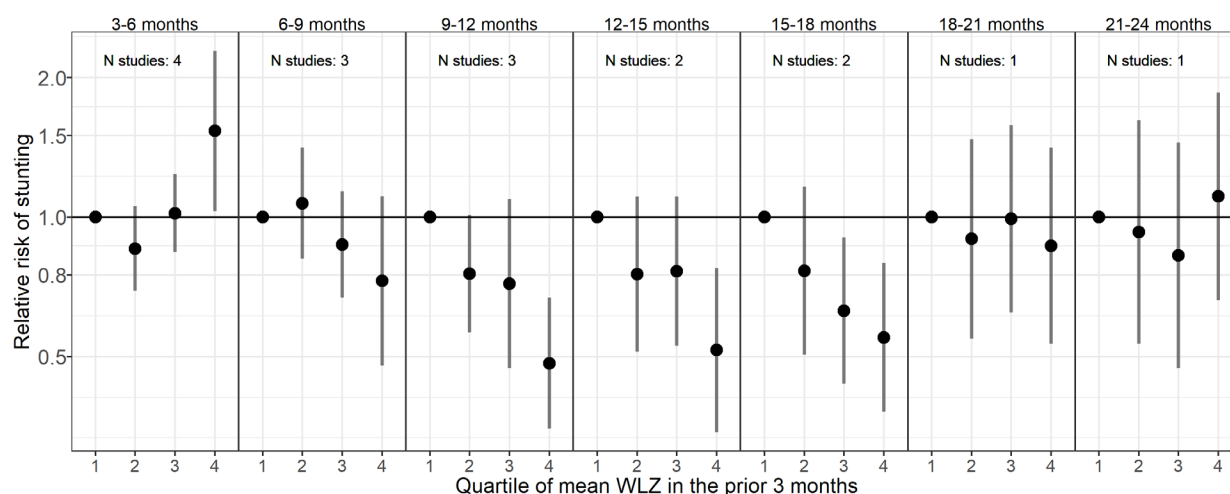


**Figure C.1 | Risks of subsequent linear-growth failure across categories of prior WLZ.**

Relative risk of stunting onset within 3-month age bands, comparing cohort-specific quartiles on mean WLZ in the prior 3-month age to the reference level of the first quartile. Children are assumed to not recover from stunting, so only the first measure of  $LAZ < -2$  is used to define stunting onset.

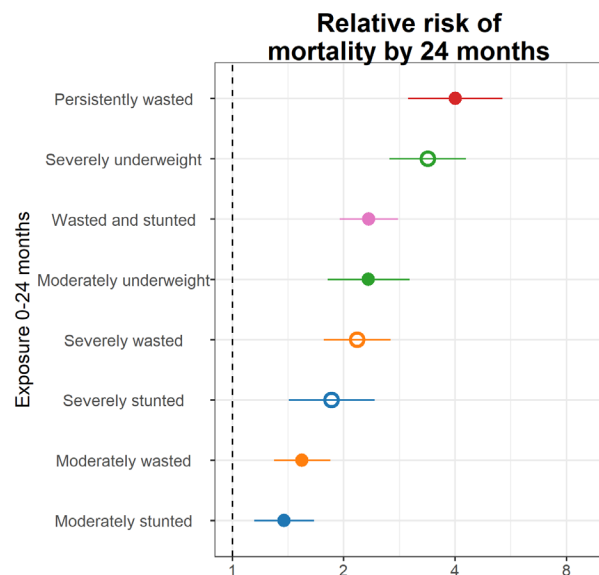


**Figure C.2 | Risks of mortality due to growth failure before 24 months.**

Relative risk of all-cause mortality by measure of growth failure before age 24 months. Estimates included up to 8 cohorts, 2,418 deaths, and 57,903 children. Estimates compare risk among children with each measure of growth failure before age 24 months with risk among children who did not experience the specific measure of growth failure.

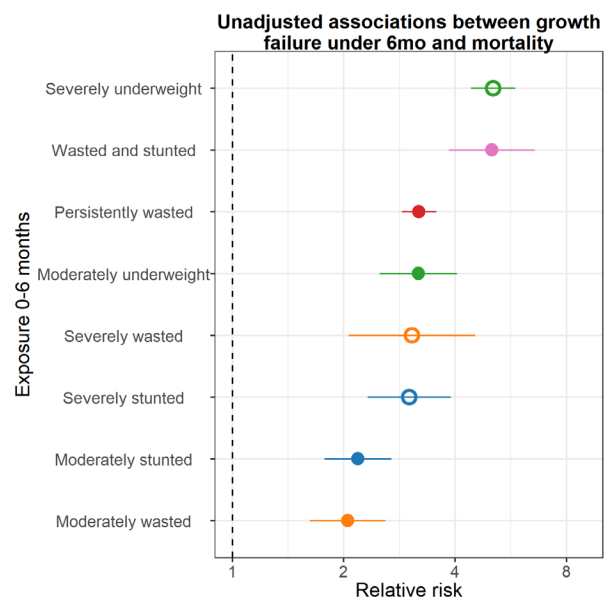


**Figure C.3 | Unadjusted risks of mortality due to growth failure before 6 months.**

Relative risk of all-cause mortality by measure of growth failure before age 6 months, unadjusted for potential confounders. Estimates included up to 8 cohorts, 2,418 deaths, and 57,903 children. Estimates compare risk among children with each measure of growth failure before age 6 months with risk among children who did not experience the specific measure of growth failure.

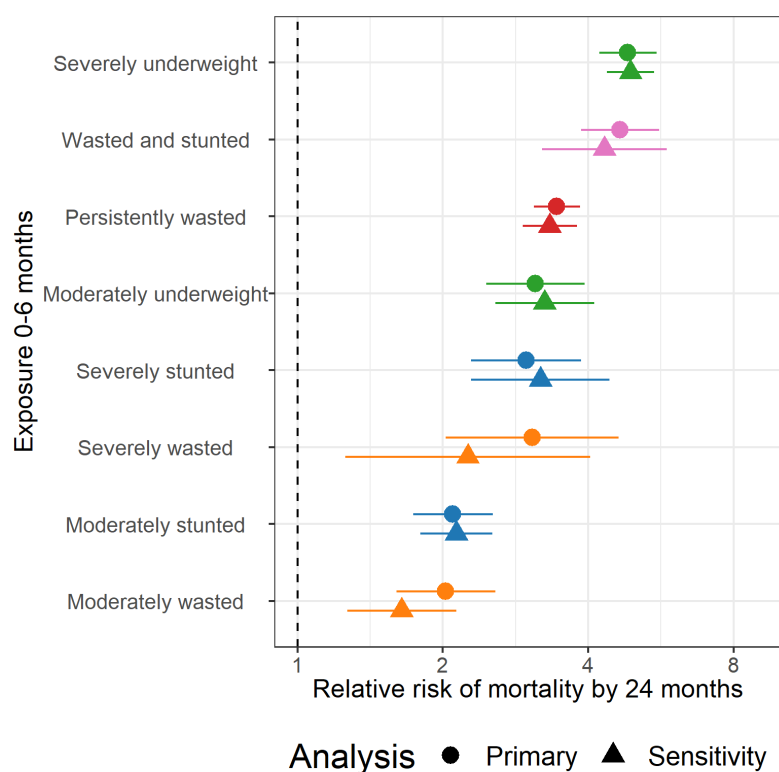


**Figure C.4 | Sensitivity analysis on the effect of excluding children without anthropometry measures in the first month of life on estimated associations between growth failure under 6mo and mortality**

Relative risk of all-cause mortality by measure of growth failure before age 6 months, estimated including (primary analysis) and excluding (sensitivity analysis) children with no anthropometry measurements in the first month of life. Estimates compare risk among children with each measure of growth failure before age 6 months with risk among children who did not experience the specific measure of growth failure. Primary estimates included up to 8 cohorts, 2,418 deaths, and 57,903 children, while sensitivity estimates included up to 7 cohorts, 2,239 deaths, and 52,219 children.

**Table C.1| Summary of KI cohorts used in Chapter 4**

| Region, Study ID                                  | Country     | Study Years    | Design             | Children Enrolled* | Anthropometry measurement ages (months)                  | Total measurements* | Primary References                         |
|---------------------------------------------------|-------------|----------------|--------------------|--------------------|----------------------------------------------------------|---------------------|--------------------------------------------|
| South Asia                                        |             |                |                    |                    |                                                          |                     |                                            |
| Biomarkers for EE                                 | Pakistan    | 2013-2015      | Prospective cohort | 380                | Birth, 1, 2, ..., 18                                     | 8918                | Iqbal et al 2018 Nature Scientific Reports |
| Resp. Pathogens                                   | Pakistan    | 2011 - 2014    | Prospective cohort | 284                | Birth, 1, 2, ..., 17                                     | 3177                | Ali_et_al-2016-Journal_of_Medical_Virology |
| Growth Monitoring Study                           | Nepal       | 2012 - Ongoing | Prospective cohort | 698                | Birth, 1, 2, ..., 24                                     | 13487               | Not yet published                          |
| MAL-ED                                            | Nepal       | 2010 - 2014    | Prospective cohort | 240                | Birth, 1, 2, ..., 24                                     | 5936                | Shrestha et al 2014 Clin Infect Dis NPL    |
| CMC Birth Cohort, Vellore                         | India       | 2002 - 2006    | Prospective cohort | 373                | Birth, 0.5, 1, 1.5, ..., 24                              | 9131                | Gladstone et al. 2011 NEJM                 |
| MAL-ED                                            | India       | 2010 - 2012    | Prospective cohort | 251                | Birth, 1, 2, ..., 24                                     | 5947                | John et al 2014 Clin Infect Dis IND        |
| Vellore Crypto Study                              | India       | 2008 - 2011    | Prospective cohort | 410                | Birth, 1, 2, ..., 24                                     | 9825                | Kattula et al. 2014 BMJ Open               |
| CMIN                                              | Bangladesh  | 1993 - 1996    | Prospective Cohort | 280                | Birth, 3, 6, ..., 24                                     | 5399                | Pathela et al 2006 Acta Paediatrica        |
| MAL-ED                                            | Bangladesh  | 2010 - 2014    | Prospective cohort | 265                | Birth, 1, 2, ..., 24                                     | 5816                | Ahmed et al 2014 Clin Infect Dis           |
| PROVIDE RCT                                       | Bangladesh  | 2011 - 2014    | Individual RCT     | 700                | Birth, 6, 10, 12, 14, 17, 18, 24, 39, 40, 52, 53 (weeks) | 12165               | Kirkpatrick et al 2015 Am J Trop Med Hyg   |
| Food Suppl RCT                                    | India       | 1995 - 1996    | Individual RCT     | 418                | Baseline, 6, 9, 12                                       | 2242                | Bhandari et al 2001 J Nutri                |
| Optimal Infant Feeding                            | India       | 1999 - 2001    | Cluster RCT        | 1535               | Birth, 3, 6, ..., 18                                     | 9539                | Bhandari et al 2004 J Nutri                |
| New Delhi Birth Cohort                            | India       | 1969 - 2002    | Prospective cohort | 7452               | Birth, 3, 6, 9, 12, 18, 24                               | 32463               | Bhargava et al 2009 NEJM                   |
| NIH Birth Cohort                                  | Bangladesh  | 2008 - 2009    | Prospective Cohort | 629                | Birth, 3, 6, ..., 12                                     | 6216                | Korpe et al. 2016 PLOS NTD                 |
| JiVitA-4 Trial                                    | Bangladesh  | 2012 - 2014    | Cluster RCT        | 5444               | 6, 9, 12, 14, 18                                         | 36167               | Christian et al 2015 IJE                   |
| JiVitA-3 Trial                                    | Bangladesh  | 2008 - 2012    | Cluster RCT        | 27342              | Birth, 1, 3, 6, 12, 24                                   | 109535              | West et al JAMA 2014                       |
| NIH Cryptosporidium Study                         | Bangladesh  | 2014 - 2017    | Prospective cohort | 758                | Birth, 3, 6, ..., 24                                     | 9774                | Steiner et al 2018 Clin Infect Dis         |
| The Cebu Longitudinal Health and Nutrition Survey | Philippines | 1983 - 2005    | Prospective cohort | 3073               | Birth, 2, 4, ..., 24                                     | 32636               | Adair 2007 Am J Hum Biol                   |

|                                                                                                                                                                             |              |             |                    |       |                           |        |                                       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-------------|--------------------|-------|---------------------------|--------|---------------------------------------|
| Africa                                                                                                                                                                      |              |             |                    |       |                           |        |                                       |
| MAL-ED                                                                                                                                                                      | Tanzania     | 2009 - 2014 | Prospective cohort | 262   | Birth, 1, 2, ..., 24      | 5857   | Mduma et al 2014 Clin Infect Dis TZN  |
| Tanzania Child 2                                                                                                                                                            | Tanzania     | 2007 - 2011 | Individual RCT     | 2400  | 1, 2, ..., 20             | 32198  | Locks et al Am J Clin Nutr 2016       |
| MAL-ED                                                                                                                                                                      | South Africa | 2009 - 2014 | Prospective cohort | 314   | Birth, 1, 2, ..., 24      | 6478   | Bessong et al 2014 Clin Infect Dis SA |
| MRC Keneba                                                                                                                                                                  | Gambia       | 1987 - 1997 | Cohort             | 2931  | Birth, 1, 2, ..., 24      | 40952  | Schoenbuchner et al. 2019, AJCN       |
| ZVITAMBO Trial                                                                                                                                                              | Zimbabwe     | 1997 - 2001 | Individual RCT     | 14104 | Birth, 6 wks, 3, 6, 9, 12 | 73651  | Malaba et al 2005 Am J Clin Nutr      |
| Lungwena Child Nutrition RCT                                                                                                                                                | Malawi       | 2011 - 2014 | Individual RCT     | 840   | Birth, 1-6 wks, 6, 12 18  | 4346   | Mangani et al. 2015, Mat Child Nutr   |
| iLiNS-Zinc Study                                                                                                                                                            | Burkina Faso | 2010 - 2012 | Cluster RCT        | 3266  | 9, 12, 15, 18             | 10552  | Hess et al 2015 PlosOne               |
| Latin America                                                                                                                                                               |              |             |                    |       |                           |        |                                       |
| MAL-ED                                                                                                                                                                      | Peru         | 2009 - 2014 | Prospective cohort | 303   | Birth, 1, 2, ..., 24      | 6442   | Yori et al 2014 Clin Infect Dis PER   |
| CONTENT                                                                                                                                                                     | Peru         | 2007 - 2011 | Prospective cohort | 215   | Birth, 1, 2, ..., 24      | 8339   | Jaganath et al 2014 Helicobacter      |
| Bovine Serum RCT                                                                                                                                                            | Guatemala    | 1997 - 1998 | Individual RCT     | 315   | Baseline, 1, 2, ..., 8    | 2551   | Begin et al. 2008, EJCNC              |
| MAL-ED                                                                                                                                                                      | Brazil       | 2010 - 2014 | Prospective cohort | 233   | Birth, 1, 2, ..., 24      | 5092   | Lima et al 2014 Clin Infect Dis       |
| INCAP Nutrition Supplementation Trial Longitudinal Study                                                                                                                    | Guatemala    | 1969 - 1977 | Cluster RCT        | 1406  | Birth, 3, 6, ..., 24      | 7965   | Habicht et al 1995 J Nutr             |
| Europe                                                                                                                                                                      |              |             |                    |       |                           |        |                                       |
| PROBIT Study                                                                                                                                                                | Belarus      | 1996 - 1997 | Cluster RCT        | 16898 | 1, 2, 3, 6, 9, 12         | 124509 | Kramer et al 2001 JAMA                |
| Mortality analysis only                                                                                                                                                     |              |             |                    |       |                           |        |                                       |
| Burkina Faso Zinc trial                                                                                                                                                     | Burkina Faso | 2010-2011   | Cluster RCT        | 7167  | 6, 10, 14, 17, 22         | 15155  | Becquey et al 2016 J Nutr             |
| Vitamin A Trial                                                                                                                                                             | India        | 1995-1996   | Cluster RCT        | 3983  | 1, 3, 6, 9, 12            | 32570  | WHO CHD Vitamin A Group 1998 Lancet   |
| iLiNS-DOSE                                                                                                                                                                  | Malawi       | 2009-2011   | Individual RCT     | 1932  | 6, 9, 12, 18              | 13801  | Maleta et al. 2015 J Nutr             |
| iLiNS-DYAD-M                                                                                                                                                                | Malawi       | 2011-2015   | Individual RCT     | 1235  | 1, 6, 12, 18              | 9207   | Ashorn et al 2015 J. Nutr             |
| *Children enrolled is for children with measurements under 2 years of age. Total measurements are number of measurements of anthropometry on children under 2 years of age. |              |             |                    |       |                           |        |                                       |

## Appendix D: Research and Methodological Lessons

There are several lessons from this dissertation beyond the subject-area findings on child growth failure, which can be considered in planning future analyses of early life growth failure, IPD meta-analyses, and in designing studies with plans to share the data to future IPD meta-analyses. This appendix discusses process and methodological lessons I learned through the process of conducting this dissertation research

### **1) Replicate the data mapping of raw data to harmonized individual participant datasets, and to check the underlying accuracy of the raw data.**

An underdiscussed issue in IPD analyses is cleaning and harmonizing the contributed datasets,<sup>225</sup> although the PRISM-IPD reporting guidelines for IPD studies recommend reporting data standardizing and data checking steps.<sup>226</sup> Data harmonization is the process of coding variables from different studies to standard scales, mapping equivalent or close to equivalent variables to the same variable names, and merging all datasets together.<sup>227</sup> Data harmonization is less challenging in the IPD analyses of intervention effects, especially with a two-stage design, because analysis of well-randomized trials do not require adjustment for confounders, so only the outcome needs to be standardized. But for IPD analyses of observational studies or non-randomized exposures measured in trials, like the analyses in chapters 2-4 of this dissertation, the data harmonization process is far more challenging. Both outcome variables and key covariates need to be standardized, and even among studies conducted on the same scientific questions, different adjustment covariates are measured, and in different ways, especially if the covariates are geographically or culturally specific like dietary intake or asset ownership.

There are two key issues in the data harmonization process: 1) quality control (QC) checking the raw contributed datasets and the coding of the mapping to a standard format, and 2) correctly harmonizing and combining variables or choosing when variables are too different to harmonize. Although there are standardized methodologies for coding of certain measures, like body-mass index, other variables are measured across a range of scales, like household food security, or are context dependent, like educational attainment. Therefore, the IPD team needs both data scientists to map raw data to a standard form and subject area-experts who understand the underlying studies. Close communication with the authors of the individual contributed studies is also critical to answer questions on the underlying data, especially if the descriptions of key variables in the metadata are meager.

In harmonizing the data contributed to the KI effort, I undertook an extensive quality control (QC) process for three reasons:

**a) To check the accuracy of the underlying contributed datasets.** Each underlying dataset should be checked to ensure it matches the description of the published study and that the values of the key variables are reasonable. Although every effort should be made to communicate with the original study authors to clarify discrepancies, understand variables, and ensure the correct data were shared, original study investigators may not have the time, inclination, or knowledge of the measurement quality of non-primary outcomes to assist with questions. In that case, it will be the IPD meta-analysis investigators' decision whether to exclude a variable or a study from

the analysis. Any exclusion based on the QC check needs to be documented and justified in supplementary materials. In the KI analysis, MAL-ED Pakistan was excluded from the analysis for inexplicable binning of the LAZ scores around -2, -1, and 0 (Figure 5.1a). Although the WAZ and WLZ scores had a normal and believable distribution the unrealistic LAZ distribution raised questions on the overall anthropometry measurement quality or data entry process in the study, so the cohort was excluded. I also checked primary results from published papers against results estimated from the individual-participant dataset, which, along with checking that I was using the correctly cleaned data, had the added benefit of being a de-facto external replication of the contributed studies.

**b) To check for errors in coding of the data mapping from raw contributed datasets to the cleaned, harmonized single datasets.** Unless an IPD analysis is done on data from an existing public repository, all shared datasets are likely to be in different formats, with different variables names and data structure. Therefore, every study needs unique coding to map the data to a standardized IPD format. This multiplies the possibility of data mapping errors by the number of studies in the IPD analysis. Raw contributed data can be in either long-form, where every row is an observation, or wide form, where every row is a child, and there is a different column for every timepoint a variable is measured. In large datasets, variables names can often be non-descriptive alpha-numeric codes which need to be carefully renamed based on the meta-data. Sometimes, a single raw dataset is shared, but sometimes multiple datasets are shared, representing different variable domains or different follow-up measurements, and each needs to be carefully cleaned and merged. Missing data are coded in different ways, sometimes as a blank or missing value depending on the type of file used for data storage, but sometimes codes such as 99 or 99.9 are used to indicate missingness. Depending on the distribution of a variable, these missing codes could appear to be reasonable data values, or just not noticed during data harmonization, but as they are numeric codes they will be included in analyses if not recoded, with potentially highly influential effects. Although no major errors were found in the raw datasets that led to study exclusions, other than the cohort detailed above, several major data mapping errors were uncovered by looking at histograms and scatter plots of anthropometry measurements, including erroneous imputation of birthweight and birth length (Figure 5.1b). Unlike errors in the raw data, once identified, errors in the data mapping process are easy to fix.

**c) To assess the relative measurement quality across the different contributed studies.**

Beyond finding discrete errors in coding that need to be fixed or clear errors in the raw data that lead to decisions on excluding an observation, variable, or study, the QC checking process can be used to assess the relative measurement quality of each study. Although there are many scales developed for AD meta-analyses that assess the strength of evidence from each study based on study design and size,<sup>228</sup> IPD give a unique opportunity to assess quality directly from the data. This is not possible for many variables, but for some standardized measurements, like clinical endpoints, there is a known amount of variability expected, or only a certain range of values are possible. In the case of child anthropometry, I focused on repeated measures of height in children. Although child weight can decrease over time, children cannot lose height. Therefore, if a child's height is lower at an older age than a younger age beyond the expected technical error of measurement (TEM), one or both of the measurements must be inaccurate.<sup>229</sup> For each cohort, I calculated the proportion of measurements where height decreased beyond the TEM (Figure 5.2c). No cohort rose to the level of concern that I considered excluding it for poor measurement quality, but this QC analysis allowed me to compare and report a relative measure of

anthropometry data quality that I could report so readers could assess the input data and see due diligence was done to avoid “garbage in, garbage out” research where inaccurate underlying data leads to incorrect results, no matter the analytical sophistication.

Beyond QC checking the data quality, most of the work to harmonize datasets focused on standardizing variables across studies so they can be combined into a single dataset. This process is trivial for some variables. Maternal weight, as a continuous variable with standard units, is easy to combine after standardizing variable names and converting the units to kilograms in each study. Others could be easily converted to the same standard scale: all anthropometry measurements were converted to Z-scores using the 2006 WHO standards so that all children’s heights and weights were compared to the same reference population. But other variables require more effort or more assumptions to standardize across studies, and variables either need to be transformed, coarsened, or aggregated to create harmonized variables with comparable meanings across studies.

It may not be possible to pre-register all the decisions made in harmonizing variables until the contributed data are received and explored, but all effort should be made to pre-register the planned coding of the cleaned variables, or use a standard instrument, indicator, or score for each variable. For example, in the raw datasets acquired through the KI effort, survey questions on household drinking water and latrine usage had disparate, and sometimes only locally relevant, response, with the number of responses varying across studies from a few to a few dozen. To harmonize the variable coding and to allow generalizability beyond the KI IPD meta-analysis, I coded indicators for access to clean drinking water and improved sanitation from the survey responses based on the widely used WHO/UNICEF joint monitoring program definitions.<sup>230</sup> For some variables, there may be multiple instruments in use and assumptions may be required to harmonize. KI studies that measured household food insecurity used three standard scales: The Household Hunger Scale (HHS),<sup>222</sup> Food Access Survey Tool (FAST),<sup>223</sup> USAID Household Food Insecurity Access Scale (HFIAS).<sup>224</sup> The first two scales were three-level, while the third scale is four level, so the middle 2 categories (mild and moderately food insecure) were assumed to represent levels of food insecurity that match the middle level of the other scales, and a hybrid three-level variable was coded from the three scales.

Other variables are culturally dependent, and so will not match across geographically diverse studies, but can be coded to be comparable within studies, and then within-study comparisons can be pooled. For example, although maternal body-mass index (BMI) is an absolute biological measure, years of maternal education is a relative socio-economic measure whose meaning may vary depending on local educational infrastructure and norms. It may therefore be reasonable to directly compare a low BMI mother in Bangladesh to a high BMI mother in Peru, but a mother’s educational attainment in Bangladesh is not necessarily comparable to a mother’s educational attainment in Peru in assessing maternal knowledge or family socio-economic status. But it is reasonable to categorize educational attainment as low, medium, or high depending on the distribution of years of schooling within a study, and then estimate the relative risks of outcomes of interest (such as child wasting or stunting), and then pooling the relative risks across cohorts, as done in Chapter 3 of this dissertation. Other modeling techniques can be used to harmonize domains of variables that either vary in meaning across studies or have different sets of variables measured across studies. Household income is not a reliable indicator of wealth in low-resource settings as it often fluctuates greatly over time

or season.<sup>231</sup> Many KI studies included indicators of asset ownership, including car, mobile phone, or bicycle ownership, but also culturally specific items like khat or chouki (types of traditional beds in Bangladesh) ownership. To compare across studies, I ran a principal components analysis across all assets measured within each study, and used quartiles of the first principal component as a within-study relative measure of wealth that can be pooled across studies.<sup>232</sup> This harmonization process for culturally-specific variables reduces their interpretability but allows them to be used as more generalizable prognostic factors when pooled across studies.

A key limitation of the data harmonization process is the risk of bias towards studying or adjusting for variables that are easily harmonized. Chapter 3 of this dissertation focuses on parental and at-birth anthropometry as strong prognostic factors for child growth failure, partly because of the strength of association and partly because they are commonly measured in most child growth cohorts. Although parental and at-birth anthropometry are strong predictors of child growth failure, poor breastfeeding practices and early-life nutrition are clear causes of child malnutrition, but indicators of breastfeeding practice were not strongly associated with child growth failure, and measures of child nutrition or complementary feeding were not included.<sup>1</sup> This may partially be due to greater measurement error and recall bias in self-reported breastfeeding practices than in weight measurements, or because I coarsened detailed longitudinal information on breastfeeding practice into indicators of exclusive or predominant breastfeeding (where a child can only consume breastmilk or other liquids) in the first 6 months of life. A few KI studies included in the analyses presented in Chapter 3 had detailed measurements of child micronutrient intake and complementary feeding, but different studies had different frequency of dietary and breastfeeding surveys and measured different sets of culturally relevant early-child foods. IPD analyses of coarsened variables may need to be supplemented with separate analyses, tailored to the structure of the raw data, within studies that have more detail in or frequency of measurement. Although outside of the IPD meta-analyses, this approach may supplement the precise estimates of the IPD meta-analyses with more nuanced evidence of how the exposure affects the outcome.

In conclusion, errors in the data harmonization effort were found towards the end of the analysis and dissertation writing, and only after the results were shared with the primary investigators of the individual's studies included in the analysis, one of whom noticed missing data at birth. The data were from an unpublished study, so no comparison with published estimates was possible. Possibly, additional metadata-sharing, including detailed analysis plans and study surveys, would have helped catch the error, but replication of the coding of the data mapping and harmonization by two independent data scientists would have likely caught the error. Replicating the data mapping is expensive and time consuming and is a decision that needs to occur at the onset of an IPD study rather than after finding errors, but it is worth the cost, especially if the individual-participant data will be made public.

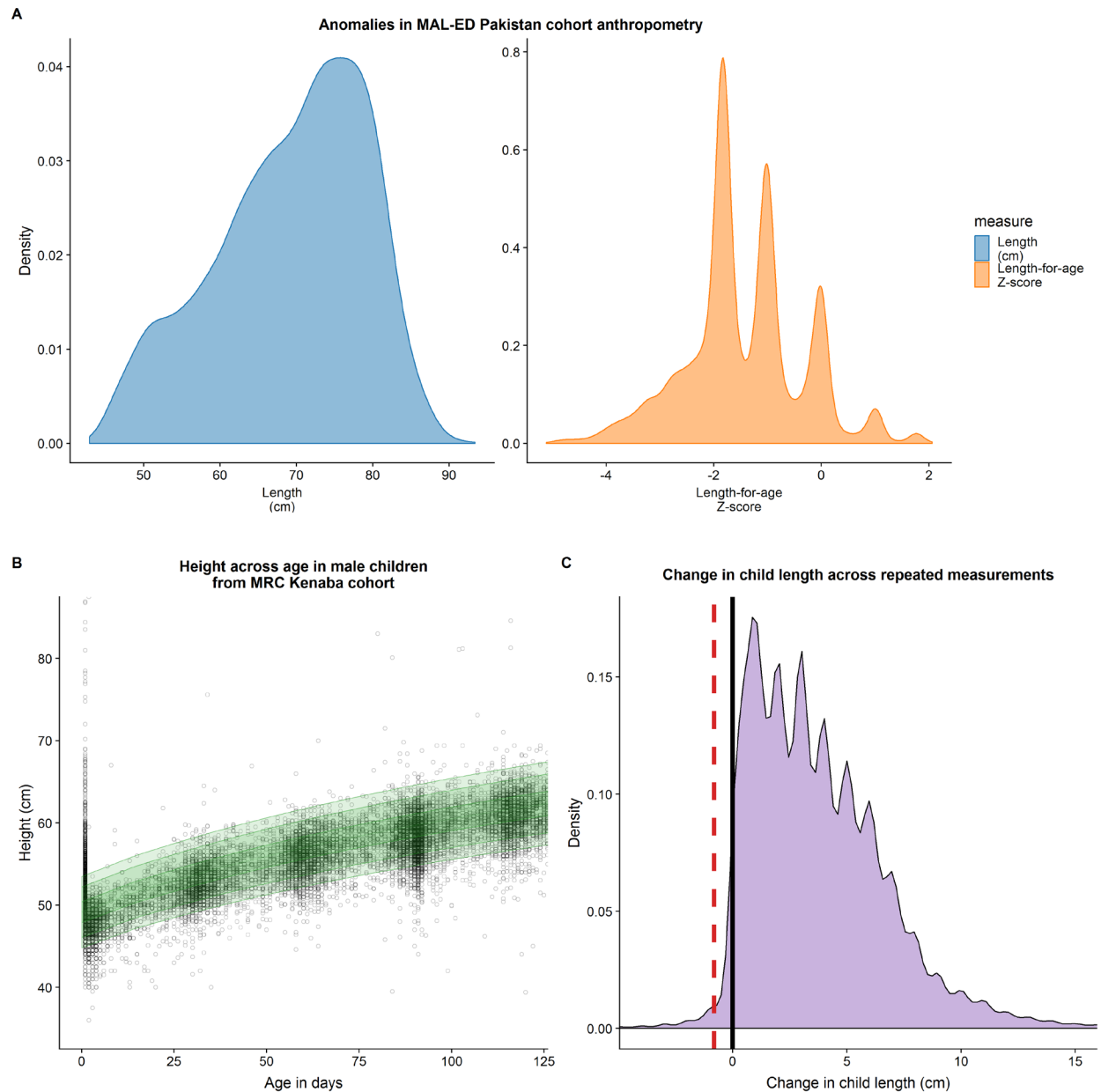


**Figure D.1 | Anthropometry measurement quality checking of KI individual-participant datasets**

**(A)** In the MAL-ED Pakistan cohort, children under two years old had lengths with a normal distribution, but when lengths were converted into length-for-age Z-scores, there was anomalous binning in the distribution of Z-scores where most Z-scores were whole numbers. There was no error found in the mapping of the contributed dataset, indicating that there were erroneous lengths or ages recorded in the original study data collection effort. The dataset was excluded from the analysis.

**(B)** Improbable lengths of children at birth in the MRC Kenaba cohort can be seen in the plot of child lengths by age. Darker to lighter green bands mark 1, 2, and 3 standard deviations from reference growth standards. After conversions to Z-scores, most were

above the 5 standard deviation outlier cutoff, and so would have been excluded from the analysis and the error would not have been caught. Upon examining the distribution of child lengths over age, the error was noticed. The data cleaning of the contributed IPD was fixed to correctly map child anthropometry at birth, and the cohort was retained in the analysis.

**(C)** Checking measurement quality of child lengths by examining length that decreased over longitudinal measurements. Although child weight can fluctuate over measurements, child length can only increase, though there is expected error in child length measurement. The dashed red line marks the technical error of measurement. To the left of the dashed lines, child length decreased across measurement greater than can be explained through expected measurement error alone.

## **2) Data coarsening and assumptions are needed when datasets aren't from studies using standardized measures**

Although there are often reasons to measure a variable using an instrument or survey question unique to a particular study, standardizing variable measurements used in primary data collection will make future IPD meta-analyses more powerful. My collaborators and I needed to make assumptions to combine some variables, like assuming categories of different household food security instruments are equivalent, and coarsen other variables into summary indicators from more nuanced measurements. For example, in Chapter 3 I analyzed an indicator of whether a child was exclusively breastfed in the first six months of life and the percent days a child had diarrheal disease in the first six months of life. Both of these variables were coded from multiple questions from repeated surveys, with different survey designs across studies. Specific timing of diarrheal disease in early childhood, or the specifics of food consumption when children are not exclusively breastfed, may both be important in causing the onset of child wasting. I coarsened these longitudinal experiences into summary variables, neither of which were strongly associated with the risk of wasting, but I may have removed important information on early child infection and nutritional experiences. Answering subtle questions on the timing of illness or feeding practices may not be possible in an IPD meta-analysis, unless combining studies with coordinated timings of observation and standardized measurements. Potentially, IPD meta-analyses should be supplemented by study-specific analyses of studies with greater measurement detail, and the coarsened IPD results qualitatively compared to the more nuanced study-specific estimates.

## **3) Increase social norms around data sharing for IPD analyses and standardize data collection, coding, and storage methods**

The advantages of IPD meta-analyses are clear but overcoming the key barrier to broader adoption of IPD meta-analyses, the access to individual-participant data, is more challenging. Although there is room for improvement of IPD meta-analysis methods, including additional research on one-step targeted learning models, and applications of marginal structural models using machine learning to estimate non-linear exposure-response relationships or time-to-events, the most impactful path to increasing the frequency and quality of an individual-participant data analysis would be to make it easier to access or create harmonized datasets. A review of IPD-

meta-analyses found that 40% of studies were not able to get IPD from all identified studies,<sup>18</sup> and a second review found a lower rate of 25%, with only 43% of IPD meta-analyses able to retrieve more than 80% of eligible data.<sup>233</sup> Additionally, rate of data retrieval success did not appear to increase over time,<sup>233</sup> even as there are increasing calls for data sharing<sup>16,234</sup> and increasing number of platforms to make data-sharing easier.<sup>235</sup> A trial found that financial incentives did not increase data sharing, so issues other than lack of compensation prevent sharing.<sup>236</sup> Primary investigators may not choose to share data due to lack of time to clean and explain data, out of concern for patient confidentiality, because of fears around scrutiny of their research, or out of desire to publish more on the dataset and worries that sharing data will lead to other teams publishing first.

To increase the rate of data sharing, research norms, investigator incentives, and funding and publication requirements need to change. Incentive structures in science need to change to make researchers feel obligated and rewarded to contribute their data to IPD meta-analyses, and to work with the investigator to understand and accurately harmonize the contributed data. The US Institute of Medicine issued a report on sharing clinical trial data, calling for a change in scientific culture to one where “data sharing is the expected norm.”<sup>237</sup> Following this, the British Medical Journal began requiring data sharing upon request for all trials published in the journal.<sup>238</sup>

Not only should data be shared with IPD meta-analysis efforts, the IPD should be made public after publications of the meta-analysis results, if possible. IPD meta-analyses require a large effort for both data contributors and the IPD investigators, and for meta-analyses of treatment effects IPD analysis may not find conclusions different from an AD analysis, but the effort has a long-term payoff if the dataset becomes a public good. There should also be career rewards for the creation of a large public data; scientific culture should shift to view cleaning, publishing, and hosting IPD data as an equivalent contribution to science as a high-profile publication. Institutions or teams who have worked on an IPD have uniquely detailed knowledge of study data, and the dataset can become a public good as a teaching resources or basis for graduate student projects. Especially in global health research, making data from IPD meta-analyses public is also a first step in increasing equity between high-income country researchers and low-income country researchers, as open data removes the institutional and networking advantages high-income country researchers have in applying for grants and in requesting data access.<sup>239,240</sup> Also, if privacy is protected, data sharing is ethical for participants as it maximized the scientific good of participants time and potential discomfort. Grant and publication stipulations that data be made public within a reasonable timeframe are a clear policy change to increase the amount of data sharing, but shifting norms could increase useful participation by the original study investigators in explaining study and data details. The peer-review process that underpins scientific publication is not mandatory, doesn’t involve financial reward, and doesn’t directly lead to career advancement, yet it is an ingrained aspect of the scientific process, so data sharing and IPD study coordination could follow a similar template.

If data sharing becomes the cultural norm, as well as a requirement for many funders and journals, then the planning for data sharing could begin at the grant proposal stage. Early planning for sharing would ensure sharable study documentation, proper consents for future data sharing, and early addressing of patient confidentiality issues. If data-sharing is planned in the study design phase, then standardized instruments, indicators, and scales can be chosen to

measure variables in order to increase study comparability and combinability, unless there is a compelling reason to measure or code a variable in a custom manner to meet study goals. Also, mirroring PRISMA, CONSORT, and STROBE reporting guidelines for meta-analyses, trials, and observational studies, each field should develop standardized guidelines for formatting, storing, and naming variables and for structuring electronically-searchable metadata so that studies from the same field can be more easily combined. And when data are shared, all possible study documentation, including analysis plans, surveys, protocols, and primary analysis code, should be shared with the IPD investigator to maximize transparency and understanding of the contributed data.

Beyond steps the primary investigators of original studies can take to improve IPD meta-analyses, the IPD investigators can improve data sharing through steps of their own. If potential data contributors are worried that other teams will complete and publish planned secondary analyses before the original investigators, then allow a reasonable embargoed time before making the dataset used in an IPD meta-analysis public. Also, report in the primary IPD publication which teams did not contribute data, and why, to allow readers to assess potential “data contribution” bias, but also as a reputational reward for data contribution. If IPD is going to be made public after publication of the meta-analysis results, then it is critical that the data mapping is accurate, and the coding of data mapping from the raw data to the harmonized IPD should be replicated by two independent data scientists. The measurement quality and accuracy of the original study data should also be checked, and original study findings replicated with the shared data. With the ongoing reproducibility crisis in science,<sup>241</sup> IPD meta-analyses also provide a unique opportunity to externally replicate published findings. I recommend that the online versions of publications of key trials and studies be updated with a note after they are externally replicated through an IPD meta-analysis. This both accurately adds evidence to the robustness of a trial and also gives trial researchers an incentive to share data as they can boost the prestige/strength of evidence of their work. If the primary research goal of the IPD meta-analysis is on a different scientific question, replicating the primary results of underlying datasets may seem extraneous, but on top of being a scientific contribution, it is a form of QC checking of the IPD harmonization process and can be a robustness check on differing variable cleaning or exclusion criteria.

#### **4) Investigate rare outcomes and subgroups in IPD meta-analyses.**

Although IPD meta-analyses can focus on better estimation of the primary outcomes of the underlying datasets, possibly the largest contribution of this dissertation is estimates of outcomes too rare to examine within individual studies. Additionally, unlike a traditional meta-analysis of intervention effects, access to individual participant data allowed me to estimate treatment effects within a broad range of subgroups standardized across trials. Although the estimates of wasting prevalence and associations between risk factors and wasting and stunting reinforce prior findings, but in a larger dataset, the large, longitudinal dataset allowed me to examine patterns in wasting within narrow slices in time, and find a higher early onset of wasting than previously understood. Additionally, concurrent stunting and wasting is associated with a much higher risk of mortality but is too rare to examine in most existing studies, and the WHO does not estimate the prevalence of concurrent wasting and stunting due to these data limitations. I was able to both describe patterns in and consequences of concurrent wasting and stunting over time, which peaks among children 1.5 years old, and is strongly associated with later mortality.

## **5) Two-step IPD meta-analyses allowed for greater flexibility in subgroup analyses and more precise estimates when using targeted learning.**

There are two primary approaches to analyze IPD: the two-step and one-step approaches.<sup>17,242</sup> In a two-step approach, which was taken in chapter 2 and 3 of this dissertation as is most common in the published literature, the IPD analyses mirrors an AD analysis.<sup>15</sup> Summary estimates of interest are calculated separately for each study and then pooled using traditional meta-analysis methods. In the KI analyses, among other parameters, I estimated mean Z-scores, mean differences in Z-scores, prevalences of growth failure, and relative risks of growth failure within each cohort, and then pooled across cohorts using random-effects meta-analysis models. In a one-step approach, which is growing in popularity,<sup>15</sup> only a single model, applied to the full set of aggregated IPD, is used to estimate the parameter of interest. Crucially, in the one-step approach the study origin of all participants must be accounted for in the model unless the participants of all studies come from the same population, a frequently unrealistic assumption. The one-step approach can mirror either random or fixed effects AD meta-analysis models, depending on if the study origin of each participant is modeled using a multilevel or mixed effect model, or as fixed indicator terms. In practice, both approaches give very similar estimates, especially when the treatment of interest is a treatment effect from randomized control trials,<sup>243–245</sup> but the estimates from these two approaches can meaningfully vary when confounding is being adjusted for.<sup>242</sup>

There are a few other situations where the one- and two-step approaches differ in utility. Random effects are easy to implement in a two-step design, regardless of design, which is not true for time-to-event models in a one-step design.<sup>246</sup> Two-step designs also allow the investigator to harness the broad set of visualizations and sensitivity analyses developed for AD meta-analyses. Once study-specific parameters of interest are estimated, forest plots can be created to visualize study-specific estimates and potential outliers, and funnel plots can be created to visualize potential publication bias. It is also easy to calculate standard heterogeneity statistics, or to exclude a study from the pooling step as a sensitivity analysis. The two-step approach also serves as a de-facto external replication of the estimates from contributed studies, while the one-step approach obscures the relative contributions of each study to the pooled estimate, depending on the complexity of the model. Lastly, the two-step approach allows AD estimates to be included into the second step if there are key studies where IPD cannot be accessed but that are important to include in estimating pooled effects.<sup>243,247</sup> The pooled estimates from the AD estimates can also be compared to the IPD estimates to assess generalizability of the contributed studies.

The one-step design allows for flexible modeling where both individual-level and study-level variables are to be included,<sup>243</sup> though study-level variables can be accounted for in a two-step approach through using meta-regressions.<sup>248</sup> Also, if exposures of interest are continuous, a one-step analysis will give more power to estimate non-linear trends.<sup>249</sup> A one-step design may be preferable depending on the size of data and type of outcome being analyzed; the two-step design may lead to bias in estimated pooled effects, standard errors, and between-study

heterogeneity if there are few studies, small studies, or rare outcomes, or when the statistical models used cannot fully account for follow-up times.<sup>250</sup>

I chose to use a two-step approach in chapters 2-4 of this dissertation because I had an interest in comparing pooled estimates between different countries and regions, as well as calculating an overall pooled estimate. In a one-step approach, different models would need to be fitted on the subject-level data to calculate global and regional estimates, which is more computationally expensive than fitting separate meta-analysis models on the summary estimates. Although one-step and two-step designs often lead to similar inferences, the IPD design should be pre-specified, as there can be differences in certain situations. If one approach is not clearly preferable to the other, the alternative design should be conducted and reported as a sensitivity analysis.<sup>251</sup>

Existing literature is on IPD analysis using parametric regression models, but the KI analysis was the first application of targeted learning to IPD meta-analyses. Targeted learning is a semi-parametric doubly-robust estimation method that targets a parameter of interest (for example, a relative risk between exclusive breastfeeding and stunting) and optimizes the bias-variance trade-off for the parameter of interest rather than for the model as a whole.<sup>60</sup> Targeted learning can utilize machine learning methods to estimate the parameter of interest, while still retaining valid inference, allowing for flexible adjustment for potential confounding and reducing the risk of model misspecification.<sup>60,133</sup> I used ensemble machine learning, which utilizes the optimal combination of multiple user-specified algorithms, and I specified a library of simple means, main effects generalized linear models, LASSO penalized regressions,<sup>134</sup> generalized additive models,<sup>135</sup> and gradient boosting machines.<sup>136</sup>

In IPD meta-analyses of randomized control trials, the targeted learning approach may not be necessary if trials are well randomized and there is no imbalance in subject characteristics across treatment arms, but in IPD meta-analyses of observational studies, targeted learning will better control for confounding, along with removing the model-building that often occurs with parametric regression models (Coyle et al., 2020). This is especially important in IPD meta-analyses, because different studies will measure different sets of potential confounders, as in the KI data presented in Chapters 3-4, and because instead of deciding on the covariates to include and how to model them within the regression framework for a single study, every study needs to be modeled, and so the automated and best-fit modeling of non-linearity and interactions that is inherent in cross-validated ensemble machine learning can estimate the exposure-response relationship more accurately and free of model selection bias.<sup>252</sup>

In chapters 3-4 of this dissertation, my collaborators and I chose to use a two-step approach, where parameters were estimated with separate targeted learning models for each study, and then pooled using traditional meta-analysis models, so that cohort-specific estimates could be compared, and pooled estimates could be quickly estimated within regional subgroups. But the flexibility of SuperLearner ensemble machine learning, combined with the cross-validated implementation to avoid over-fitting, may allow for estimation of pooled parameters with greater efficiency when using a one-step approach.<sup>249</sup>

I compared estimates of population attributable differences of different exposures on length for age Z-scores (LAZ) and weight-for-length Z-score (WLZ) in 24-month-old children, estimated with targeted learning models using either a two-step or a one-step approach (Figure

5.1). I estimated both using fixed-effect models, where in the two-step approach, parameters are pooled using inverse-variance weighting, without random effects for study, and in the one-step approach study is included as indicator terms in the SuperLearner model. For both approaches, missingness of factors were coded as a missing category, while missingness in continuous variables were imputed with study-specific medians, and an indicator term marking missingness was created. Although this method can induce bias in generalized linear models, it is a bias-free solution in machine learning procedures,<sup>253–255</sup> though multiple imputation is recommended for IPD analysis using parametric models.<sup>256</sup> Rather than examine differences in one- and two-step estimation approaches using targeted learning with more commonly estimated relative risks or Z-score differences, I compared estimation of population attributable differences, the expected difference in population mean Z-score if all children had the reference level of an exposure rather than the observed distribution, because it allows rank ordering of a single numeric summary of exposure's association with child growth.

I found that two-step estimation was preferable to the one-step approach in this finite sample example. Although estimates were broadly similar, there were a few large differences in attributable differences estimated through the two approaches (Figure 5.2a). The two-step estimates, however, were more precise on average, with an average SE difference of 0.2 (Figure 5.2b). Additionally, when estimates differed substantially, the one-step estimates were more extreme in most cases (Figure 5.2c), which may be finite sample bias from sparse data.<sup>257</sup> Future research should confirm these findings through simulation studies, and extend the work to other parameters, but this analysis indicates that the two-step approach is appealing in finite sample, even when the samples are large as in this analysis of 31 cohorts (see Chapter 3). This is on top of the appealing ease in which pooled estimates can be estimated in subgroups without re-running computationally expensive machine learning models, so I recommend the two-step approach in future IPD meta-analyses that use targeted learning for better confounder adjustment.

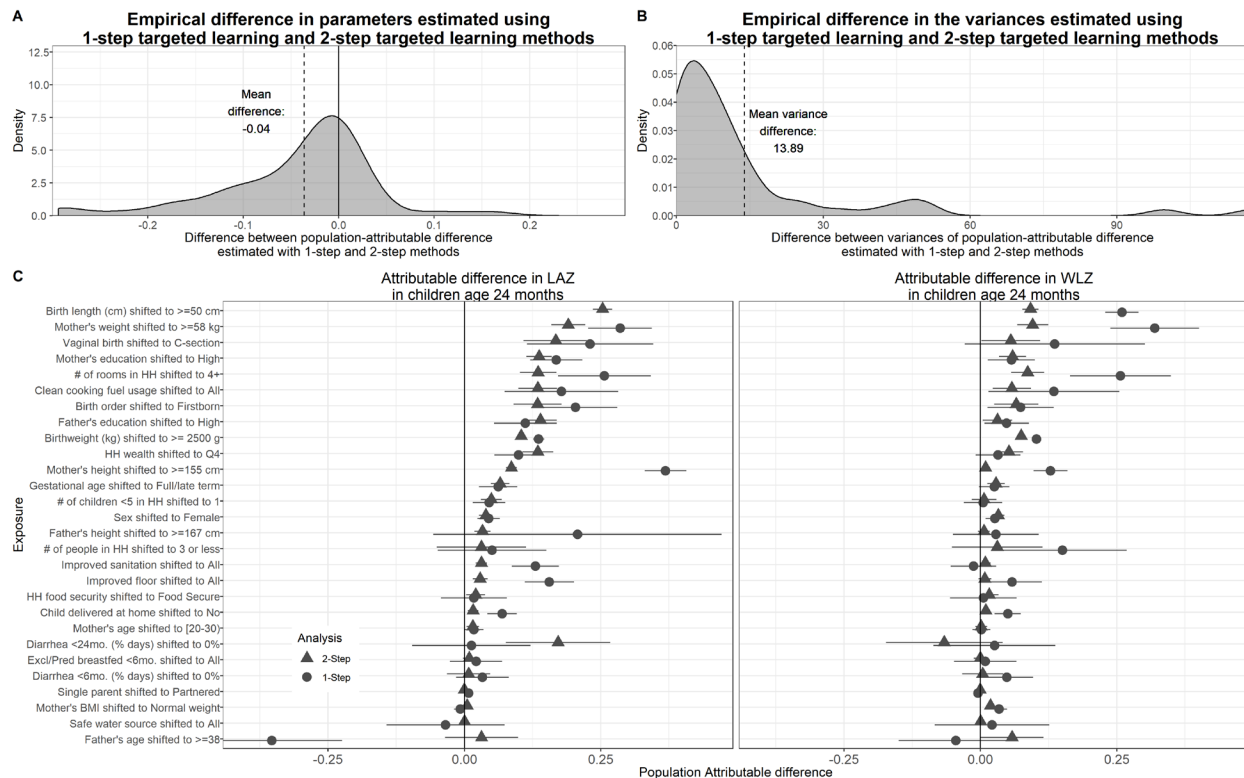


**Figure D.2 | Comparing population attributable differences estimated using targeted learning using either a one-step or two-step approach.**

(A) Density plot of differences in population attributable differences estimated via one- or two-step analyses for all exposures examined in Chapter 3 of this dissertation, using the outcomes of LAZ and WLZ at 24-months of age. Estimates are pooled using fixed-effects models in the second stages of the two-step analyses, cohort membership is modeled as a fixed-effect in the one-step analyses. The population attributable difference is the expected difference in population mean Z-score if all children had the reference level of the exposure rather than the observed distribution. There was broad agreement in the estimates of population-attributable differences estimated by the one-step or two-step approach, though there are a few outlier estimates where the approaches differed substantially. On average, the two-step estimates were 0.4 larger than the one-step estimates. In the two-step approach, estimates were pooled across studies using fixed effects models, and the one-step estimates were estimated using targeted learning models with indicators for study membership included.

(B) Differences in the variances of population-attributable differences estimated either through a one-step or two-step approach show that the two-step method had more precise estimates. On average, the variances of the two-step estimates were 13.89 smaller than the one-step estimates.

(C) Comparison of exposure-specific population attributable differences estimated with one-step or two-step models for the outcomes of LAZ and WLZ at 24 months of child age. Triangle points represent the 2-step estimate, and the circle points directly below the triangles represent the equivalent one-step estimate, with the exposure listed on the Y-axis along with the reference level the population is shifted to in the estimation of the attributable risk. In general, when the estimates diverged substantially, the one-step estimates were more extreme while two-step estimates were closer to the null.