FROM THE AMERICAN PEOPLE

DHS WORKING PAPERS

Alternative and Complementary Metrics of Linear Growth for Tracking Global Progress in Child Nutritional Status

Ashley Aimone
Diego G. Bassani
Huma Qamar
Nandita Perumal
Sorrel M.L. Namaste
Daniel E. Roth

2021 No. 153

September 2021

This document was produced for review by the United States Agency for International Development.

DEMOGRAPHIC AND HEALTH SURVEYS

DHS Working Papers No. 153

Alternative and Complementary Metrics of Linear Growth for Tracking Global Progress in Child Nutritional Status

Ashley Aimone[1]
Diego G. Bassani[1]
Huma Qamar[1]
Nandita Perumal[1]
Sorrel M.L. Namaste[2]
Daniel E. Roth[1]

The DHS Program
ICF
Rockville, Maryland, USA

September 2021

[1] The Hospital for Sick Children, Centre for Global Child Health, Toronto, Ontario, Canada
[2] The DHS Program, ICF, Rockville, Maryland, USA

Corresponding author: Ashley Aimone, Centre for Global Health Research, Toronto, Ontario, Canada, aaimone@gmail.com

Acknowledgements:
The authors would like to acknowledge the Growth Metrics Project Advisory Group (in alphabetical order): Edward Frongillo (University of South Carolina), Jef Leroy (International Food and Policy Research Institute), Erin Milner (United States Agency for International Development), Kuntal Saha (World Health Organization), and Julia Krasevec (United Nations Children's Fund). The Advisory Group provided guidance on the initial research protocol and interpretation of some preliminary results, but was not involved in the review or approval of the final results or the manuscript. The authors would also like to thank Rukundo Benedict (The DHS Program, ICF) for her input during the conceptualization of the project.

Editor: Kerry Aradhya
Document Production: Chris Gramer

This study was implemented with support provided by the United States Agency for International Development (USAID) through The DHS Program (#720-OAA-18C-00083). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. For additional information about The DHS Program, contact: The DHS Program, ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: 301-572-0950; fax: 301-572-0999 / 301-407-6501; email: reports@dhsprogram.com; internet: www.dhsprogram.com.

Recommended citation:
Aimone, Ashley, Diego G. Bassani, Huma Qamar, Nandita Perumal, Sorrel M.L. Namaste, and Daniel E. Roth. 2021. *Alternative and Complementary Metrics of Linear Growth for Tracking Global Progress in Child Nutritional Status.* DHS Working Papers No. 153. Rockville, Maryland, USA: ICF

CONTENTS

TABLES

FIGURES

ACRONYMS

DHS	Demographic and Health Surveys
GDP	Gross domestic product
HAD	Height-for-age difference
HAZ	Height-for-age z-score
IGME	Inter-agency Group for Child Mortality Estimation
IQR	Interquartile range
LMICs	Low- and middle-income countries
LR	Less restrictive
MOB	Month of birth
MoL	Measure of location
PCA	Principal component analysis
SD	Standard deviation
SMART	Standardized Monitoring and Assessment of Relief and Transitions
TEAM	Technical Expert Advisory Group on Nutrition Monitoring
U5M	Under-5 mortality
UN	United Nations
UNICEF	United Nations Children's Fund
WHO	World Health Organization

ABSTRACT

Stunting prevalence is a core indicator of child health that is conventionally estimated as the proportion of children with height-for-age z-score (HAZ) values below -2 standard deviations (SDs), based on the World Health Organization (WHO) growth standards. Despite its widespread use in public health, stunting prevalence is conceptually problematic because it is often used to identify a subgroup of the population that is affected by undernutrition, rather than being correctly interpreted as a characteristic of the entire population. From a statistical perspective, stunting prevalence is a measure of location (MoL) of the HAZ distribution of an observed population relative to the distribution of a healthy standard population. A higher stunting prevalence usually represents a downward whole-population shift of the HAZ distribution, whereby even the tallest children in the population are shorter than expected. Moreover, stunting prevalence is based on cut-off values at the tails of the HAZ distribution and may therefore be more sensitive to imprecise data on height and date of birth than MoLs of central tendency (e.g., means, medians). We hypothesized that other linear growth metrics based on child height data may be less sensitive (i.e., more robust) to distribution asymmetry or the presence of influential outliers and therefore suitable alternatives or complementary approaches for describing the location of an observed HAZ distribution relative to international standards.

The objectives of this study were to (1) identify and describe a range of candidate linear growth metrics that could be used as alternative or complementary indicators for assessing population childhood linear growth and nutritional status and (2) assess and compare these potential metrics of child linear growth based on the relative strengths of their associations with other key population indicators and on the robustness of these associations against variations in anthropometric data quality. Height and date of birth data of children under 5 years of age from Demographic and Health Surveys (DHS) from 64 countries (2000 to 2018) were used to generate two types of linear growth metrics: estimates of descriptive statistics based on observed distributions (e.g., MoLs such as means and stunting prevalence) and regression model-derived estimates (e.g., predicted means at discrete ages or slopes of decline within a defined age range). DHS data were also used to generate indices for anthropometric data quality based on principal component analysis.

Correlations between each candidate linear growth metric and stunting prevalence among children under 5 were compared using the absolute value of Pearson correlation coefficients. Absolute values of Spearman rank correlations were used to compare pairwise associations between each linear growth metric and each of six population indicators of health and well-being (e.g., under-5 mortality [U5M], gross domestic product). Data quality was measured using indices composed of either three or six individual anthropometric data quality indicators (referred to as the 3Q index and 6Q index, respectively). Relationships between the metrics and population indicators were assessed using Spearman rank correlations (for a subset of three indicators) and linear mixed effects models (for all six indicators) to test their robustness against variations in anthropometric data quality. All analyses were performed using four approaches for identifying (i.e., flagging) and excluding HAZ outliers: (1) no flagging, and therefore no exclusions; (2) less restrictive flagging, which excluded HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median; (3) WHO flagging, which excluded HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median; and (4) Standardized Monitoring and Assessment of Relief and Transitions (SMART) flagging, which excluded HAZ values <-3 SDs and >+3 SDs from the observed sample mean.

Results showed that descriptive linear growth metrics and model-derived predicted HAZ values at 2 years, 3 years, and 5 years were strongly correlated with stunting prevalence, with absolute values of the Pearson correlation coefficients exceeding 0.90 across all flagging approaches. Pearson correlations for model-derived slopes of HAZ and height-for-age difference (HAD) from birth to 3 years of age ranged from moderate to strong (absolute values ≥0.46 and ≤0.77, respectively). Predicted HAZ at birth had one of the weakest coefficients and the largest range of correlations across flagging approaches (-0.35 to -0.68).

Spearman rank correlation coefficients for relationships between linear growth metrics and population indicators (i.e., absolute values) ranged from 0.01 to 0.73 with little variation across flagging approaches for each metric. Correlation strengths of the descriptive metrics tended to cluster near the midpoint of this range, with stunting prevalence generally ranking higher than other descriptive metrics. Conversely, correlations with model-derived metrics varied more widely; for example, predicted HAZ at birth had the weakest correlations and HAD slope from 0-3 years of age often had the strongest.

In linear mixed effects models, the associations between descriptive linear growth metrics and any of the population indicators were not significantly modified by the 3Q index. When data quality was defined using the 6Q index, significant modifying effects were observed for associations between stunting and the percentage of the population with an improved source of drinking water. The 3Q index had a significant modifying effect on the associations between U5M and four of the model-derived linear growth metrics when the WHO or SMART flagging approaches were used; however, these interactions were not significant when using the 6Q index. With few exceptions, the 3Q index had no observable modifying effect on the associations between the model-derived linear growth metrics and other population indicators; these findings were similar for analyses using the 6Q index.

Correlations between linear growth metrics and U5M were generally weaker with high (versus low) data quality, defined using the 3Q index; results were similar when data quality categorization was based on the 6Q index. However, differences between high- and low-quality strata were usually minor. Patterns for the other population indicators were inconsistent with the findings observed for U5M: For gross domestic product, the correlation strengths were greater at high (versus low) quality for half of the candidate metrics when using the 6Q index; for maternal education, correlation strengths were generally lower at high (versus low) quality using the 3Q index (similar to U5M), but this pattern was not consistent for many linear growth metrics when using the 6Q index.

In conclusion, numerous alternative linear growth metrics may be derived from the same population-level child height data that are widely used to estimate stunting prevalence. The majority of these metrics correlate strongly with stunting prevalence and may therefore have no substantial empirical advantages. Findings to date suggest that some of the alternative metrics may outperform stunting prevalence in terms of the strength of their correlations with other important population indicators. However, it was methodologically challenging to assess the sensitivity of metric performance to variations in anthropometric survey data quality. Therefore, findings so far are insufficient to lead to a recommendation to adopt one or more of the alternative linear growth metrics for use in tracking country or regional improvements in child growth. In further research, we will apply a revised methodological approach with a more comprehensive set of alternative metrics to further our understanding of the potential application of alternative linear growth metrics for assessing and tracking child health and nutritional status in low-and middle-income countries.

Key words: Alternative linear growth metrics, anthropometric data quality, nutritional status, Demographic and Health Surveys, global targets, height-for-age, linear growth, population-based surveys, stunting, Sustainable Development Goals

1 INTRODUCTION

Reducing the burden of child growth faltering is a global public health priority represented within the United Nations Sustainable Development Goals. Indicators based on child height have a prominent role in efforts to track progress in nutritional status and related programming, such as the United States Agency for International Development Multi-sectoral Nutrition Strategy.[1] It is conventional practice to use cutoffs at the tails of the population height distribution to represent the proportion of children who are stunted—those with heights more than 2 standard deviations (SDs) below the median for age and sex (i.e., height-for-age z-score [HAZ] <-2) based on the World Health Organization (WHO) Child Growth Standards—to estimate the prevalence of poor nutritional status in world regions/countries, as well as to track secular trends in and responses to public health interventions. The Demographic and Health Surveys (DHS) Program is one of the primary sources of population-representative data for low- and middle-income countries (LMICs) used to measure and track progress on child nutritional status.

An analysis of 179 DHS datasets (1993–2015) from 64 countries showed linear growth faltering as a predominantly whole-population condition, characterized by shifts in the entire population HAZ distribution rather than deviations of a small, high-risk subgroup at the tail end as implied by the use of the stunting indicator.[2] This observation highlighted the potential utility of describing population-level distributions of child height using alternative measures of location, including measures of central tendency (e.g., means, medians) and regression model-derived estimates that convey the displacement of an observed HAZ distribution relative to the WHO Child Growth Standards or a trajectory of change in height with age in that population (e.g., predicted mean HAZ at a discrete age, slope of decline of HAZ in a defined age range). A related concern is that metrics of early childhood growth are potentially susceptible to the quality of anthropometric measurements, which can vary within and between surveys.[3,4] An indicator that is relatively robust against excess dispersion and outliers due to errors in anthropometric measurements may improve the accuracy with which progress in linear growth is tracked and compared at the country and global levels.

1.1 Research objectives

The overall goal of this study was to identify metrics of early childhood linear growth at the population level that may be used as alternatives or complements to the conventional indicator of under-5 stunting prevalence. The specific objectives were to:

(1) Identify and describe a range of candidate linear growth metrics that could be used as alternative or complementary metrics for assessing population childhood growth and nutritional status.

(2) Assess and compare potential metrics of child linear growth based on the relative strengths of their associations with other key population indicators and on the robustness of these associations against variations in anthropometric data quality.

This working paper focuses on preliminary applications of a novel framework for the evaluation of candidate metrics based on child height data in a multi-country DHS dataset. Future extensions of this project will refine the methods and broaden the array of candidate linear growth metrics.

2 DATA AND METHODS

2.1 Anthropometric data sources

Individual-level height and corresponding sex and date-of-birth data were obtained from 145 DHS surveys conducted between 2000 (i.e., post DHS Phase III) and the public release of data as of January 2019 (Table 1 and Supplementary Table A1), representing 64 LMICs. The rationale for these inclusion criteria was to ensure consistency in the sampling frame, as anthropometric measures in previous DHS phases (i.e., phases I–III) were collected only from the children of women included in the Women's Questionnaire. In subsequent DHS phases, anthropometric measures were collected for all children in the surveyed households. Survey data were retained in the present analysis even in the rare cases when anthropometric data were excluded from the DHS final reports due to data quality issues (i.e., 2011-12 Benin and 2007 Jordan DHS surveys) or data quality issues were noted in the final reports. Furthermore, data analyses were restricted to de facto children as defined in the DHS methodology (i.e., children 0-59 months who slept in the household the night prior to survey administration), and anthropometric variables were generated only for children who had complete data on month and year of birth. If the day of birth was missing, the day of birth was imputed as 15 to calculate the age (in days) at which anthropometric data were collected.

Table 1 Characteristics of the analytical samples of 145 Demographic and Health Surveys conducted between 2000 and 2018 in 64 low- and middle-income countries, by outlier flagging approach

	No flagging[1]	Less restrictive flagging[2]	WHO flagging[3]	SMART flagging[4]
Median number (min, max) of children under 5 years of age per survey	5,718 (1,317, 244,170)	5,674 (1,309, 243,137)	5,472 (1,290, 239,588)	5,102 (1,152, 218,278)
Total number of children under 5 years of age across all surveys	1,237,430	1,229,768	1,208,759	1,115,119
Total percentage of data excluded [min, max]	–	0.62 [0.0, 4.5]	2.2 [0.03, 11.6]	9.2 [0.76, 31.4]

Note: SMART = Standardized Monitoring and Assessment of Relief and Transitions; WHO = World Health Organization

[1] No exclusion of outlying values.

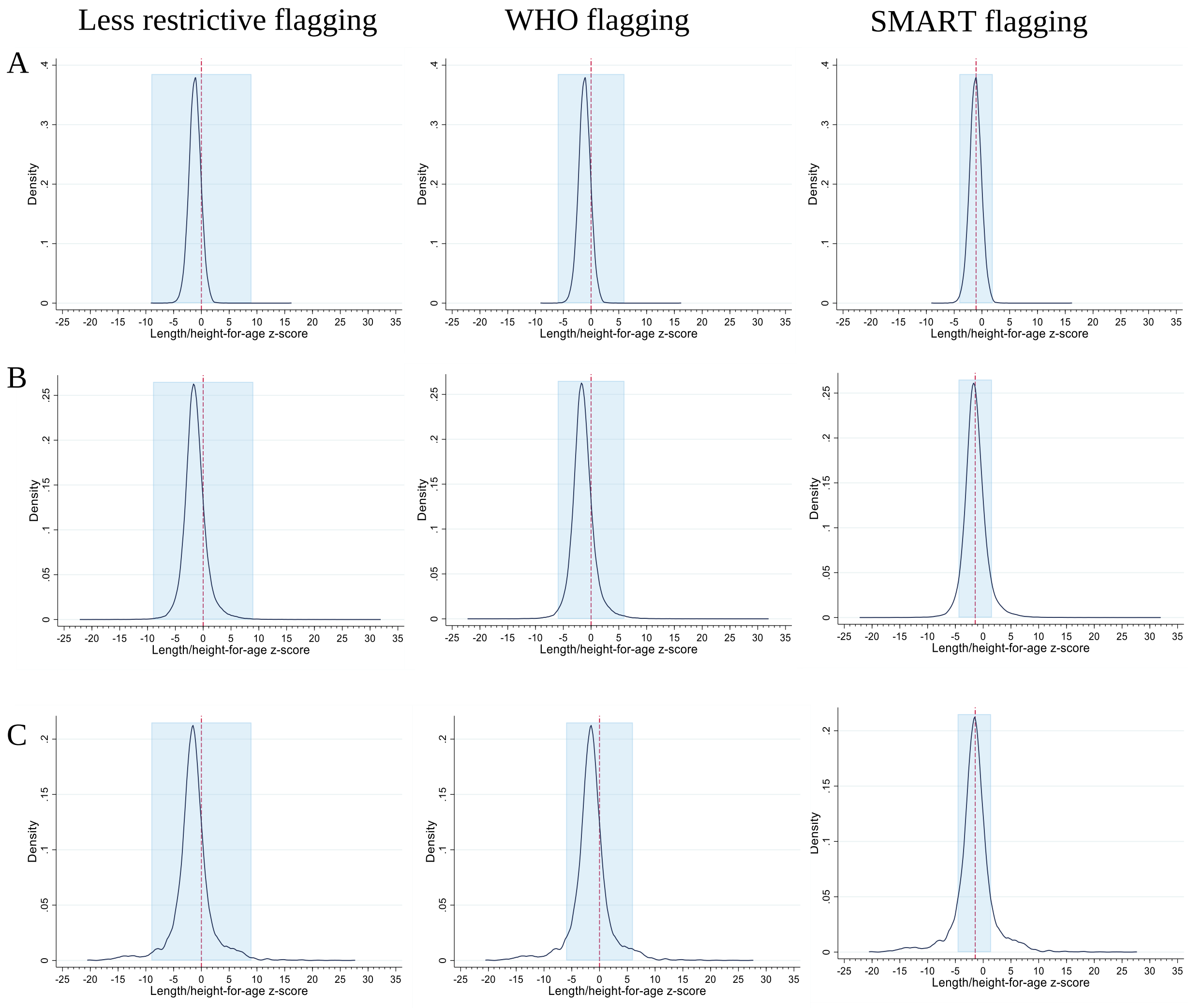
[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Linear growth metrics were generated using four approaches for identifying (i.e., flagging) and excluding implausible anthropometric values: (1) no flagging (no exclusion of outlying values), (2) less restrictive flagging (excluded HAZ values <-9 SDs and >+9 SDs from the age- and sex- specific WHO standard median), (3) WHO flagging (excluded HAZ values <-6 SDs and >+6 SDs from the age- and sex-specific WHO standard median), and (4) Standardized Monitoring and Assessment of Relief and Transitions (SMART) flagging (excluded HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean). Figure 1 presents three examples of the differential effect of the flagging approaches on the amount and distribution of data retained.

Figure 1 **Comparison of flagging approaches using data from the (A) 2012 Peru, (B) 2015 India, and (C) 2013 Sierra Leone Demographic and Health Surveys**



Note: Distributions are Kernel density plots. Shaded areas indicate the range of data retained for each flagging approach: less restrictive flagging (excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median), WHO flagging (excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median), and SMART flagging (excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean). Dotted lines represent the midpoint used.

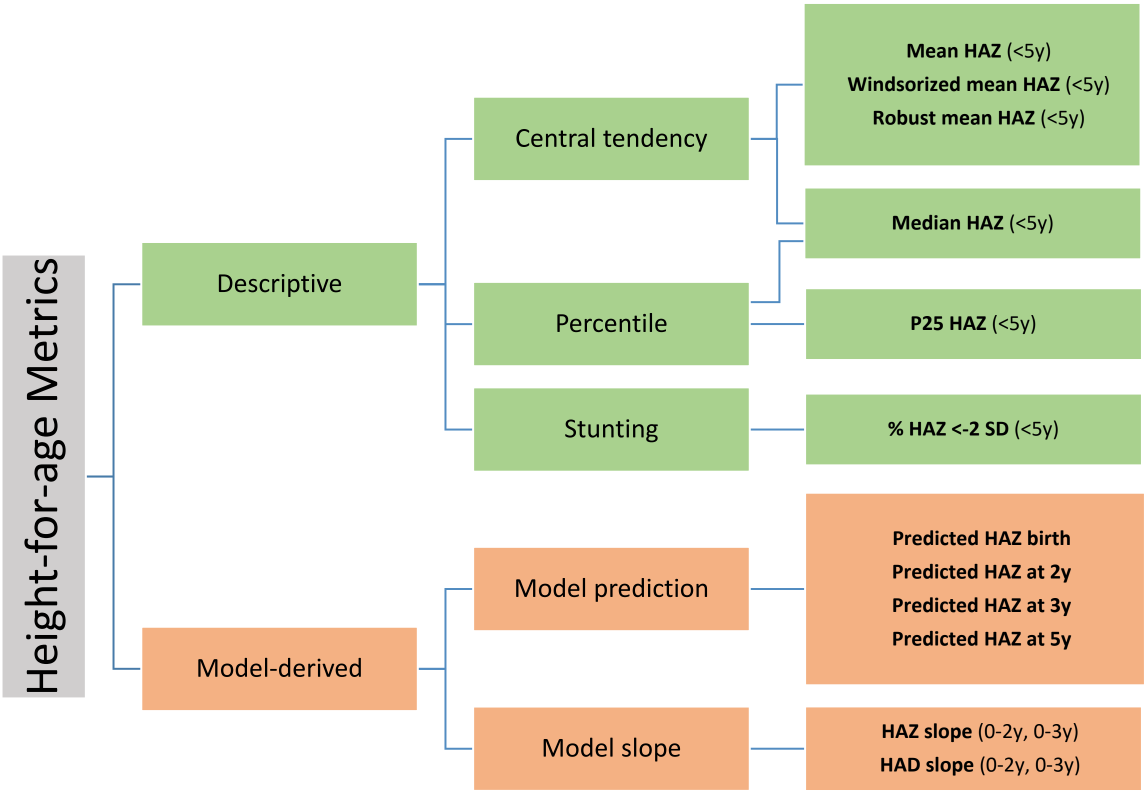
2.2 Identification, estimation, and prediction of candidate linear growth metrics

We scanned online resources and peer-reviewed literature to identify candidate statistical metrics for population-level early childhood linear growth. We focused on metrics that could be estimated empirically or derived (i.e., estimated or predicted) using regression modelling, which could be used to evaluate and assess variations in height distributions between populations. We were particularly interested in finding metrics that would be minimally influenced by data dispersion and outliers in child height datasets. Given the importance of accounting for complex survey design to obtain representative estimates using survey data, all selected candidate metrics can be generated using statistical tools that appropriately weight the data for sampling.

Using individual-level data for children under-5 from DHS surveys in LMICs, we generated two types of linear growth metrics for each survey: "descriptive" metrics (i.e., those estimated directly from the observed HAZ distribution without the use of regression models) and "model-derived" metrics (i.e., estimates of model predictions for specific ages and regression slopes based on raw height or HAZ data) (Figure 2, Table 2). All metrics were generated accounting for the complex design of DHS surveys.

The descriptive metrics were further divided into three classes: (1) measures of central tendency, (2) percentiles, and (3) stunting prevalence. Each metric was then estimated for all children under 5 years of age, within specified age intervals and at discrete ages. Model-derived metrics were classified into two groups: (1) predicted HAZ at specific ages based on a regression model and (2) regression slopes of HAZ or height-for-age difference (HAD) as a function of age within specific age bands (Figure 2). The model-derived predictions of mean HAZ or HAD at discrete ages (i.e., birth, 2 years, 3 years, 4 years, 5 years) were derived using a linear spline model regressing HAZ (or HAD) on age in days, with a knot at 730 days (2 years of age). Predicted HAZ, predicted HAD, and predicted mean height at discrete ages were expected to perform similarly in correlation analyses or regression models, so results are only shown for predicted HAZ. Fixed-effect slopes for each survey were estimated from linear regression models regressing HAZ or HAD on age within specified age intervals (0-2 years, 0-3 years).

Figure 2 Candidate descriptive and model-derived linear growth metrics



Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SD = standard deviation

Table 2 Definitions and methods of derivation of candidate linear growth metrics

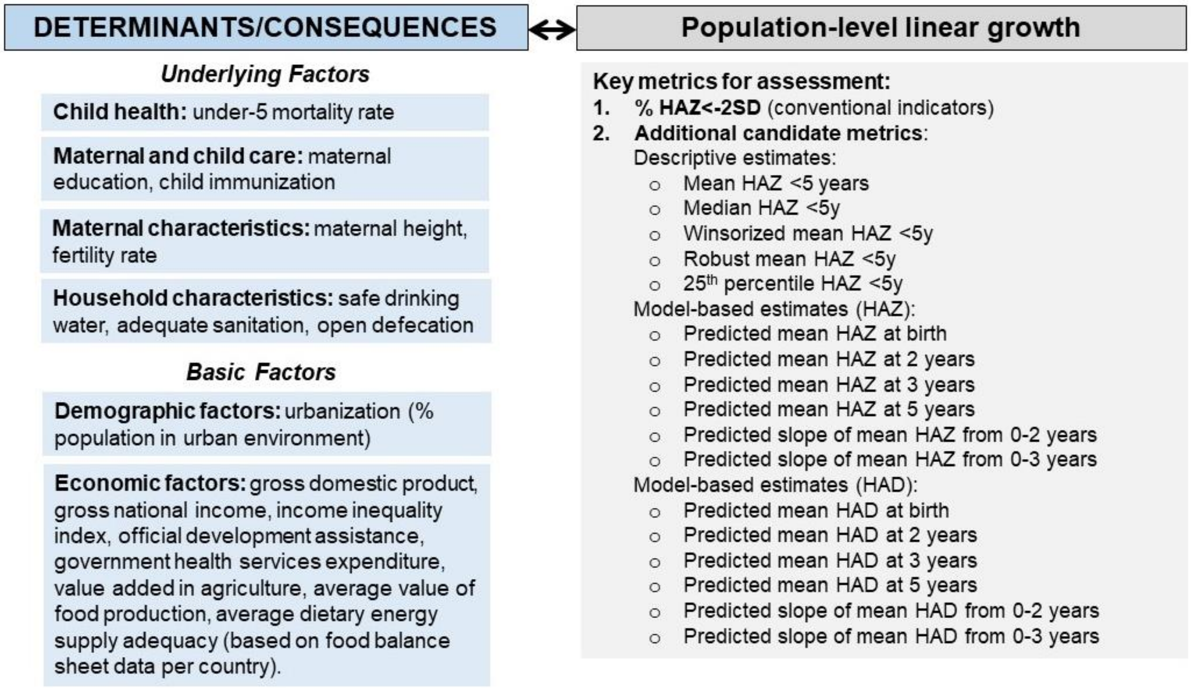
Mean HAZ	**Sum of all individual HAZ values, divided by the number of values**
Winsorized mean	Mean of HAZ distribution after setting the highest and lowest values to the 95th and 5th percentile, respectively
Robust Huber mean	Mean of a robust M-estimator, obtained by minimizing a loss function
Median HAZ	HAZ value that divides the ordered data in half
Percentile HAZ	Values at the Xth percentile of the HAZ distribution
Predicted HAZ	Model-derived HAZ at a specified age
Slope of HAZ	Model-derived slope within a specified age interval
Slope of HAD	Model-derived slope of HAD within a specified age interval; HAD was calculated for each child as the difference between the median sex- and age-specific height from the WHO Child Growth Standards and the measured height

2.3 Population indicators

A conceptual framework of childhood linear growth at the population level was developed using evidence from studies that examined cross-national variation in HAZ/stunting (e.g., via ecological or multi-level analyses)[5–11] (Figure 3). This framework informed the selection of population indicators used as the outcome variables in multivariable analyses to assess the extent to which metrics of child linear growth explained cross-national variations in the population indicators and to compare the strength of these relationships across metrics. To generate a parsimonious model, we selected variables from these studies that were shown to have a statistically significant association with child linear growth.

Since the intent was to identify candidate metrics for assessing cross-national variations in child linear growth, the framework was not designed to demonstrate causal effects of risk factors on child growth; rather, we aimed to identify a list of population-level determinants and consequences that were already known to be associated with child linear growth based on current evidence. Variables from previous studies that examined the effect of individual-level characteristics on child height (e.g., infant feeding practices) were not included in the framework.

Figure 3 **Conceptual framework of population-level childhood linear growth, including the candidate descriptive and model-based linear growth metrics (derived from height-for-age z-score or height-for-age difference)**



Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SD = standard deviation. Results for predicted mean HAD at specified ages were not reported because they are perfectly correlated with predicted mean HAZ.

The population indicators (and their respective data sources) selected for inclusion in the analyses are described in Table 3. Maternal education was estimated using the DHS survey data (internal), and data for the other indicators were obtained using external data sources matched by calendar year. Indicators were selected from a more complete list (Figure 3), based on consensus among authors, input from an external advisory group (the Growth Metrics Project Review Group), and the availability of data across all countries and survey years.

Table 3 Population indicators selected for inclusion in validation analyses

Domain	Variable/Indicator	Operational definition	Data source
Child health	Under-5 mortality rate	Number of deaths before 5 years of age per 1,000 live births	UN IGME
Economy	Gross domestic product	Measure of a country's economic output per capita, adjusted for purchasing power parity in 2011 constant international dollars	World Bank
Maternal characteristics	Maternal education	% of women who completed secondary school or higher	DHS
Household characteristics related to water, sanitation, and hygiene	% improved drinking water source	Proportion of the population with access to safely managed drinking water (from an improved source located on premises, available when needed, and free from fecal and priority chemical contamination)	WHO/ UNICEF JMP
	% improved sanitation facilities	Proportion of the population that uses safely managed sanitation services (improved facilities that are not shared with other households and for which excreta are safely disposed of in situ or transported and treated off-site)	WHO/ UNICEF JMP
	% open defecation	Proportion of the population that disposes of human feces in fields, forests, bushes, open bodies of water, beaches, and other open spaces, or disposes of them with solid waste	WHO/ UNICEF JMP

Note: DHS = Demographic and Health Surveys; IGME = United Nations Inter-agency Group for Child Mortality Estimation; JMP = Joint Monitoring Programme for Water Supply, Sanitation, and Hygiene; UN = United Nations; UNICEF = United Nations Children's Fund; WHO = World Health Organization

2.4 Data quality index

Several indicators of anthropometric data quality are recommended for use in population-based surveys by the WHO/United Nations Children's Fund (UNICEF) Anthropometric Data Quality Working Group under the Technical Expert Advisory Group on Nutrition Monitoring (TEAM).[12] From among the recommended indicators, we selected three that were not directly related to the HAZ distribution to construct a data quality index (3Q index): (1) proportion of observations with complete date of birth, (2) proportion of observations with anthropometry measured, and (3) digit preference for height (Table 4). We restricted our selection to indicators not directly related to the HAZ distribution because the candidate linear growth metrics were themselves based on the HAZ distribution and may have been intrinsically related to indicators of quality derived from the same data.

We previously developed a composite data quality index for HAZ using six individual anthropometric data quality indicators (i.e., the 6Q index) for use in multi-survey epidemiologic analyses of child nutritional status.[13] The 6Q index included the same three indicators used in the 3Q index, as well as three additional data quality indicators related to the dispersion of HAZ: (1) the proportion of flagged values, (2) the absolute difference in HAZ by reported month of birth (MOB), and (3) the SD of HAZ (Table 4). The 6Q index therefore has the advantage of capturing anthropometric data quality more comprehensively than the 3Q index, particularly when measurement error in anthropometric data may be due to poor technique in height measurements. However, given the expected collinearity between the linear growth metrics and the indicators of data quality in the 6Q index that are based on the HAZ distribution (i.e., the proportion of flagged values, SD), we assumed that using a 6Q index in our analyses would variably attenuate estimates of the associations between linear growth metrics and population indicators, thereby causing inconsistent performance of the metrics across surveys. The 3Q index, which is primarily based on indicators related to survey non-response, largely circumvents this issue of collinearity between the data quality index and candidate metrics of linear growth.

Table 4 Individual data quality indicators included in the anthropometric data quality indices

Indicator	Operational definition
Completeness of date of birth	Complete data for date of birth was defined as the percentage of children 0-59 months with complete month and year of birth. Birth dates with missing day of birth were not considered to be incomplete. We replaced missing day of birth with "15" because the 2006 WHO Child Growth Standards can be used to generate anthropometric indices based on age in months.
Completeness of height measurement	Complete measurement was defined as the percentage of children 0-59 months with height data recorded. Data for participants who refused, were not present, or were not measured for another reason were considered incomplete.
Digit preference for height (any digit 0-9)	The index of dissimilarity was used to numerically characterize digit preference for height. It was calculated as the sum of the absolute difference between the observed and expected percentages divided by 2. $Index\ of\ dissimilarity = \sum \text{abs (expected percentage} - \text{observed percentage) / 2}$
Absolute difference in HAZ by MOB	This absolute difference was the difference between the mean HAZ value in the month of January and the mean HAZ value in the month of December.[14] It should have been close to zero if there were no systematic errors in date of birth reporting.
Biologically implausible (i.e., WHO-flagged) values	This was the proportion of values more than six standard deviations above or below the median z-score of the reference population according to the WHO flagging convention.
Dispersion of HAZ	The distribution of HAZ was captured using the standard deviation after the removal of flagged values based on the WHO flagging convention. The higher the standard deviation was above 1.0, which is the expected value for a Gaussian distribution, the more likely there was a data quality issue.

Note: HAZ = height-for-age z-score; MOB = month of birth; WHO = World Health Organization

2.4.1 Method for data quality index construction

We generated each data quality indicator for each survey using the unweighted samples (i.e., without accounting for the complex survey design). Principal component analysis (PCA) was used to generate an index that summarized the largest variability in the data using the correlation matrix of a linear combination of variables in the first component. Following PCA analysis, we generated a predicted factor score and used the values of the score to assess anthropometric data quality across surveys. Lower values reflected lower data quality, and higher values reflected higher data quality relative to other surveys included in the analysis. Characteristics of each anthropometric data quality indicator across the 145 DHS surveys and factor loadings for each indicator are summarized in Table 5. We also estimated the Pearson correlation coefficients for all pairwise correlations among the indicators of data quality (Table 6).

Table 5 Summary statistics and factor loadings in principal component analyses of six indicators of anthropometric data quality used to develop anthropometric data quality indices for height-for-age in 145 Demographic and Health Surveys from 64 countries

	Summary statistics			PCA	
Data quality indicator	**Median**	**IQR (25th, 75th)**	**Range (min, max)**	**Factor Loadings 3Q index**	**Factor Loadings index**
Completeness of date of birth, %	99	98, 100	20 (80, 100)	0.710	0.504
Completeness of height measurement, %	96	93, 98	30 (70, 100)	0.728	0.334
Digit preference for height (any digit 0-9), index of dissimilarity %	15	10, 24	80 (3.1, 83)	0.679	0.665
Biologically implausible, %	1.6	0.72, 2.9	11.6 (0.03, 11.6)	-	-0.199
Absolute difference in HAZ by month of birth, z-score	0.25	0.13, 0.38	0.90 (0.001, 0.90)	-	0.886
Standard deviation of HAZ, z-score	1.59	1.41, 1.78	1.4 (1.07, 2.47)	-	0.534

Note: HAZ = height-for-age z-score; IQR = interquartile range; PCA = principal component analysis

Overall, completeness of date of birth and anthropometric measures was high, ranging from 70% to 100% across all surveys (Table 5). A higher index of dissimilarity reflected more evidence of digit preference. The overall index of dissimilarity for height was 15%, suggesting a non-negligible average degree of digit preference; however, it ranged widely across surveys from 3% (Guatemala 2015) to 83% (Sao Tome and Principe 2008). The percentage of biologically implausible values for HAZ (based on the WHO definition) ranged from less than 1% to almost 12%. The absolute difference in HAZ by MOB, which would be 0 if there were no systematic errors in date of birth reporting, was 0.25 overall and up to 0.90. The overall range (difference between the 25th and 75th percentile) of HAZ SD was 1.4, with a maximum value of 2.47. A healthy standard population is expected to have a mean HAZ of 0 and SD of 1, and we would expect dispersion in high-quality studies of child height to typically be closer to SD of 1, even in settings where the population mean HAZ is well below 0. SDs well above 1 are therefore assumed to reflect imprecision in height or age data, but may also be due to population heterogeneity.[3]

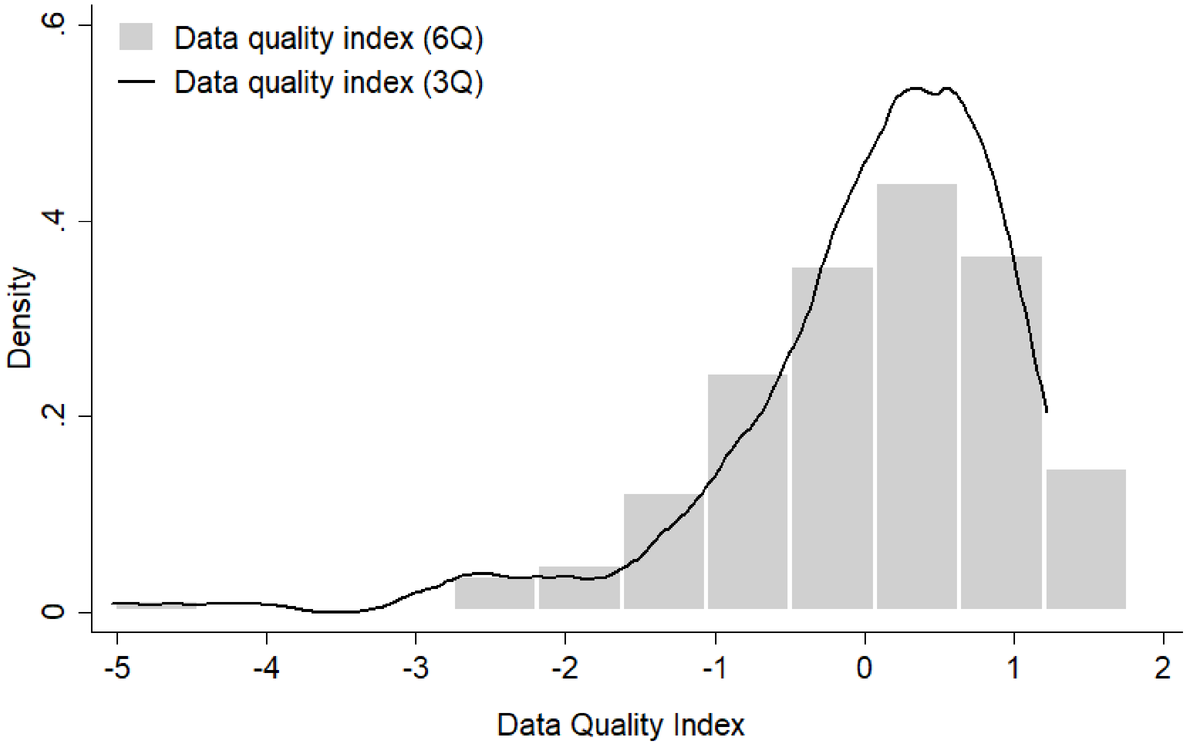
Table 6 Pearson correlation coefficient matrix of data quality indices and six individual metrics of anthropometric data quality in 145 Demographic and Health Surveys from 64 countries

Data quality indicator	**Data quality index (6Q)**	**Data quality index (3Q)**	**Date of birth complete**	**Anthro-pometry measured**	**Digit preference for height**	**Biologically implausible values**	**Difference in HAZ by MOB**	**HAZ SD**
Data quality index (6Q)	1.00							
Data quality index (3Q)	0.70	1.00						
Completeness of date of birth	0.51	0.71	1.00					
Completeness of height measurements	0.33	0.73	0.27	1.00				
Digit preference height	-0.67	-0.68	-0.24	-0.26	1.00			
Biologically implausible values	-0.89	-0.49	-0.35	-0.24	0.48	1.00		
Difference in HAZ by MOB	-0.53	-0.09	-0.09	0.03	0.13	0.37	1.00	
SD of HAZ	-0.88	-0.37	-0.28	-0.06	0.47	0.80	0.49	1.00

Note: HAZ = height-for-age z-score; MOB = month of birth; SD = standard deviation

The restricted 3Q index (which did not include any indicators of data quality related to the dispersion of HAZ) and the 6Q index were moderately correlated with one another (0.70) (Figures 4 and 5). Countries were ranked in descending order of quality using the most recent survey using both 6Q and 3Q indices (Figure 5). Where possible, these rankings were compared with external determinations of survey quality based on DHS reports and author expertise. Although there were notable differences in the values assigned to each survey based on the two data quality indices (Figure 5), the restricted 3Q index had no or a weak relationship with mean HAZ and prevalence of stunting in comparison to the stronger correlations between the 6Q index and mean HAZ and stunting prevalence (See Section 3.3). These findings supported a decision to use the 3Q index to assess robustness of candidate metrics against poor anthropometric data quality. Therefore, we primarily relied on the 3Q index to assess the extent to which the associations between each linear growth metric and the population indicators were robust against variations in anthropometric data quality. We additionally used the 6Q index in sensitivity analyses to test the robustness of inferences regarding the alternative metrics and anthropometric data quality.

Figure 4 **Distribution of anthropometric data quality indices for 145 Demographic and Health Surveys based on six individual indicators of anthropometric data quality for height-for-age z-score (6Q) or a restricted index based on three individual data quality indicators unrelated to metrics of height-for-age z-score (3Q)**



Note: Lower index values reflect relatively worse data quality (compared with other surveys in the analysis).

Figure 5 Country ranking based on anthropometric data quality indices (6Q and 3Q) for height-for-age z-scores for the most recent Demographic and Health Surveys from 64 countries

Note: Lower index values reflect relatively worse data quality (compared with other surveys in the analysis).

2.5 Statistical analysis

Pearson correlation coefficients were used to describe the strength of the relationships between linear growth metrics and stunting prevalence. Spearman rank correlation coefficients were used to compare the linear growth metrics in terms of their covariation with the population indicators listed in Table 3. Higher correlations with population indicators known to be associated with child growth and nutritional status were interpreted as indicating both greater validity of the linear growth metric and potential utility as an alternative or complement to stunting. Spearman rank correlation coefficients were used because, unlike the correlations of alternative linear growth metrics with stunting, several metric-indicator relationships were non-linear according to visual assessment of scatterplots and LOWESS curves. Exploratory descriptive analyses of the relationship between stunting or mean HAZ and under-5 mortality (U5M) also showed that surveys from Benin were highly influential outliers; therefore, sensitivity analyses were conducted to confirm that the inclusion of this survey did not meaningfully impact the relative strengths of the correlation coefficients across metrics (data not shown).

Generalized linear mixed models with country-specific random intercepts were used to model each population indicator listed in Table 3 (i.e., outcome) as a function of the linear growth metric, the anthropometric data quality index, and the interaction between data quality and growth metric. A significant interaction (i.e., modifying effect of quality index on the association between the population indicator and the growth metric) was interpreted as an indication that the metric may be sensitive to variations in anthropometric data quality, and thus less suitable for assessing child growth and nutritional status using population-level survey data. Since the main effects were from models that included the interaction terms, the point estimates for the main effects were interpreted as the magnitude of the association between the linear growth metric and the population indicator when the anthropometric data quality index was equal to 0, which is approximately at the midpoint within the distribution of anthropometric data quality index values (Figure 5). Since the 2011-12 survey from Benin was an influential outlier, and these data were also suppressed from DHS reports, this survey was removed from all regression analyses. All models were replicated across the four flagging approaches for identifying and excluding implausible anthropometric values (i.e., no flagging, less restrictive flagging, WHO flagging, and SMART flagging).

To further assess the robustness of linear growth metrics against variations in data quality, Spearman rank correlation coefficients for the association between each candidate metric and a subset of the population indicators (i.e., U5M, gross domestic product [GDP], and maternal education) were estimated and compared across strata of low and high anthropometric data quality (defined as quality index values below and above the median, respectively). For each population indicator, the magnitude and directionality of the correlation were compared descriptively across metrics, strata of data quality (defined by using either the 3Q or 6Q index), and the four flagging approaches (i.e., no flagging, less restrictive flagging, WHO flagging, and SMART flagging). A difference in the strength or direction of an association between low- and high-quality indices was interpreted as sensitivity of the metric to data quality.

The most recent surveys from each country were used to compare linear growth metrics and to assess relative strengths of their associations with population indicators. Analyses to characterize anthropometric data quality indices and to assess the robustness of the associations between linear growth metrics and population indicators against variations in anthropometric data quality included all surveys. Three population health indicators (improved drinking water source, improved sanitation facilities, and open defecation) were missing for 15 surveys and maternal education was missing for one survey (Table A1).

Analyses were conducted using STATA version 15.1 software (StataCorp LLC, College Station, Texas, USA), and the code was independently verified by a second data analyst.

3 RESULTS

3.1 Correlations of candidate linear growth metrics with stunting prevalence

Most of the linear growth metrics explored were strongly inversely correlated with stunting prevalence, and correlation coefficients for the descriptive metrics (e.g., mean HAZ) all had magnitudes of 0.95 or higher across all flagging approaches (Table 7). Pearson correlation coefficients for the model-derived metrics, including predicted HAZ at 2, 3, and 5 years were also strong, whereas slopes of HAZ and HAD from 0-2 or 0-3 years were moderately to strongly correlated with stunting. Predicted HAZ at birth had the weakest coefficients at -0.35 (i.e., for no flagging) and -0.37 (i.e., for WHO flagging) and the widest range of correlation strengths across flagging approaches (Table 7).

Table 7 **Pearson correlation coefficients for the relationships between linear growth metrics and stunting prevalence among children <5 years of age in the most recent Demographic and Health Surveys from 64 countries**

Height-for-age metric	No flagging[1]	Less restrictive flagging[2]	WHO flagging[3]	SMART flagging[4]
Mean HAZ	-0.95	-0.95	-0.95	-0.98
Median HAZ	-0.97	-0.97	-0.97	-0.98
25th percentile HAZ	-0.96	-0.96	-0.97	-0.99
Windsor mean HAZ	-0.96	-0.96	-0.96	-0.98
Robust mean HAZ	-0.97	-0.97	-0.97	-0.98
Predicted HAZ at birth	-0.35	-0.41	-0.37	-0.68
Predicted HAZ at 2 years	-0.92	-0.92	-0.91	-0.95
Predicted HAZ at 3 years	-0.95	-0.94	-0.94	-0.97
Predicted HAZ at 5 years	-0.92	-0.92	-0.92	-0.94
HAZ slope 0-2 years	-0.51	-0.50	-0.56	-0.48
HAZ slope 0-3 years	-0.52	-0.51	-0.55	-0.46
HAD slope 0-2 years	-0.68	-0.68	-0.72	-0.73
HAD slope 0-3 years	-0.74	-0.74	-0.74	-0.77

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

We also assessed the correlations between population indicators and under-5 mean HAZ, stunting prevalence, and data quality indices (Table 8). As expected, mean HAZ and stunting prevalence were relatively strongly correlated with population indicators. The data quality indices were also correlated with population indicators. The higher correlation coefficients for relationships between the 6Q index and population indicators suggested that countries with worse population indicators had worse anthropometric data quality. In comparison, the 3Q index was relatively weakly correlated with population indicators. These findings further supported our decision to base the primary inferences regarding robustness of linear growth metrics against variations in data quality on the 3Q index.

Table 8 **Spearman rank correlations of under-5 mean height-for-age z-score, stunting prevalence, and data quality indices with population indicators from 49 countries**

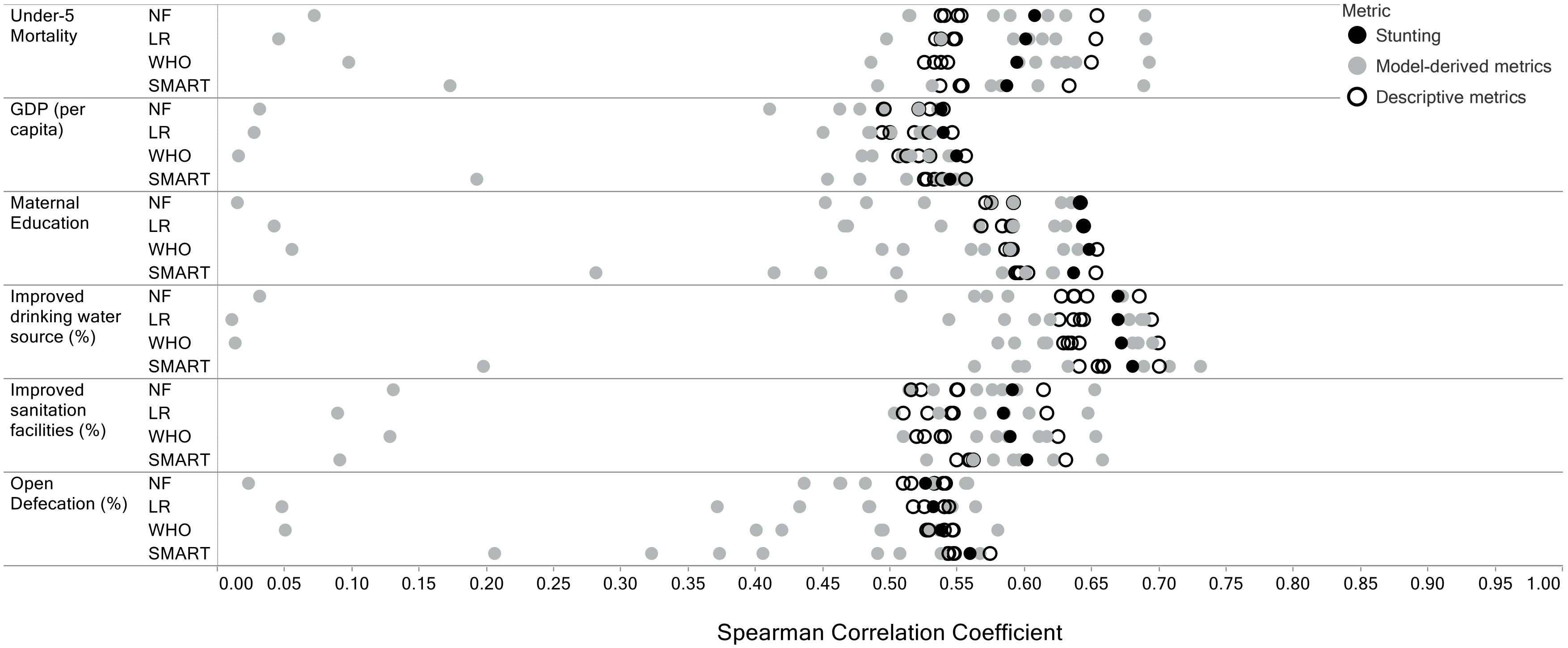
Population indicators	Mean HAZ	Stunting prevalence	Data quality index (3Q)	Data quality index (6Q)
Under-5 mortality rate	-0.55	0.65	0.02	-0.40
Gross domestic product	0.59	-0.63	0.01	0.26
Maternal education	0.62	-0.70	0.15	0.31
% improved drinking water source	0.64	-0.68	-0.06	0.28
% improved sanitation facilities	0.56	-0.63	-0.05	0.30
% open defecation	-0.54	0.54	-0.08	-0.18

Note: HAZ = height-for-age z-score

3.2 Correlations between linear growth metrics and population indicators

Absolute values of Spearman rank correlation coefficients for correlations between linear growth metrics and population indicators (Figure 6) ranged from 0.05 (predicted HAZ at birth, less restrictive flagging) to 0.69 (HAD slope 0-3 years, all flagging) for U5M; from 0.02 (predicted HAZ at birth, WHO flagging) to 0.56 (25th percentile [25P] HAZ, WHO flagging) for gross GDP; from 0.02 (predicted HAZ at birth, no flagging) to 0.65 (P25 HAZ, WHO flagging) for maternal education; from 0.01 (predicted HAZ at birth, less restrictive flagging) to 0.73 (HAD slope 0-3 years, SMART flagging) for improved drinking water source; from 0.09 (predicted HAZ at birth, less restrictive flagging) to 0.66 (HAD slope 0-3 years, SMART flagging) for access to improved sanitation facilities; and from 0.02 (predicted HAZ at birth, no flagging) to 0.58 (HAD slope 0-3 years, WHO flagging) for open defecation. The correlation strengths of most descriptive metrics tended to cluster near the middle of the range, whereas the model-derived metrics accounted for the spread from the weakest to the strongest values (Figure 6). Across all linear growth metrics, the metrics that frequently demonstrated the strongest correlations with population indicators were stunting prevalence, HAD slope 0-3 years, predicted HAZ at 3 years, and P25 HAZ (Figure A1). Stunting prevalence generally ranked higher than other descriptive metrics (except for P25 HAZ) and was among the most strongly correlated with GDP and maternal education (across all flagging approaches), as well as with U5M (no flagging and SMART flagging only).

Figure 6 **Absolute values of Spearman rank correlation coefficients of stunting and other descriptive or model-derived linear growth metrics and six population indicators using data from the most recent Demographic and Health Surveys in 64 low- and middle-income countries, applying four flagging approaches for identifying and removing extreme height-for-age z-scores values**



Note: GDP = gross domestic product; NF = no flagging; LR = less restrictive; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization; Due to missing data the sample sizes were lower for maternal education (N=63 countries) and for improved drinking water source, improved sanitation facilities, and open defecation (N=50 countries).

3.3 Correlations between linear growth metrics and anthropometric data quality

The restricted 3Q index had weak or null relationships with all candidate linear growth metrics (Table 9a). The absolute values of the Spearman rank correlation coefficients ranged from <0.01 to 0.09 without flagging, 0.01 to 0.13 with less restrictive flagging, 0.02 to 0.09 with WHO flagging, and <0.01 to 0.21 with SMART flagging (Table 9a). Spearman rank correlations with the 6Q index tended to be of moderate strength, with absolute values ranging from 0.05 to 0.37 without flagging, 0.03 to 0.36 with less restrictive flagging, 0.05 to 0.32 with WHO flagging, and <0.01 to 0.28 with SMART flagging (Table 9b).

Table 9a Spearman rank correlation coefficients for the relationships between linear growth metrics and the restricted (3Q) survey data quality index in the most recent Demographic and Health Surveys from 64 countries

Height-for-age metric	No flagging[1]	Less restrictive flagging[2]	WHO flagging[3]	SMART flagging[4]
Stunting prevalence	-0.07	-0.06	-0.05	-0.04
Mean HAZ	<0.01	-0.02	-0.03	<-0.01
Median HAZ	-0.01	-0.02	-0.03	-0.02
25th Percentile HAZ	0.09	0.08	0.06	0.04
Winsorized m**ean HAZ**	0.01	-0.01	-0.03	<-0.01
Robust Huber HAZ	0.01	0.01	-0.02	-0.01
Predicted HAZ at birth	0.08	0.07	0.06	0.13
Predicted HAZ at 2 years	-0.02	-0.06	-0.09	-0.03
Predicted HAZ at 3 years	-0.02	-0.05	-0.06	-0.03
Predicted HAZ at 5 years	-0.01	-0.04	-0.08	-0.09
HAZ slope 0-2 years	-0.08	-0.13	-0.07	-0.18
HAZ slope 0-3 years	-0.06	-0.08	-0.07	-0.21
HAD slope 0-2 years	-0.07	-0.11	-0.07	-0.13
HAD slope 0-3 years	-0.01	-0.04	-0.06	-0.10

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Table 9b **Spearman rank correlation coefficients for the relationships between linear growth metrics and the extended composite (6Q) anthropometric data quality index in the most recent Demographic and Health Survey from 64 countries**

Height-for-age metric	No flagging[1]	Less restrictive flagging[2]	WHO flagging[3]	SMART flagging[4]
Stunting prevalence	-0.31	-0.30	-0.28	-0.24
Mean HAZ	0.15	0.12	0.10	0.16
Median HAZ	0.14	0.14	0.12	0.14
25th percentile HAZ	0.37	0.36	0.32	0.28
Winsorized mean HAZ	0.18	0.15	0.11	0.16
Robust Huber HAZ	0.17	0.16	0.13	0.15
Predicted HAZ at birth	-0.05	-0.03	-0.05	0.24
Predicted HAZ at 2 years	0.16	0.11	0.07	0.14
Predicted HAZ at 3 years	0.17	0.13	0.10	0.13
Predicted HAZ at 5 years	0.17	0.14	0.09	0.07
HAZ slope 0-2 years	0.18	0.13	0.18	<0.01
HAZ slope 0-3 years	0.19	0.15	0.14	-0.07
HAD slope 0-2 years	0.19	0.12	0.15	0.06
HAD slope 0-3 years	0.22	0.19	0.15	0.08

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

3.4 Assessment of anthropometric data quality as a modifier of the associations between linear growth metrics and population indicators

3.4.1 Linear mixed effects regression analyses of the modifying effect of anthropometric data quality on the associations between linear growth metrics and population indicators

In linear mixed effects models including all available surveys except Benin 2012 (n=144), associations between stunting and any of the six population indicators were not significantly modified by anthropometric data quality, as defined using the 3Q index (Supplementary Figure A2). When data quality was defined using the 6Q index, a significant interaction between the quality index and stunting was observed in the model for percent improved drinking water source (WHO flagging) (Supplementary Figure A3).

In models for U5M, the 3Q index had a significant modifying effect on the associations between U5M and predicted HAZ at 2 years and predicted HAZ at 3 years using the WHO flagging approach (Supplementary Figure A4). With SMART flagging, significant modifying effects of the 3Q index were observed for associations between U5M and predicted HAZ at 5 years, and between U5M and HAD slope 0-2 years (Supplementary Figure A4). The association between predicted HAZ at 5 years and maternal education was also significantly modified by the 3Q index when SMART flagging was used (Supplementary Figure A6). The 6Q index demonstrated effect modification for predicted HAZ at birth when using the SMART flagging approach in models with U5M and GDP (Supplementary Figures A10 and A11), as well as for the association between mean HAZ and U5M when using the no flagging approach (Supplementary Figure A10). For all other models, no effect modification by the 3Q or 6Q indices was observed for associations between linear growth metrics and percent improved drinking water source, improved sanitation facilities, or open defecation (Supplementary Figures A7 to A9 and A13 to A15).

3.4.2 Spearman rank correlation coefficients for the associations between linear growth metrics and population indicators at low and high anthropometric data quality

Among the most recent surveys from all 64 countries, the magnitudes of the Spearman correlation coefficients for linear growth metrics and U5M were generally lower among high quality surveys than low quality surveys as defined using the 3Q index for anthropometric data quality (Table 10a). This indicates that surveys with higher data quality demonstrated relatively weaker associations between linear growth and U5M. However, the average magnitude of the differences across flagging approaches was generally less than 0.15, and the average maximum was 0.27 for predicted HAZ at 5 years. As expected, linear growth metrics (other than stunting) were inversely associated with U5M; however, predicted HAZ at birth was an exception, as it demonstrated weakly positive associations with U5M when the no flagging, less restrictive flagging, and WHO flagging approaches were used.

Results were similar when data quality was defined using the 6Q index (Table 10b), although average differences in the strengths of correlations between high- and low-quality surveys were slightly larger (up to 0.19) than with the 3Q index. The highest average difference between low- and high-quality surveys was 0.31 for predicted HAZ at 5 years. Similar to the 3Q index, predicted HAZ at birth was the only growth metric that was more strongly correlated with U5M among high quality surveys than low quality surveys and showed positive associations with U5M when the no flagging, less restrictive flagging, and WHO flagging approaches were used.

Table 10a Spearman rank correlation coefficients for the relationships between linear growth metrics and under-5 mortality at high and low data quality (defined as above or below the 50th percentile of the restricted [3Q] anthropometric data quality index) in the most recent Demographic and Health Survey from 64 countries

	No flagging[1]		Less restrictive flagging[2]		WHO flagging[3]		SMART flagging[4]	
Height-for-age metric	**Low**	**High**	**Low**	**High**	**Low**	**High**	**Low**	**High**
Stunting prevalence	0.66	0.54	0.65	0.53	0.62	0.53	0.64	0.51
Mean HAZ	-0.59	-0.47	-0.61	-0.46	-0.61	-0.45	-0.62	-0.48
Median HAZ	-0.62	-0.46	-0.62	-0.46	-0.60	-0.46	-0.62	-0.46
25th percentile HAZ	-0.68	-0.60	-0.68	-0.60	-0.68	-0.60	-0.67	-0.57
Winsorized mean HAZ	-0.60	-0.47	-0.61	-0.47	-0.61	-0.47	-0.61	-0.48
Robust Huber HAZ	-0.62	-0.47	-0.62	-0.46	-0.61	-0.47	-0.62	-0.47
Predicted HAZ at birth	0.13	0.05	0.05	0.06	0.17	0.10	-0.16	-0.16
Predicted HAZ at 2 years	-0.68	-0.59	-0.68	-0.58	-0.65	-0.56	-0.68	-0.55
Predicted HAZ at 3 years	-0.65	-0.55	-0.67	-0.56	-0.67	-0.55	-0.65	-0.51
Predicted HAZ at 5 years	-0.64	-0.37	-0.64	-0.36	-0.64	-0.33	-0.59	-0.38
HAZ slope 0-2 years	-0.53	-0.51	-0.52	-0.53	-0.64	-0.60	-0.56	-0.50
HAZ slope 0-3 years	-0.62	-0.53	-0.65	-0.56	-0.68	-0.57	-0.61	-0.57
HAD slope 0-2 years	-0.57	-0.57	-0.54	-0.60	-0.66	-0.60	-0.64	-0.52
HAD slope 0-3 years	-0.65	-0.66	-0.68	-0.67	-0.70	-0.68	-0.71	-0.67

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

Note: Sample sizes: low quality surveys = 26; high quality surveys = 38.

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Table 10b Spearman rank correlation coefficients for the relationships between linear growth metrics and under-5 mortality at high and low data quality (defined as above or below the 50th percentile of the extended composite [6Q] anthropometric data quality index) in the most recent Demographic and Health Survey from 64 countries

	No flagging[1]		Less restrictive flagging[2]		WHO flagging[3]		SMART flagging[4]	
Height-for-age metric	Low	High	Low	High	Low	High	Low	High
Stunting prevalence	0.59	0.46	0.58	0.46	0.55	0.46	0.55	0.45
Mean HAZ	-0.57	-0.37	-0.56	-0.37	-0.56	-0.38	-0.56	-0.41
Median HAZ	-0.55	-0.39	-0.54	-0.39	-0.52	-0.40	-0.53	-0.39
25th percentile HAZ	-0.64	-0.50	-0.63	-0.49	-0.63	-0.50	-0.60	-0.49
Winsorized mean HAZ	-0.58	-0.37	-0.57	-0.37	-0.57	-0.39	-0.56	-0.41
Robust Huber HAZ	-0.56	-0.40	-0.55	-0.40	-0.54	-0.40	-0.56	-0.41
Predicted HAZ at birth	0.08	0.15	0.01	0.16	0.10	0.20	-0.11	-0.05
Predicted HAZ at 2 years	-0.70	-0.46	-0.67	-0.48	-0.65	-0.48	-0.60	-0.48
Predicted HAZ at 3 years	-0.67	-0.43	-0.65	-0.44	-0.63	-0.46	-0.58	-0.44
Predicted HAZ at 5 years	-0.61	-0.29	-0.61	-0.28	-0.59	-0.25	-0.55	-0.28
HAZ slope 0-2 years	-0.53	-0.46	-0.54	-0.50	-0.63	-0.55	-0.53	-0.53
HAZ slope 0-3 years	-0.59	-0.49	-0.63	-0.53	-0.67	-0.56	-0.64	-0.59
HAD slope 0-2 years	-0.59	-0.50	-0.59	-0.53	-0.63	-0.55	-0.57	-0.51
HAD slope 0-3 years	-0.69	-0.61	-0.70	-0.60	-0.70	-0.62	-0.71	-0.63

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

Note: Sample sizes: low quality surveys = 31; high quality surveys = 33.

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Spearman correlation coefficients for the relationships between 10 of the 14 candidate linear growth metrics and GDP were of lower magnitude among high (versus low) quality surveys using the 3Q index, for at least two of the flagging approaches (Table 11a). However, the average magnitude of these differences were less than 0.10 for every metric when averaged across flagging approaches. Correlations for predicted HAZ at birth were inconsistent in directionality between low- and high-quality surveys, but overall were very close to the null.

When data quality was defined using the 6Q index, the pattern of the Spearman correlation strengths was less consistent, as almost half of the metrics had stronger correlations among higher quality surveys for at least two flagging approaches (Table 11b). The average differences between low- and high-quality surveys were generally less than 0.10 across flagging approaches, with the exception of predicted HAZ at 5 years and all HAZ and HAD slope metrics (with average differences ranging from 0.10 to 0.22). Similar to the correlations with U5M, a change in directionality was observed between low- and high-quality strata for predicted HAZ at birth when no flagging or less restrictive flagging were used; however, overall, the correlations were closer to the null relative to other metrics, and differences were very small.

Table 11a **Spearman rank correlation coefficients for the relationships between linear growth metrics and gross domestic product at high and low data quality (defined as above or below the 50th percentile of the restricted [3Q] anthropometric data quality index) in the most recent Demographic and Health Survey from 64 countries**

	No flagging[1]		Less restrictive flagging[2]		WHO flagging[3]		SMART flagging[4]	
Height-for-age metric	**Low**	**High**	**Low**	**High**	**Low**	**High**	**Low**	**High**
Stunting prevalence	-0.57	-0.49	-0.57	-0.49	-0.58	-0.51	-0.55	-0.53
Mean HAZ	0.52	0.46	0.52	0.47	0.52	0.50	0.53	0.52
Median HAZ	0.53	0.52	0.52	0.52	0.52	0.53	0.54	0.53
25th percentile HAZ	0.57	0.49	0.57	0.50	0.57	0.53	0.54	0.53
Winsorized mean HAZ	0.50	0.46	0.51	0.47	0.50	0.50	0.53	0.52
Robust Huber HAZ	0.54	0.50	0.53	0.50	0.54	0.51	0.54	0.52
Predicted HAZ at birth	-0.04	0.02	0.00	0.06	-0.08	0.09	0.16	0.21
Predicted HAZ at 2 years	0.53	0.45	0.54	0.45	0.53	0.47	0.54	0.53
Predicted HAZ at 3 years	0.52	0.47	0.54	0.46	0.53	0.52	0.54	0.53
Predicted HAZ at 5 years	0.54	0.49	0.53	0.48	0.51	0.53	0.51	0.57
HAZ slope 0-2 years	0.38	0.44	0.46	0.43	0.51	0.46	0.44	0.43
HAZ slope 0-3 years	0.45	0.48	0.47	0.50	0.45	0.52	0.43	0.52
HAD slope 0-2 years	0.43	0.48	0.42	0.49	0.51	0.51	0.54	0.51
HAD slope 0-3 years	0.47	0.56	0.50	0.55	0.49	0.57	0.54	0.57

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization
Note: Sample sizes: low quality surveys = 26; high quality surveys = 38.
[1] No exclusion of outlying values.
[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.
[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.
[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Table 11b **Spearman rank correlation coefficients for the relationships between linear growth metrics and gross domestic product at high and low data quality (defined as above or below the 50th percentile of the extended composite [6Q] anthropometric data quality index) in the most recent Demographic and Health Survey from 64 countries**

	No flagging[1]		Less restrictive flagging[2]		WHO flagging[3]		SMART flagging[4]	
Height-for-age metric	**Low**	**High**	**Low**	**High**	**Low**	**High**	**Low**	**High**
Stunting prevalence	-0.59	-0.56	-0.59	-0.56	-0.59	-0.57	-0.57	-0.57
Mean HAZ	0.54	0.45	0.54	0.46	0.55	0.48	0.54	0.52
Median HAZ	0.54	0.52	0.53	0.52	0.54	0.52	0.55	0.52
25th percentile HAZ	0.54	0.55	0.56	0.55	0.57	0.58	0.59	0.59
Winsorized mean HAZ	0.53	0.46	0.55	0.46	0.54	0.49	0.54	0.53
Robust Huber HAZ	0.55	0.50	0.55	0.50	0.56	0.52	0.54	0.53
Predicted HAZ at birth	0.09	-0.12	0.09	-0.03	0.01	0.00	0.17	0.23
Predicted HAZ at 2 years	0.44	0.54	0.50	0.52	0.49	0.54	0.51	0.56
Predicted HAZ at 3 years	0.52	0.50	0.54	0.49	0.56	0.54	0.55	0.54
Predicted HAZ at 5 years	0.61	0.42	0.60	0.42	0.59	0.44	0.57	0.48
HAZ slope 0-2 years	0.29	0.58	0.38	0.51	0.46	0.48	0.45	0.42
HAZ slope 0-3 years	0.32	0.66	0.41	0.57	0.42	0.56	0.45	0.54
HAD slope 0-2 years	0.38	0.61	0.43	0.55	0.48	0.54	0.53	0.51
HAD slope 0-3 years	0.39	0.71	0.44	0.65	0.45	0.65	0.49	0.62

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization
Note: Sample sizes: low quality surveys = 31; high quality surveys = 33.
[1] No exclusion of outlying values.
[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.
[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.
[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Spearman correlation analyses for maternal education using the 3Q index showed generally lower magnitudes of correlation among high (versus low) quality surveys, but average differences across flagging approaches were generally less than 0.10 (Table 12a). Predicted HAZ at birth was the only exception, as correlation strengths changed in directionality with the no flagging, less restrictive flagging, and WHO flagging approaches. Analyses using the 6Q index to define data quality strata demonstrated conflicting results with stronger correlations among high (versus low) quality surveys for most of the metrics (Table 12b).

Table 12a **Spearman rank correlation coefficients for the relationships between linear growth metrics and maternal education at high and low data quality (defined as above or below the 50th percentile of the restricted [3Q] anthropometric data quality index) in the most recent Demographic and Health Survey from 63 countries**

	No flagging[1]		Less restrictive flagging[2]		WHO flagging[3]		SMART flagging[4]	
Height-for-age metric	**Low**	**High**	**Low**	**High**	**Low**	**High**	**Low**	**High**
Stunting prevalence	-0.67	-0.57	-0.67	-0.57	-0.67	-0.59	-0.64	-0.59
Mean HAZ	0.60	0.54	0.63	0.53	0.64	0.56	0.60	0.57
Median HAZ	0.59	0.58	0.60	0.58	0.60	0.58	0.61	0.58
25th percentile HAZ	0.65	0.59	0.65	0.59	0.65	0.62	0.65	0.62
Winsorized mean HAZ	0.59	0.54	0.63	0.54	0.63	0.57	0.61	0.58
Robust Huber HAZ	0.61	0.56	0.62	0.56	0.63	0.57	0.62	0.58
Predicted HAZ at birth	-0.19	0.07	-0.13	0.13	-0.14	0.16	0.16	0.31
Predicted HAZ at 2 years	0.67	0.62	0.68	0.60	0.69	0.62	0.63	0.60
Predicted HAZ at 3 years	0.69	0.58	0.71	0.58	0.71	0.61	0.64	0.59
Predicted HAZ at 5 years	0.69	0.46	0.70	0.45	0.68	0.48	0.62	0.57
HAZ slope 0-2 years	0.59	0.42	0.61	0.41	0.64	0.44	0.45	0.37
HAZ slope 0-3 years	0.64	0.38	0.65	0.36	0.68	0.37	0.59	0.38
HAD slope 0-2 years	0.61	0.51	0.60	0.52	0.61	0.52	0.57	0.48
HAD slope 0-3 years	0.69	0.54	0.71	0.52	0.72	0.51	0.70	0.55

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

Note: Sample sizes: low quality surveys = 25; high quality surveys = 38.

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

Table 12b Spearman rank correlation coefficients for the relationships between linear growth metrics and maternal education at high and low data quality (defined as above or below the 50th percentile of the extended composite [6Q] anthropometric data quality index) in the most recent Demographic and Health Survey from 63 countries

	No flagging[1]		Less restrictive flagging[2]		WHO flagging[3]		SMART flagging[4]	
Height-for-age metric	**Low**	**High**	**Low**	**High**	**Low**	**High**	**Low**	**High**
Stunting prevalence	-0.59	-0.62	-0.59	-0.62	-0.58	-0.64	-0.56	-0.64
Mean HAZ	0.54	0.54	0.53	0.54	0.55	0.58	0.52	0.59
Median HAZ	0.51	0.60	0.51	0.60	0.50	0.61	0.52	0.61
25th percentile HAZ	0.57	0.62	0.56	0.63	0.58	0.66	0.58	0.67
Winsorized mean HAZ	0.53	0.54	0.55	0.54	0.56	0.57	0.52	0.60
Robust Huber HAZ	0.53	0.57	0.53	0.57	0.55	0.59	0.53	0.61
Predicted HAZ at birth	-0.04	0.00	-0.01	0.06	-0.06	0.13	0.12	0.30
Predicted HAZ at 2 years	0.63	0.63	0.62	0.60	0.63	0.61	0.57	0.62
Predicted HAZ at 3 years	0.62	0.58	0.62	0.58	0.64	0.63	0.57	0.62
Predicted HAZ at 5 years	0.60	0.46	0.62	0.46	0.59	0.50	0.54	0.57
HAZ slope 0-2 years	0.43	0.52	0.45	0.48	0.54	0.47	0.38	0.43
HAZ slope 0-3 years	0.47	0.44	0.53	0.39	0.58	0.38	0.52	0.42
HAD slope 0-2 years	0.49	0.61	0.50	0.58	0.54	0.56	0.46	0.54
HAD slope 0-3 years	0.58	0.59	0.59	0.55	0.60	0.52	0.59	0.58

Note: HAD = height-for-age difference; HAZ = height-for-age z-score; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization

Note: Sample sizes: low quality surveys = 30; high quality surveys = 33.

[1] No exclusion of outlying values.

[2] Excludes HAZ values <-9 SDs and >+9 SDs from the age/sex-specific WHO standard median.

[3] Excludes HAZ values <-6 SDs and >+6 SDs from the age/sex-specific WHO standard median.

[4] Excludes HAZ values <-3 SDs and >+3 SDs from the survey-specific sample mean.

4 DISCUSSION

The goal of this work was to explore alternative data-derived metrics of linear growth to quantify early childhood nutritional status at the population level. This working paper has provided a preliminary demonstration of (1) the application of a framework to empirically compare linear growth metrics based on the relative strengths of their associations with other population indicators, (2) the further development of two composite anthropometric data quality indices, and (3) the application of the data quality indices to assess the extent to which anthropometric data quality may influence the associations between linear growth metrics and population indicators. The present discussion focuses on the strengths and limitations of the methods, summarizes notable findings that have emerged thus far in the analyses, and describes future directions for this research. This working paper does not rank metrics or draw conclusions about one or a set of prioritized metrics for public health applications because the framework is being further developed and additional candidate metrics are being considered.

4.1 Alternative linear growth metrics vary in the strengths of their correlations with stunting prevalence

Given that stunting prevalence is the most widely used metric to depict childhood nutritional status of a population, we first examined the correlations of each alternative metric with stunting. We considered that the most promising metrics may be those with moderate (rather than strong) correlations with stunting, as metrics highly correlated with stunting would be unlikely to offer substantially different information than is conveyed by stunting. We observed strong correlations of stunting with descriptive metrics (i.e., mean, median, robust Huber mean, Winsorized mean, 25P HAZ). The particularly strong correlations between these metrics and stunting prevalence are consistent with the notion that higher stunting prevalence is a result of a downward shift of essentially the entire height distribution in many LMIC settings.[2] Conversely, the regression model-derived metrics (i.e., predicted HAZ values at discrete ages, HAZ and HAD slopes) did not correlate as strongly with stunting. The model-derived metrics reflect the change in the position of the HAZ (or HAD) distribution with age, since in LMICs, postnatal faltering is commonly observed as a decline in mean HAZ (or HAD) and an increase in stunting prevalence with age.[15] Anthropometric measurements, particularly recumbent length, tend to be more difficult to obtain with consistently high quality at younger ages than older ages. Age-related variations in the precision or accuracy of measurement may partly account for the particularly low correlation strength we observed for predicted HAZ at birth when compared with predicted HAZ at later ages. For each given metric, correlation coefficients for under-5 stunting were generally consistent across flagging approaches. However, the notable exception was predicted HAZ at birth, for which coefficients differed substantially between the no flagging and SMART flagging approaches.

4.2 Alternative linear growth metrics vary in the strengths of their correlations with conventional population indicators

Initial assessment of candidate alternative metrics was based on the relative strength of their correlations with selected population indicators known to be conceptually and empirically associated with child linear growth and nutritional status in LMICs.[5–11] Analyses completed thus far demonstrated that stunting tended to have stronger associations with the population indicators examined than with most other descriptive linear growth metrics; however, many of the correlation coefficients were within 0.05 of one another, indicating that other measures of central tendency do not necessarily offer an advantage over stunting because they essentially capture the same variations in height across surveys. Nonetheless, there were some metrics that tended to outperform stunting (i.e., had stronger correlations with population indicators), most of which were model-derived metrics, such as predicted HAZ at discrete ages and HAD-by-age slopes. In addition to stronger empirical correlations with population indicators, predicted HAZ at discrete ages and slopes may have conceptual advantages as alternatives to stunting as they represent the age-dependent process of linear growth faltering, for which variations among countries may reflect differences in the underlying causes of poor growth, even when overall under-5 stunting prevalences are similar. Moreover, the model-derived metrics appropriately convey the downward shift in the entire population height distribution in LMICs rather than highlight a subset of children below an arbitrary HAZ cutoff. In most cases, the correlations observed were only slightly affected by different approaches for flagging outliers; therefore, these analyses did not provide clear guidance with respect to the optimal approach for defining and handling outliers or implausible values in survey data.

The core assumption underlying these analyses was that anthropometry-based indicators that provide valid representations of the health of the population should have relatively robust associations with other indicators of population health and societal well-being. The intention was not to address questions of causality (e.g., whether higher GDP leads to increased HAZ), nor were we attempting to make comparisons across population indicators (e.g., whether HAZ is more strongly correlated with U5M than with GDP). The selection of population indicators was guided by conceptual considerations but also influenced by data availability in all countries in all years for which anthropometric surveys were conducted. For example, although we examined relationships with all proximal “underlying” population indicators of child nutritional status, we could not include child immunization status and female fertility rate in our analysis because of limitations in availability of historical multi-national data (Figure 3). With the exception of maternal education data, which were drawn from DHS surveys, we also used population indicator data that could be extracted from external sources (i.e., external to DHS datasets), so that errors in the data on height and date of birth would be unlikely to be correlated with errors in the external indicator datasets. We acknowledge that the limited array of indicators does not represent all domains of health and development for which height-based metrics may act as surrogates. For example, it may be advantageous to include indicators of early childhood cognitive development and school readiness, yet population-based measures of these domains were not readily available for all countries in all survey years.

4.3 A composite anthropometric data quality index was developed from a selected set of data quality indicators

We constructed indices of anthropometric data quality using a set of data quality indicators recommended for use by the WHO/UNICEF Anthropometric Data Quality Working Group.[13] The 6Q index, which included six indicators of anthropometric data quality (Table 4), was a more comprehensive index than the 3Q index for assessing variations in anthropometric data quality between surveys. However, the correlations between the 6Q index and candidate metrics of linear growth as well as population indicators was problematic, as use of the 6Q index interfered with a consistent interpretation of the models aimed at estimating interactions between alternative linear growth metrics and data quality. As such, for pragmatic purposes, we primarily used the 3Q index based on the three data quality indicators, which are independent of the distributional properties of HAZ: (1) the proportion of observations with complete date of birth,

(2) the proportion of observations with anthropometry measured, and (3) digit preference for height. Although the 3Q index reflects a restricted dimension of anthropometric data quality across surveys (i.e., issues related to non-response), it provided an alternative index for assessing relative anthropometric data quality given that collinearity with the growth metrics was a concern. This approach was necessary for the present study to achieve consistent inferences regarding the robustness of candidate metrics against variations in data quality. Nonetheless, because the 3Q index only incorporates data quality indicators related to non-response, we conducted sensitivity analyses using the 6Q index. In applications in which comparing metrics of child nutritional status is not the specific objective (as in this study), the more comprehensive 6Q index is recommended.[13]

4.4 Effects of variation in anthropometric data quality on the performance of linear growth metrics are inconclusive

Linear mixed effects models were used to empirically evaluate the relative robustness of the linear growth metrics against variations in anthropometric data quality. Robustness was based on whether the survey-specific 3Q index had a significant modifying effect on the association between a given metric and a population indicator. In contrast to the correlation analyses, which used only the most recent surveys from each country, the mixed effects models used data from all countries and survey years (with the exception of Benin, given that the 2011-12 survey from Benin was an influential outlier).

Outputs from the regression analyses indicated that most associations between candidate linear growth metrics and population indicators were not significantly modified by anthropometric data quality, as measured by either the 3Q index (primary analyses) or 6Q index (sensitivity analyses). There were also no observable trends across flagging approaches; however, we noted that six of the seven significant modifying effects observed in models with U5M, GDP, or maternal education as the outcome were observed when using SMART flagging. Although assessing the sensitivity of alternative linear growth metrics to variations in data quality would have provided an important reference point of comparison to stunting, these comparisons were ultimately inconclusive due to the following methodological limitations: (1) the 3Q index did not provide comprehensive relative scoring based on all important domains of anthropometric measurement quality in surveys; (2) associations of data quality with both the growth metrics and the population indicators were unavoidable (e.g., the quality of anthropometric data tends to be poor in countries with higher stunting prevalence and in countries with high child mortality rates), even when using the 3Q index; and (3) data quality had non-linear effects on the metric-indicator associations across the ranges of those variables. Therefore, the regression models did not allow us to confidently prioritize specific metrics on the basis of their robustness against variations in data quality, and we do not endorse the further use of modeling approaches using interactions between growth metrics and anthropometric data quality to draw conclusions about the relative performance of candidate linear growth metrics. Nonetheless, data quality indices such as the 6Q index, which more comprehensively captures multiple domains of anthropometric data quality, provide a useful measure for evaluating data quality in multi-survey studies of child nutritional status.

Given the inconclusive results from the regression models, we used Spearman rank correlation analyses to descriptively assess the robustness of the linear growth metrics against variations in anthropometric data quality (defined using either the 3Q or 6Q index). This approach did not overcome the limitations of the regression modeling but provided a more straightforward method of assessing whether the correlations differed among surveys classified as having low versus high anthropometric data quality. In general, we saw only minor differences in the strength of correlations between low- and high-quality surveys. Most correlation strengths were attenuated in high (versus low) quality surveys for the 3Q index. Predicted HAZ at birth tended to be the most unstable metric in terms of variations in the strength and direction of correlations from low to high data quality strata. Similar to the main analysis, the model-derived metrics (particularly HAD slope 0-3 years) tended to outperform the descriptive statistics in terms of strength of correlations among both low and high-quality surveys, and this was fairly consistent across three population indicators (i.e., U5M, GDP, maternal education) and all flagging approaches.

4.5 Strengths and limitations

This preliminary study demonstrates a new framework for evaluating candidate metrics of linear growth as indicators of child nutritional status, including the application of a composite anthropometric data quality index to assess the relative robustness of the candidate metrics against variations in survey quality.

As described in section 4.4, we faced difficulties with using the 3Q index, particularly when incorporating it as a covariate in the regression models for assessing relative performance of candidate linear growth metrics. Another potential limitation of this work is the use of only DHS data and not other common sources of child health survey data, such as Multiple Indicator Cluster Surveys.

5 CONCLUSION AND POLICY IMPLICATIONS

Although we cannot yet recommend the adoption of one or more alternative linear growth metrics, the results presented here demonstrate that it is possible to derive alternative metrics that perform at least as well as stunting with respect to correlations with population indicators using the same child height datasets. In the next phase of this research, we will refine the methodological approach for identifying and selecting alternative growth metrics, including an expanded list of model-derived candidate metrics. Considerations affecting the selection of alternative metrics will include the empirical factors assessed here (i.e., correlation with other population indicators), the conceptual properties of each metric (e.g., age slopes indicate the process, not just the result, of growth faltering), the interpretability of the metrics by target audiences (e.g., policymakers, the public), and the potential for alternative metrics to complement or strengthen the interpretation of population-level stunting prevalence.

REFERENCES

1. USAID. 2014. *USAID Multi-sectoral Nutrition Strategy Takes Aim at Stunting.* https://www.usaid.gov/global-waters/november-2014/nutrition-strategy-takes-aim-at-stunting

2. Roth, D. E., A. Krishna, M. Leung, J. Shi, D. G. Bassani, and A. J. D. Barros. 2017. "Early Childhood Linear Growth Faltering in Low-Income and Middle-Income Countries as a Whole-Population Condition: Analysis of 179 Demographic and Health Surveys from 64 Countries (1993–2015)." *Lancet Glob Health* 5 (12): e1249–57. http://dx.doi.org/10.1016/S2214-109X(17)30418-7

3. Assaf, S., M. T. Kothari, and T. Pullum. 2015. *An Assessment of the Quality of DHS Anthropometric Data, 2005-2014.* DHS Methodological Reports No. 16. Rockville, Maryland, USA: ICF International. http://dhsprogram.com/pubs/pdf/MR16/MR16.pdf

4. Corsi, D. J., J. M. Perkins, and S. V. Subramanian. 2017. "Child Anthropometry Data Quality from Demographic and Health Surveys, Multiple Indicator Cluster Surveys, and National Nutrition Surveys in the West Central Africa region: Are We Comparing Apples and Oranges?" *Glob Health Action* 10 (1): 1328185.

5. Pelletier, D. L., E. A. Frongillo, and J. P. Habicht. 1993. "Epidemiologic Evidence for a Potentiating Effect of Malnutrition on Child Mortality. *Am J Public Health* 83 (8): 1130–3.

6. Olofin, I., C. M. McDonald, M. Ezzati, S. Flaxman, R. E. Black, W. W. Fawzi, et al. 2013. "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 (5): e64636.

7. Mejia-Guevara, I., D. J. Corsi, J. M. Perkins, R. Kim, and S. V. Subramanian. 2018. "Variation in Anthropometric Status and Growth Failure in Low- and Middle-Income Countries." *Pediatrics* 141(3): e29172183.

8. Frongillo, E. A., M. de Onis, and K. M. P. Hanson. 1997. "Community and International Nutrition Socioeconomic and Demographic Factors Are Associated with Worldwide Patterns of Stunting and Wasting of Children." *J Nutr* 127 (12): 2302–9.

9. Milman, A., E. A. Frongillo, M. de Onis, and J.-Y. Hwang. 2005. "Differential Improvement Among Countries in Child Stunting Is Associated with Long-Term Development and Specific Interventions." *J Nutr* 135 (6):1415–22.

10. Kim, R., I. Mejía-Guevara, D. J. Corsi, V. M. Aguayo, and S. V. Subramanian. 2017. "Relative Importance of 13 Correlates of Child Stunting in South Asia : Insights from Nationally Representative Data from Afghanistan, Bangladesh, India, Nepal, and Pakistan." *Soc Sci Med* 187:144–54.

11. Ozaltin, E., K. Hill, and S. Subramanian. 2010. “Association of Maternal Stature with Offspring Mortality, Underweight, and Stunting in Low- to Middle-Income Countries.” *JAMA* 303 (15): 1507–16.

12. WHO and UNICEF. 2019. *Recommendations for Data Collection, Analysis and Reporting on Anthropometric Indicators in Children Under 5 Years Old.* Geneva, Switzerland: WHO and UNICEF.

13. Perumal, N., S. Namaste, H. Qamar, A. Aimone, D. G. Bassani, and D. E. Roth. 2020. “Anthropometric Data Quality Assessment in Multisurvey Studies of Child Growth.” *Am J Clin Nutr* 112 (Suppl 2): 806S–15S.

14. Larsen, A.F., D. Headey, and W. A. Masters. 2019. “Misreporting Month of Birth: Diagnosis and Implications for Research on Nutrition and Early Childhood in Developing Countries.” *Demography* 56 (2): 707–28.

15. Victora, C.G., M. de Onis, P. C. Hallal, M. Blössner, and R. Shrimpton. 2010. “Worldwide Timing of Growth Faltering: Revisiting Implications for Interventions. *Pediatrics* 125 (3): e473–80.

APPENDIX

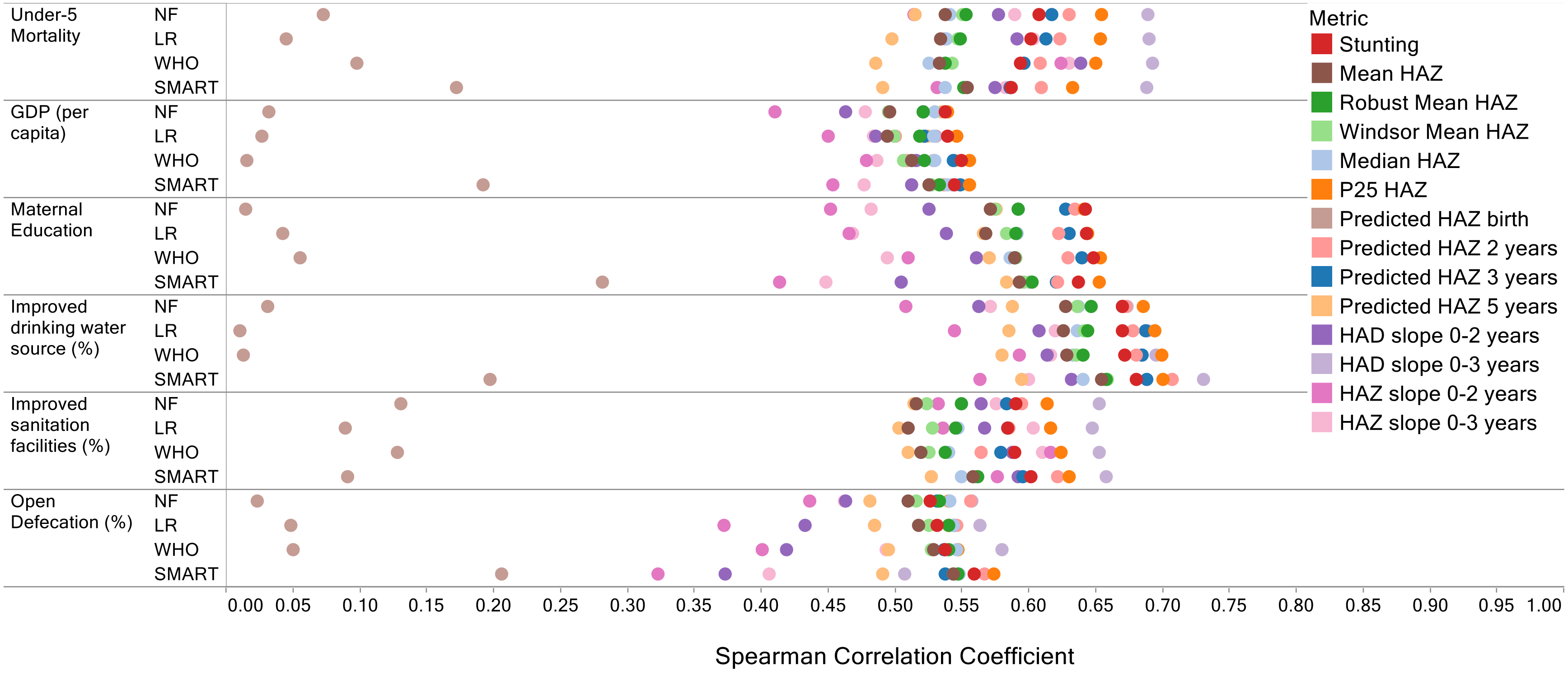
Table A1 Included data from Demographic and Health Survey and external population indicators

		Year of available external data					
Country	Survey year	Under-5 mortality	Gross domestic product	Maternal education	Improved drinking water source	Improved sanitation facilities	Open defecation
Albania	2009	2009	2009	2009	2009	2009	2009
Albania	2017	2017	2017	2017			
Angola	2015	2015	2015	2015	2015	2015	2015
Armenia	2005	2005	2005	2005	2005	2005	2005
Armenia	2010	2010	2010	2010	2010	2010	2010
Armenia	2016	2016	2016	2016			
Azerbaijan	2006	2006	2006	2006	2006	2006	2006
Bangladesh	2004	2004	2004	2004	2004	2004	2004
Bangladesh	2007	2007	2007	2007	2007	2007	2007
Bangladesh	2011	2011	2011	2011	2011	2011	2011
Bangladesh	2014	2014	2014	2014	2014	2014	2014
Benin	2001	2001	2001	2001	2001	2001	2001
Benin	2006	2006	2006	2006	2006	2006	2006
Benin	2012	2012	2012	2012	2012	2012	2012
Bolivia	2003	2003	2003	2003	2003	2003	2003
Bolivia	2008	2008	2008	2008	2008	2008	2008
Burkina Faso	2003	2003	2003	2003	2003	2003	2003
Burkina Faso	2010	2010	2010	2010	2010	2010	2010
Burundi	2010	2010	2010	2010	2010	2010	2010
Burundi	2016	2016	2016	2016			
Cambodia	2000	2000	2000	2000	2000	2000	2000
Cambodia	2005	2005	2005	2005	2005	2005	2005
Cambodia	2010	2010	2010	2010	2010	2010	2010
Cambodia	2014	2014	2014	2014	2014	2014	2014
Cameroon	2004	2004	2004	2004	2004	2004	2004
Cameroon	2011	2011	2011	2011	2011	2011	2011
Chad	2015	2015	2015	2015	2015	2015	2015
Colombia	2010	2010	2010	2010	2010	2010	2010
Comoros	2012	2012	2012	2012	2012	2012	2012
CongoBZ	2005	2005	2005	2005	2005	2005	2005
CongoBZ	2011	2011	2011	2011	2011	2011	2011
CongoDR	2007	2007	2007	2007	2007	2007	2007
CongoDR	2013	2013	2013	2013	2013	2013	2013
Côte d'Ivoire	2012	2012	2012	2012	2012	2012	2012
Dominican Republic	2002	2002	2002	2002	2002	2002	2002
Dominican Republic	2007	2007	2007	2007	2007	2007	2007
Dominican Republic	2013	2013	2013	2013	2013	2013	2013
Egypt	2000	2000	2000	2000	2000	2000	2000
Egypt	2003	2003	2003	2003	2003	2003	2003
Egypt	2005	2005	2005	2005	2005	2005	2005
Egypt	2008	2008	2008	2008	2008	2008	2008
Egypt	2014	2014	2014	2014	2014	2014	2014
Eritrea	2002	2002	2002	2002	2002	2002	2002
Ethiopia	2000	2000	2000	2000	2000	2000	2000
Ethiopia	2005	2005	2005	2005	2005	2005	2005
Ethiopia	2011	2011	2011	2011	2011	2011	2011
Ethiopia	2016	2016	2016	2016			
Gabon	2012	2012	2012	2012	2012	2012	2012
Gambia	2013	2013	2013	2013	2013	2013	2013
Ghana	2008	2008	2008	2008	2008	2008	2008
Ghana	2014	2014	2014	2014	2014	2014	2014
Guatemala	2015	2015	2015	2015	2015	2015	2015
Guinea	2012	2012	2012	2012	2012	2012	2012
Guyana	2009	2009	2009	2009	2009	2009	2009
Haiti	2000	2000	2000	2000	2000	2000	2000
Haiti	2006	2006	2006	2006	2006	2006	2006
Haiti	2012	2012	2012	2012	2012	2012	2012

		Year of available external data					
Country	**Survey year**	**Under-5 mortality**	**Gross domestic product**	**Maternal education**	**Improved drinking water source**	**Improved sanitation facilities**	**Open defecation**
Haiti	2017	2017	2017	2017			
Honduras	2006	2006	2006	2006	2006	2006	2006
Honduras	2012	2012	2012	2012	2012	2012	2012
India	2006	2006	2006	2006	2006	2006	2006
India	2015	2015	2015	2015	2015	2015	2015
Jordan	2002	2002	2002	2002	2002	2002	2002
Jordan	2007	2007	2007	2007	2007	2007	2007
Jordan	2009	2009	2009	2009	2009	2009	2009
Jordan	2012	2012	2012	2012	2012	2012	2012
Kenya	2003	2003	2003	2003	2003	2003	2003
Kenya	2009	2009	2009	2009	2009	2009	2009
Kenya	2014	2014	2014	2014	2014	2014	2014
Kyrgyzstan	2012	2012	2012	2012	2012	2012	2012
Lesotho	2009	2009	2009	2009	2009	2009	2009
Lesotho	2014	2014	2014	2014	2014	2014	2014
Liberia	2007	2007	2007	2007	2007	2007	2007
Liberia	2013	2013	2013	2013	2013	2013	2013
Madagascar	2004	2004	2004	2004	2004	2004	2004
Madagascar	2009	2009	2009	2009	2009	2009	2009
Malawi	2000	2000	2000	2000	2000	2000	2000
Malawi	2004	2004	2004	2004	2004	2004	2004
Malawi	2010	2010	2010	2010	2010	2010	2010
Malawi	2015	2015	2015	2015	2015	2015	2015
Maldives	2009	2009	2009	2009	2009	2009	2009
Mali	2001	2001	2001	2001	2001	2001	2001
Mali	2006	2006	2006	2006	2006	2006	2006
Mali	2012	2012	2012	2012	2012	2012	2012
Moldova	2005	2005	2005	2005	2005	2005	2005
Morocco	2003	2003	2003	2003	2003	2003	2003
Mozambique	2003	2003	2003	2003	2003	2003	2003
Mozambique	2011	2011	2011	2011	2011	2011	2011
Myanmar	2016	2016	2016	2016			
Namibia	2000	2000	2000	2000	2000	2000	2000
Namibia	2007	2007	2007	2007	2007	2007	2007
Namibia	2013	2013	2013	2013	2013	2013	2013
Nepal	2001	2001	2001	2001	2001	2001	2001
Nepal	2006	2006	2006	2006	2006	2006	2006
Nepal	2011	2011	2011	2011	2011	2011	2011
Nepal	2016	2016	2016	2016			
Nicaragua	2001	2001	2001	2001	2001	2001	2001
Niger	2006	2006	2006	2006	2006	2006	2006
Niger	2012	2012	2012	2012	2012	2012	2012
Nigeria	2003	2003	2003	2003	2003	2003	2003
Nigeria	2008	2008	2008	2008	2008	2008	2008
Nigeria	2013	2013	2013	2013	2013	2013	2013
Pakistan	2012	2012	2012	2012	2012	2012	2012
Pakistan	2018	2018	2018	2018			
Peru	2005	2005	2005	2005	2005	2005	2005
Peru	2008	2008	2008	2008	2008	2008	2008
Peru	2009	2009	2009	2009	2009	2009	2009
Peru	2010	2010	2010	2010	2010	2010	2010
Peru	2011	2011	2011	2011	2011	2011	2011
Peru	2012	2012	2012	2012	2012	2012	2012
Rwanda	2000	2000	2000	2000	2000	2000	2000
Rwanda	2005	2005	2005	2005	2005	2005	2005
Rwanda	2010	2010	2010	2010	2010	2010	2010
Rwanda	2015	2015	2015	2015	2015	2015	2015
Sao Tome and Principe	2008	2008	2008	2008	2008	2008	2008
Senegal	2005	2005	2005	2005	2005	2005	2005
Senegal	2010	2010	2010	2010	2010	2010	2010
Senegal	2013	2013	2013	2013	2013	2013	2013
Senegal	2014	2014	2014	2014	2014	2014	2014
Senegal	2015	2015	2015	2015	2015	2015	2015

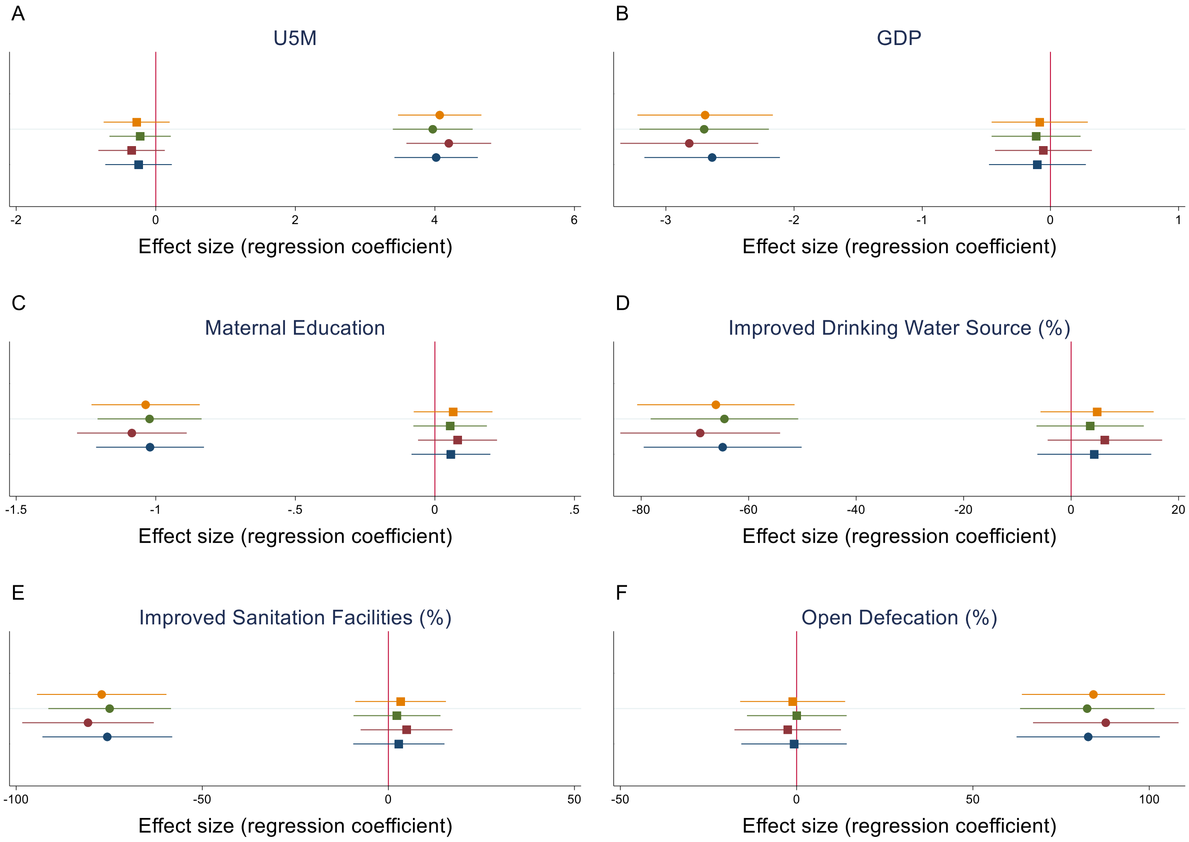
		Year of available external data					
Country	**Survey year**	**Under-5 mortality**	**Gross domestic product**	**Maternal education**	**Improved drinking water source**	**Improved sanitation facilities**	**Open defecation**
Senegal	2016	2016	2016	2016			
Senegal	2017	2017	2017	2017			
Sierra Leone	2008	2008	2008	2008	2008	2008	2008
Sierra Leone	2013	2013	2013	2013	2013	2013	2013
South Africa	2016	2016	2016	2016			
Swaziland	2006	2006	2006	2006			
Tajikistan	2012	2012	2012	2012	2012	2012	2012
Tajikistan	2017	2017	2017	2017			
Tanzania	2004	2004	2004	2004	2004	2004	2004
Tanzania	2010	2010	2010	2010	2010	2010	2010
Tanzania	2015	2015	2015	2015	2015	2015	2015
Timor-Leste	2009	2009	2009	2009	2009	2009	2009
Timor-Leste	2016	2016	2016	2016			
Togo	2014	2014	2014	2014	2014	2014	2014
Uganda	2000	2000	2000	2000	2000	2000	2000
Uganda	2006	2006	2006	2006	2006	2006	2006
Uganda	2011	2011	2011	2011	2011	2011	2011
Uganda	2016	2016	2016	2016			
Yemen	2013	2013	2013		2013	2013	2013
Zambia	2002	2002	2002	2002	2002	2002	2002
Zambia	2007	2007	2007	2007	2007	2007	2007
Zambia	2013	2013	2013	2013	2013	2013	2013
Zimbabwe	2005	2005	2005	2005	2005	2005	2005
Zimbabwe	2010	2010	2010	2010	2010	2010	2010
Zimbabwe	2015	2015	2015	2015	2015	2015	2015

Figure A1 Absolute values of Spearman rank correlation coefficients of linear growth metrics and six population indicators using data from the most recent Demographic and Health Surveys in 64 low- and middle-income countries, applying four flagging approaches for identifying and removing extreme height-for-age z-score values



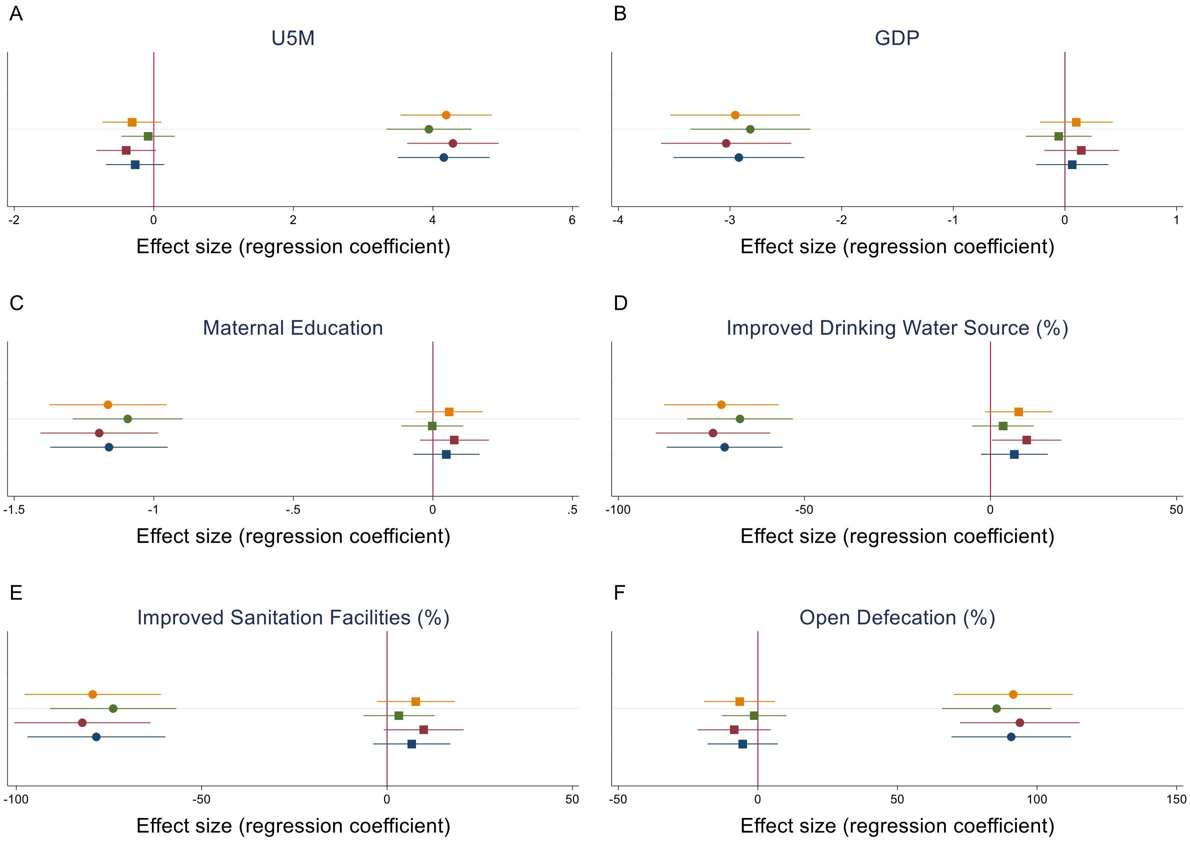
Note: GDP = gross domestic product; NF = no flagging; LR = less restrictive; SMART = standardized monitoring and assessment of relief and transitions; WHO = World Health Organization; Due to missing data the sample sizes were lower for maternal education (N=63 countries) and for improved drinking water source, improved sanitation facilities, and open defecation (N=50 countries).

Figure A2 **Regression coefficients from linear mixed effects models for the main effect of *stunting prevalence* and the modifying effect of the restricted (3Q) anthropometric data quality index for six different population indicators**



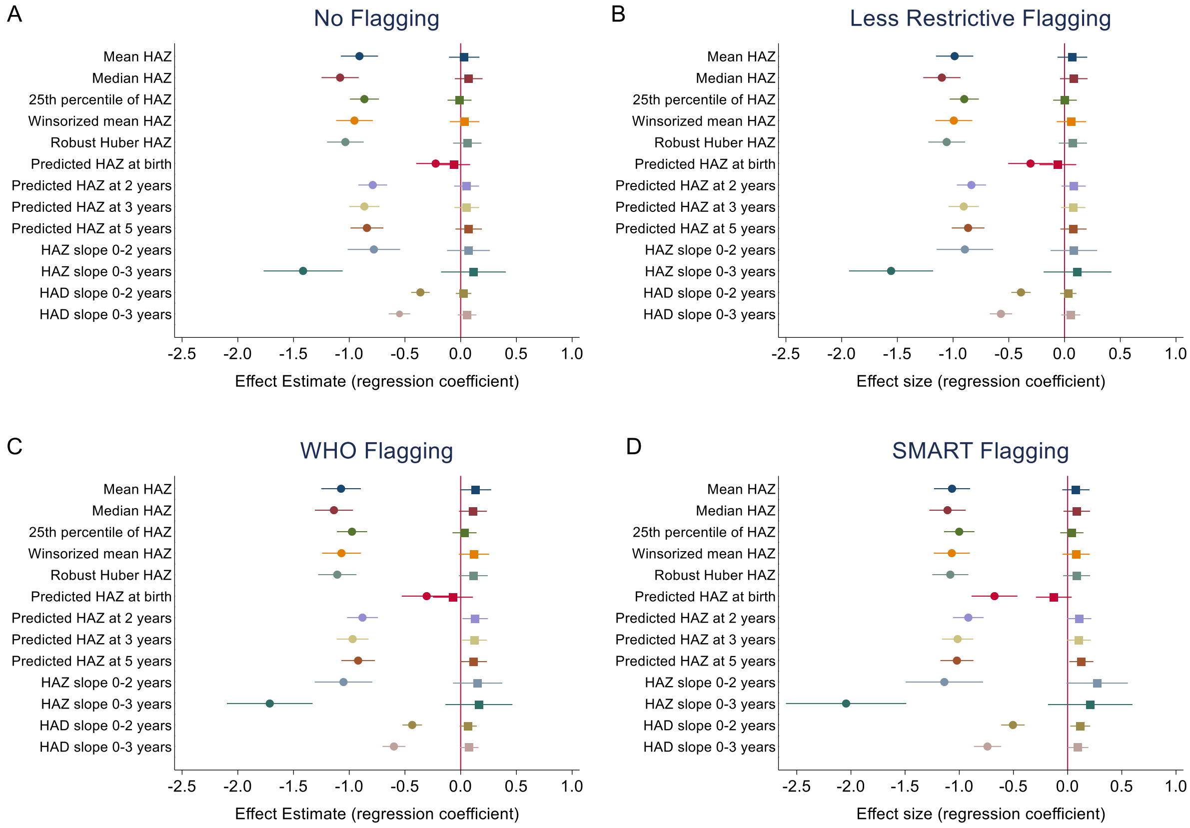
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the restricted 3Q index. (A) under-5 mortality, (B) gross domestic product, (C) maternal education, (D) % improved drinking water source, (E) % improved sanitation facilities, and (F) % open defecation. The main effects are from the models that include the interaction terms with the quality index; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5). Analyses are based on data from: 144 Demographic and Health Surveys from 64 low- and middle-income countries where U5M and GDP are the outcome; 143 Demographic and Health Surveys from 63 low- and middle-income countries where maternal education is the outcome; and, 129 Demographic and Health Surveys from 61 low- and middle-income countries where improved drinking water source, improved sanitation facilities and open defecation are the outcome. Results are shown for four flagging approaches for identifying and removing implausible HAZ values: no flagging (blue), less restrictive flagging (orange), WHO flagging (red), and SMART flagging (green).

Figure A3 **Regression coefficients from linear mixed effects models for the main effect of *stunting prevalence* and the modifying effect of the extended (6Q) anthropometric data quality index for six different population indicators**



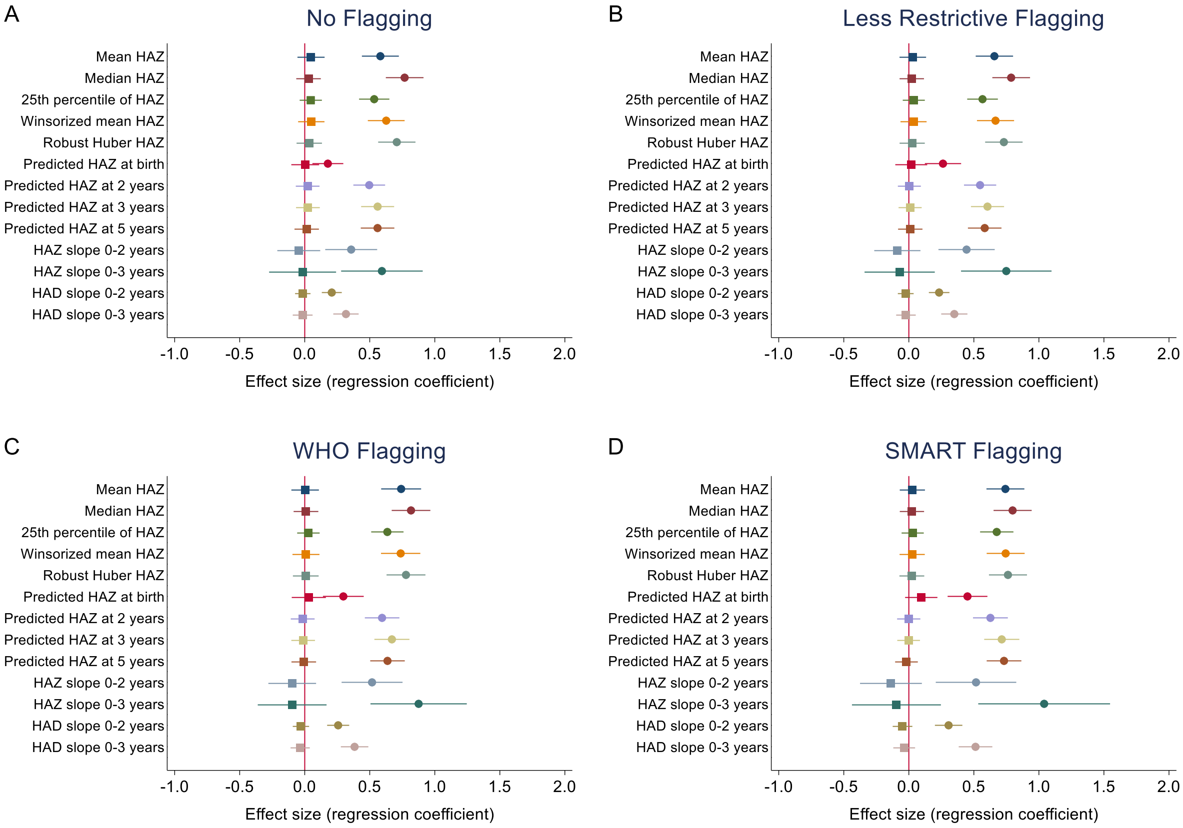
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the extended 6Q index anthropometric data quality index (squares). (A) under-5 mortality, (B) gross domestic product, (C) maternal education, (D) % improved drinking water source, (E) % improved sanitation facilities, and (F) % open defecation. For the interaction term, point estimates to the right (under-5 mortality) or to the left (all other indicators) of the null indicate that as the data quality score increases (i.e., quality improves), the association between stunting and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5). Analyses are based on data from: 144 Demographic and Health Surveys from 64 low- and middle-income countries where U5M and GDP are the outcome; 143 Demographic and Health Surveys from 63 low- and middle-income countries where maternal education is the outcome; and, 129 Demographic and Health Surveys from 61 low- and middle-income countries where improved drinking water source, improved sanitation facilities and open defecation are the outcome. Results are shown for four flagging approaches for identifying and removing implausible HAZ values: no flagging (blue), less restrictive flagging (orange), WHO flagging (red), and SMART flagging (green).

Figure A4 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *under-5 mortality rate* and the modifying effect of the restricted (3Q) anthropometric data quality index using data from 144 Demographic and Health Surveys from 64 low- and middle-income countries and four flagging approaches for identifying and removing implausible HAZ values**



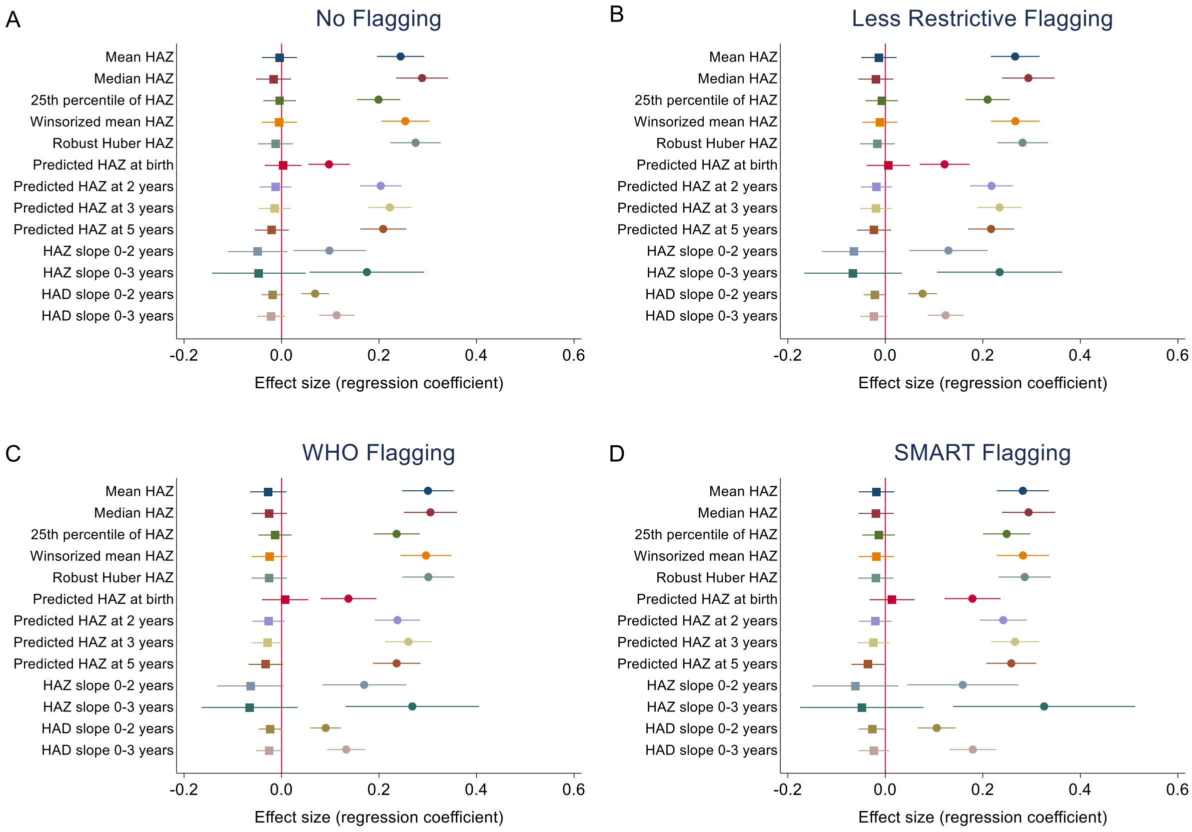
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 3Q index. For the interaction term, point estimates to the left of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval was interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A5 Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *gross domestic product* and the modifying effect of the restricted (3Q) anthropometric data quality index (squares) using data from 144 Demographic and Health Surveys from 64 low- and middle-income countries and four flagging approaches for identifying and removing implausible HAZ values



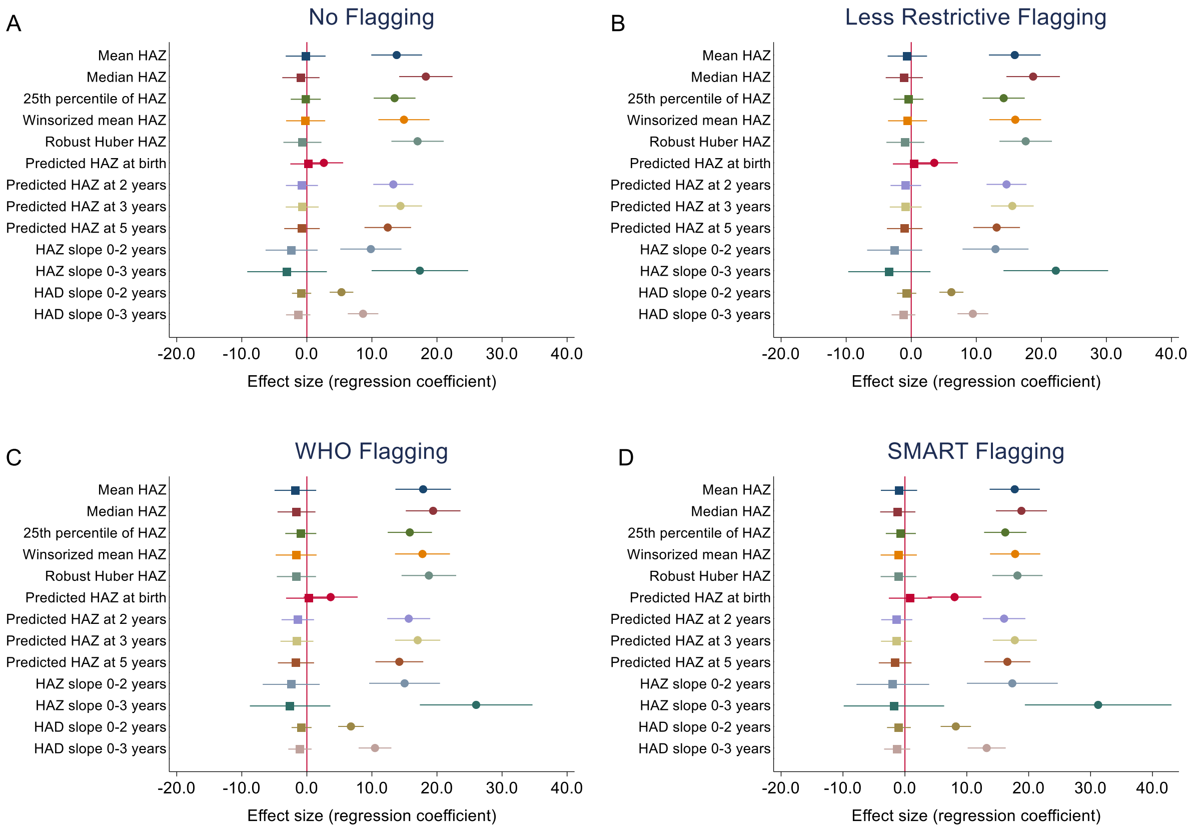
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 3Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and gross domestic product strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A6 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *maternal education* and the modifying effect of the restricted (3Q) anthropometric data quality index using data from 143 Demographic and Health Surveys from 63 low- and middle-income countries and four flagging approaches for identifying and removing implausible HAZ values**



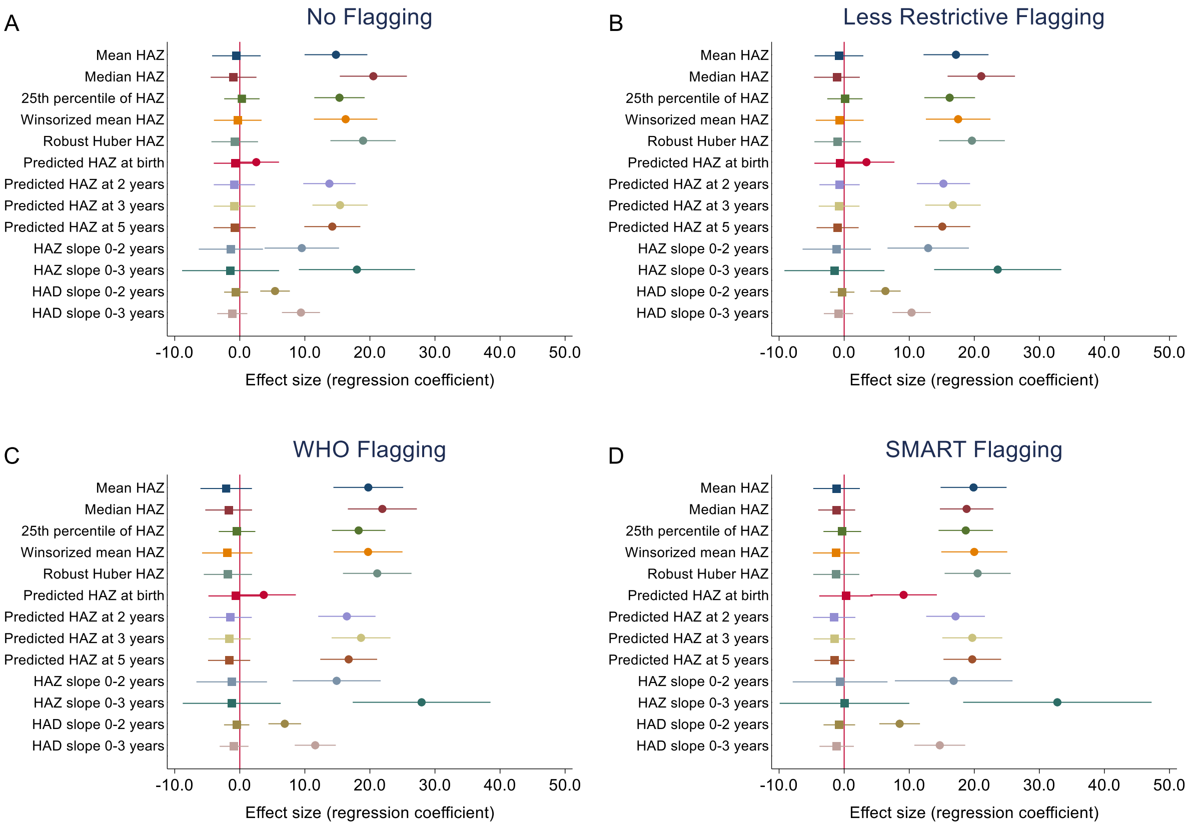
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 3Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and maternal education strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A7 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *prevalence of access to improved sanitation facilities* and the modifying effect of the extended (6Q) anthropometric data quality index (squares) using data from 129 Demographic and Health Surveys from 61 low- and middle-income countries and four flagging approaches for identifying and removing implausible HAZ values**



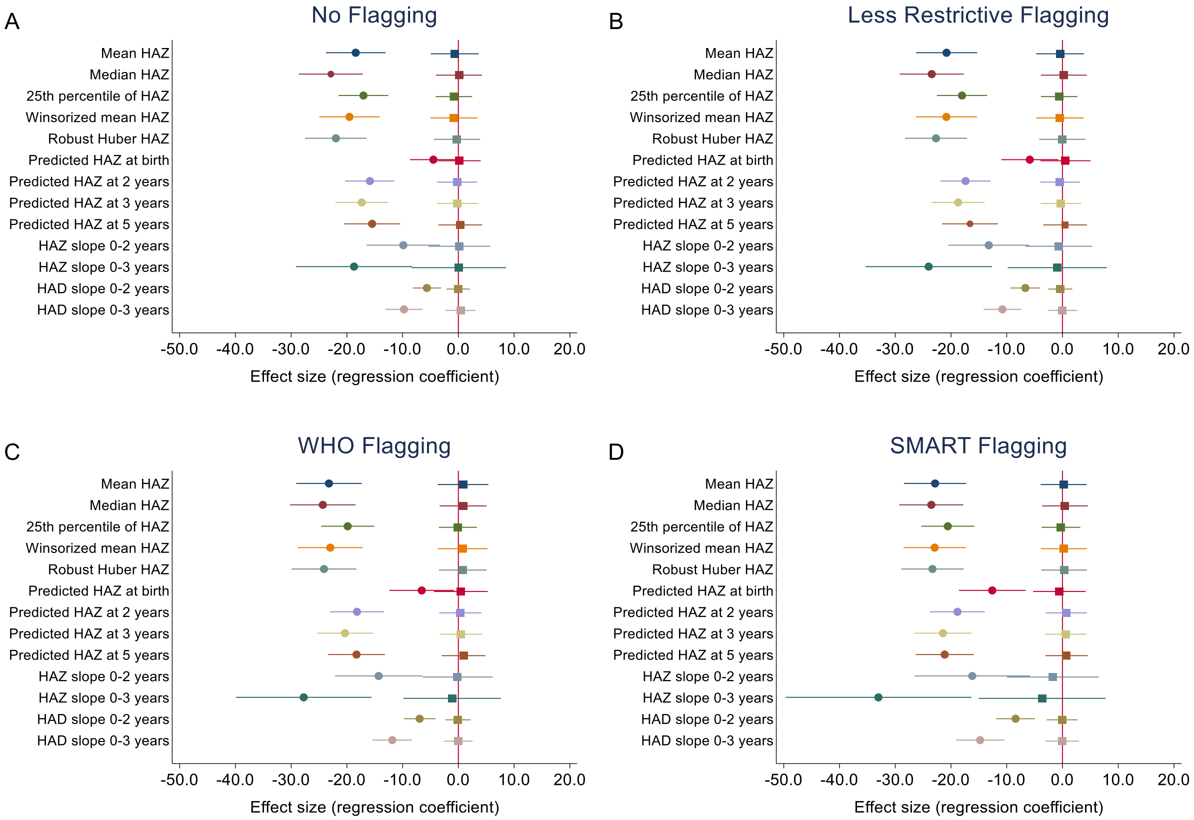
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 3Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the prevalence of access to improved drinking water source strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A8 **Regression coefficients from linear mixed effects models for the main effect (with 95% confidence intervals) of alternative HAZ measures of location and model-derived slopes on *prevalence of access to improved sanitation facilities* and the modifying effect of restricted (3Q) anthropometric data quality index (squares) using data from 129 Demographic and Health Surveys from 61 low- and middle-income countries and four flagging approaches for identifying and removing implausible HAZ values**



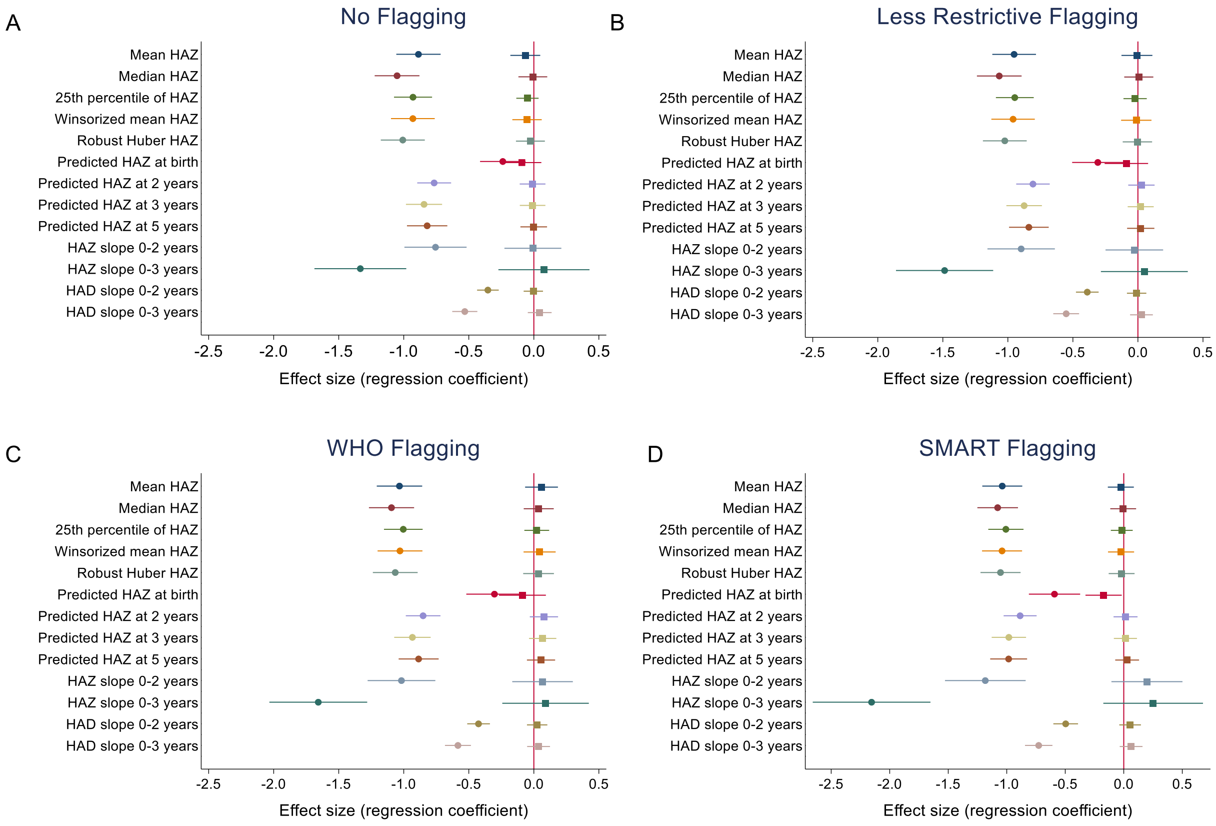
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 3Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the prevalence of access to improved sanitation facilities strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A9 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *prevalence of open defecation* and the modifying effect of the restricted (3Q) anthropometric data quality index using data from 129 Demographic and Health Surveys from 61 low- and middle-income countries and four flagging approaches for identifying and removing implausible HAZ values**



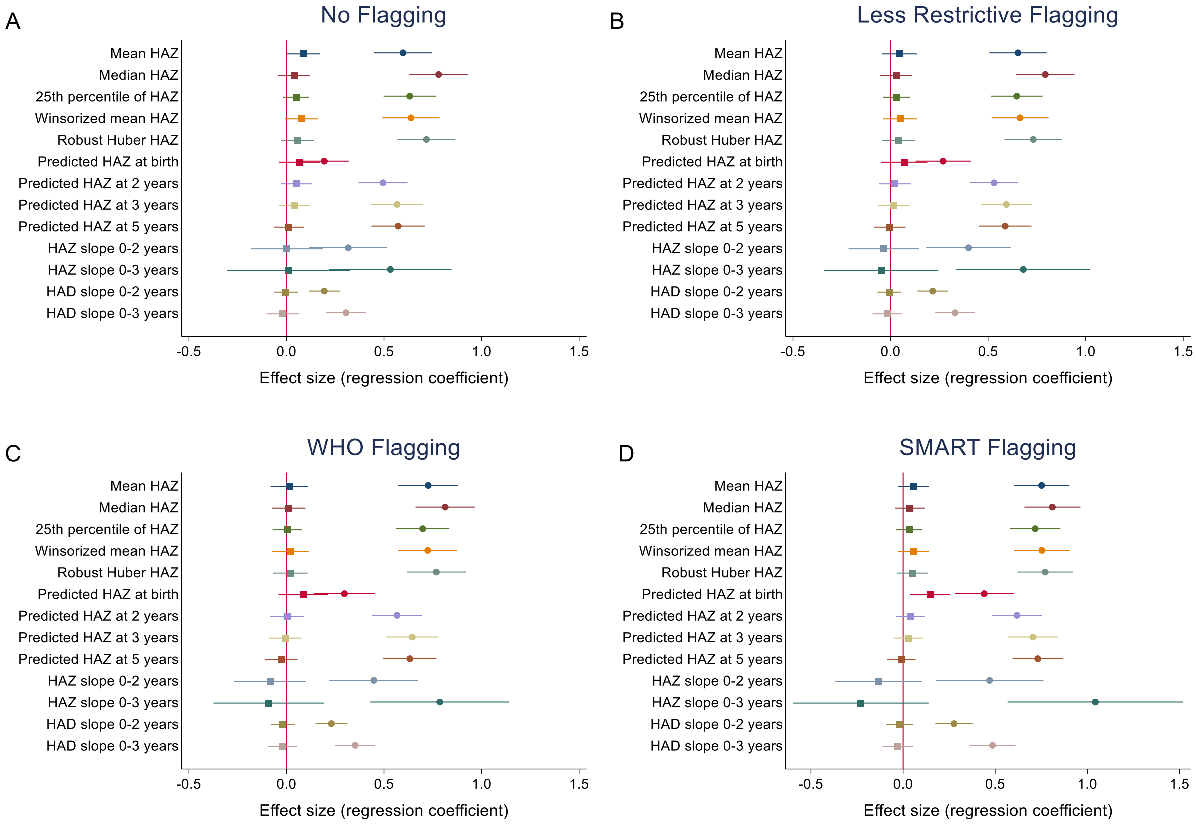
Notes: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 3Q index. For the interaction term, point estimates to the left of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the prevalence of open defecation strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the linear growth metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A10 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *under-5 mortality rate* and the modifying effect of the extended (6Q) anthropometric data quality index (squares) using data from 144 Demographic and Health Surveys from 64 low- and middle-income countries**



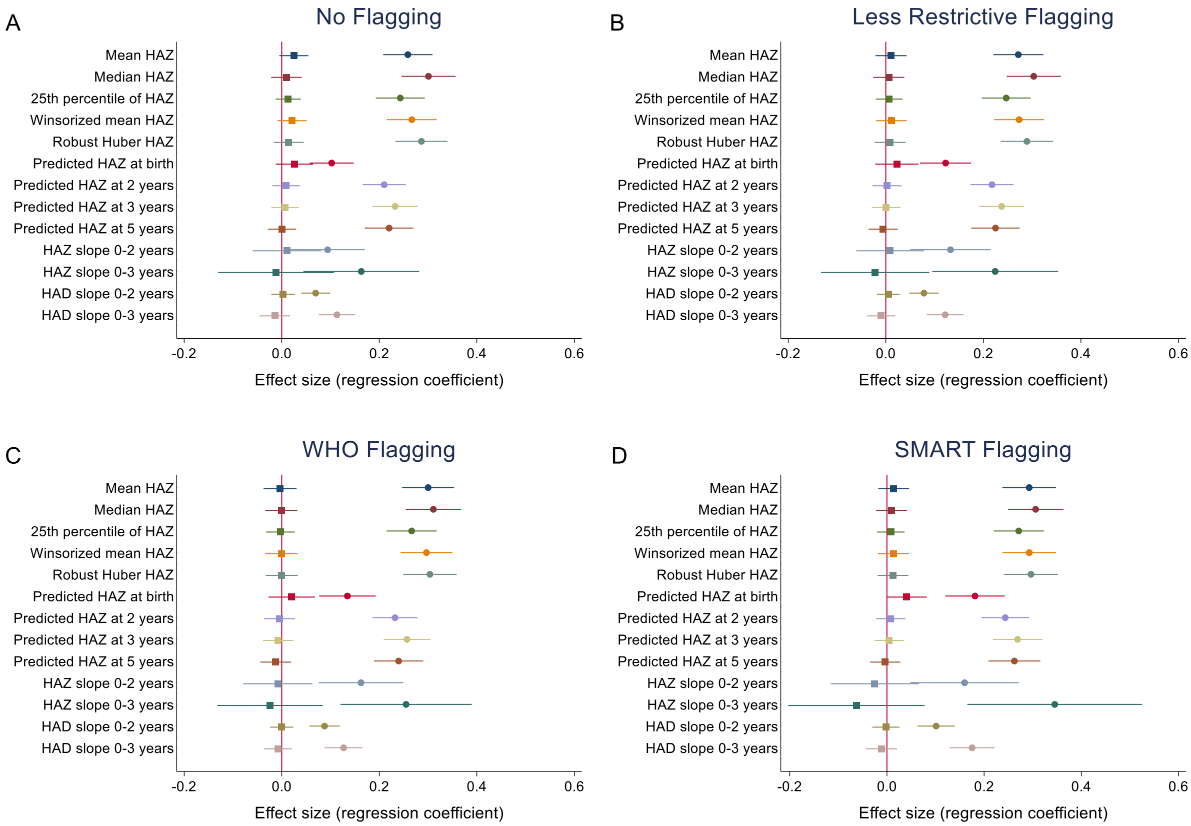
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 6Q index. For the interaction term, point estimates to the left of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A11 Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *gross domestic product* and the modifying of the extended (6Q) anthropometric data quality index (squares) using data from 144 Demographic and Health Surveys from 64 low- and middle-income countries



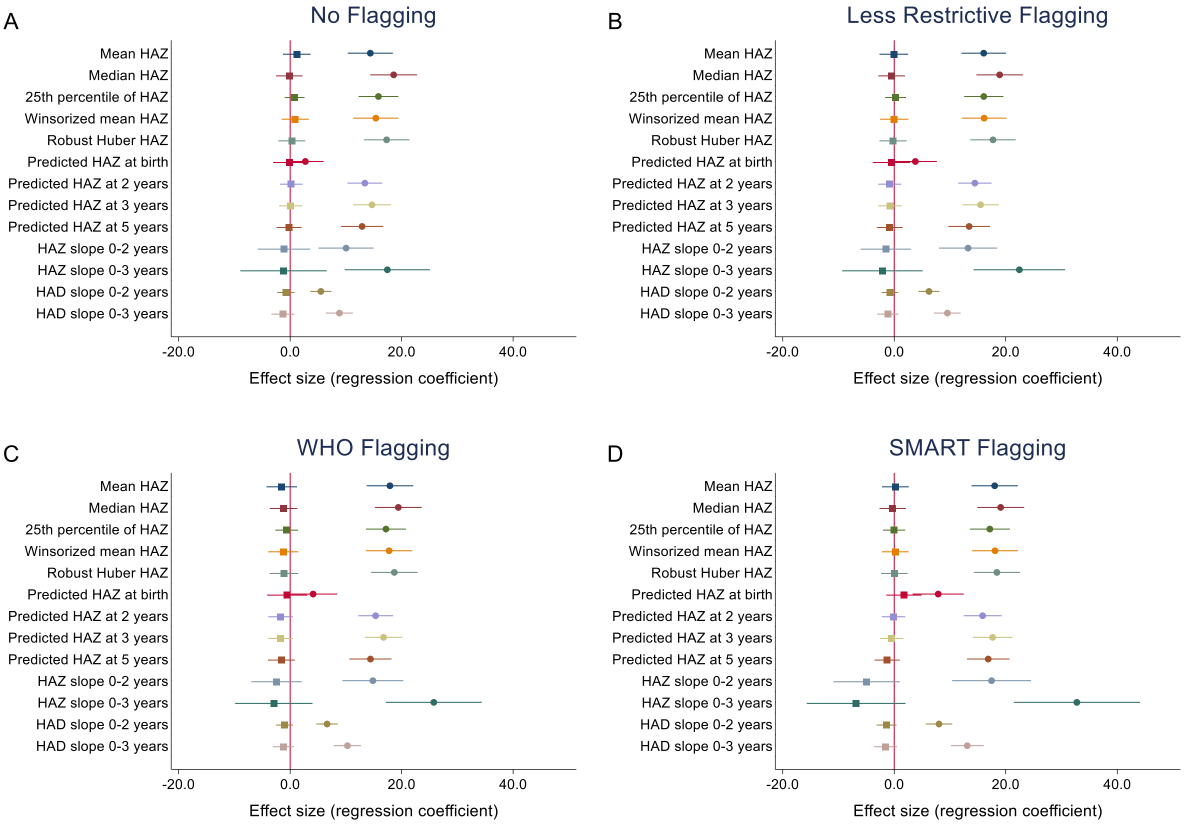
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 6Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A12 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *maternal education* and the modifying effect (with 95% confidence intervals) of the extended (6Q) anthropometric data quality index using data from 143 Demographic and Health Surveys from 63 low- and middle-income countries**



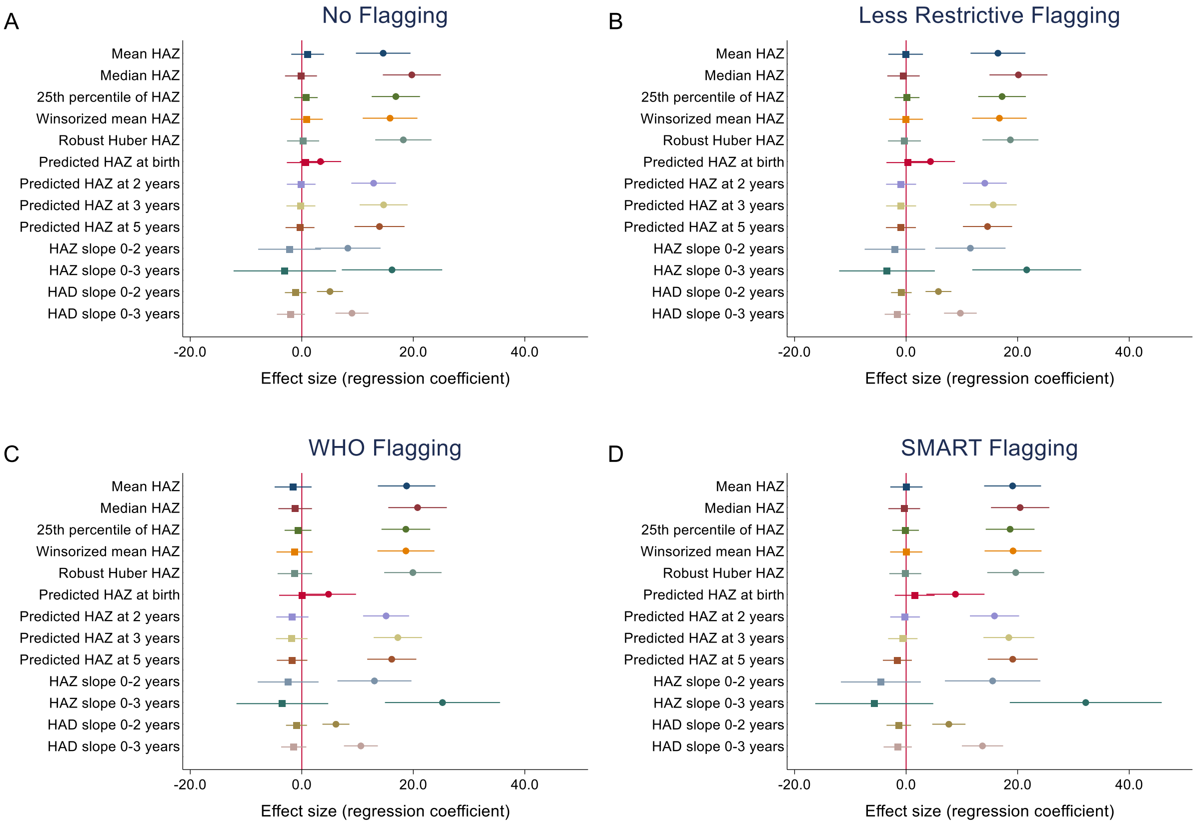
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 6Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A13 **Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *prevalence of access to improved drinking water source* and the modifying effect of the extended (6Q) anthropometric data quality index using data from 129 Demographic and Health Surveys from 61 low- and middle-income countries**



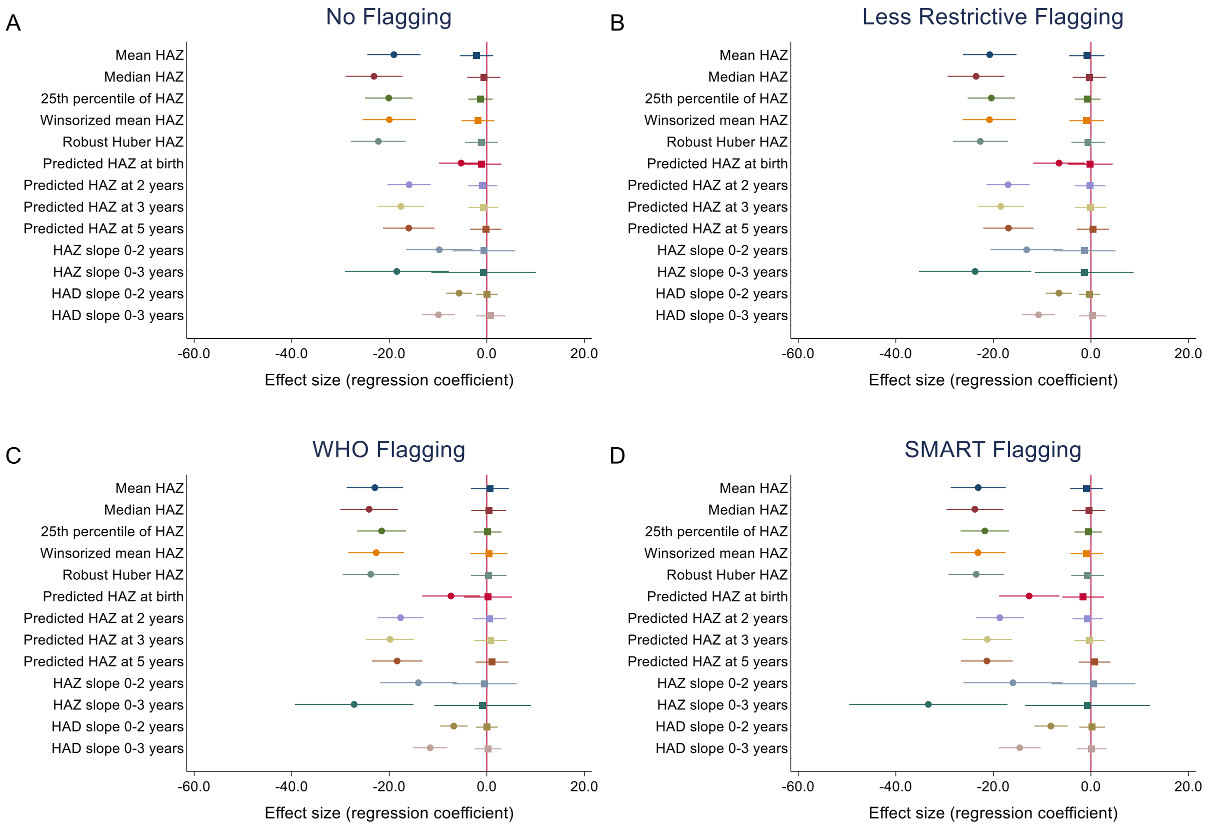
Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 6Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A14 Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *prevalence of access to improved sanitation facilities* and the modifying effect of the extended (6Q) anthropometric data quality index (squares) using data from 129 Demographic and Health Surveys from 61 low- and middle-income countries



Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 6Q index. For the interaction term, point estimates to the right of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).

Figure A15 Regression coefficients from linear mixed effects models for the main effect of alternative HAZ measures of location and model-derived slopes on *prevalence of open defecation* and the modifying effect of the extended (6Q) anthropometric data quality index using data from 129 Demographic and Health Surveys from 61 low- and middle-income countries



Note: Circles represent the main effect (with 95% confidence intervals) and squares represent the modifying effect (with 95% confidence intervals) of the 6Q index. For the interaction term, point estimates to the left of the null indicate that as the data quality score increases (i.e., quality improves), the association between the metric and the indicator strengthens. Exclusion of the null from the 95% confidence interval would be interpreted as a statistically significant modifying effect of data quality. The main effects are from the models that include the interaction terms; therefore, the point estimates for the main effects are interpreted as the magnitude of the association between the HAZ metric and the population indicator when data quality equals zero, which is approximately at the midpoint within the distribution of the data quality index (see Figure 5).