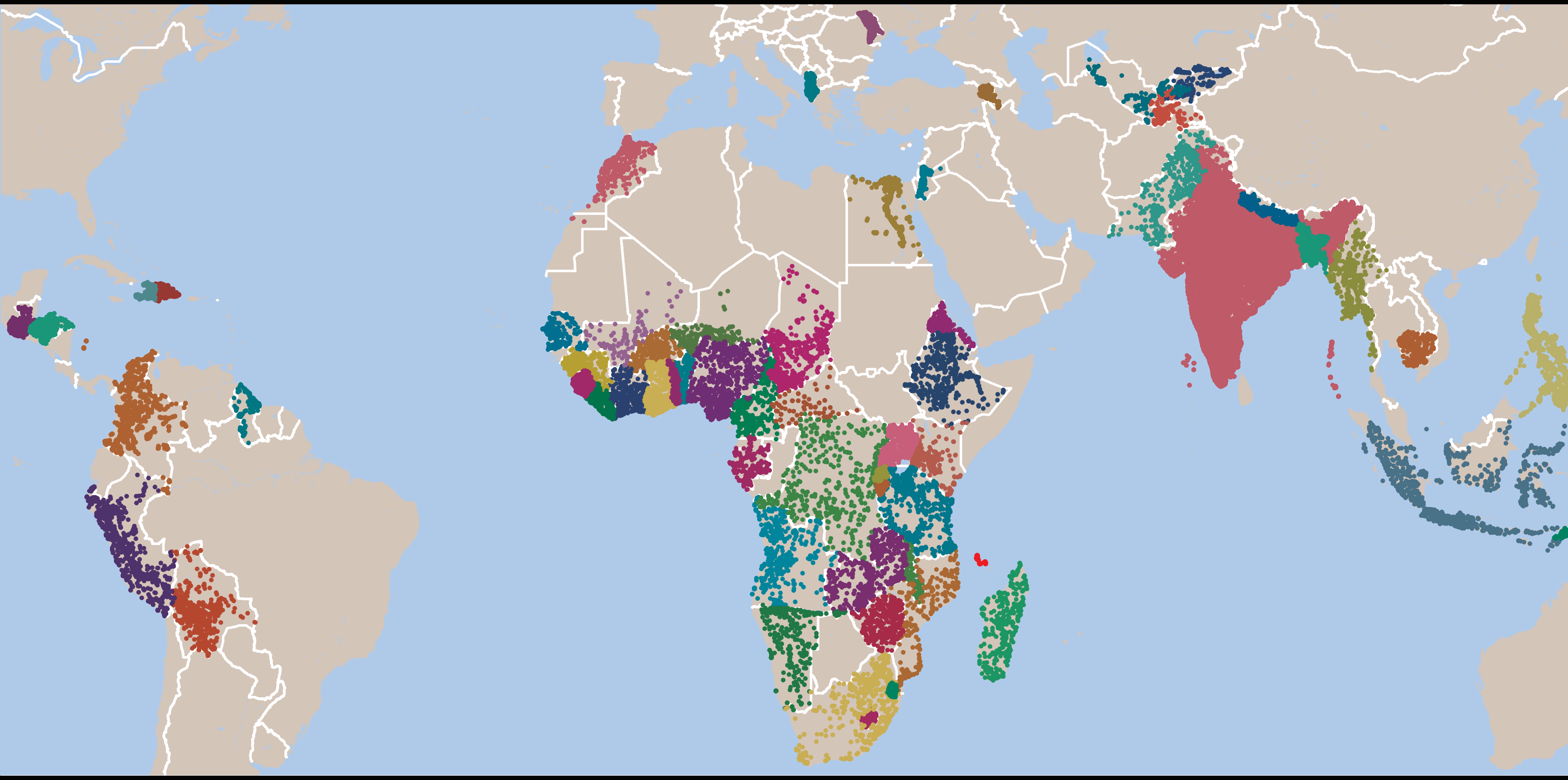
USAID
FROM THE AMERICAN PEOPLE

COMMUNITY IMPROVED SANITATION COVERAGE AND CHILDHOOD STUNTING

DHS SPATIAL ANALYSIS REPORTS 23



SEPTEMBER 2023

This publication was produced for review by the United States Agency for International Development (USAID). The report was prepared by Rose E. Donohue, Rachael Church, Shireen Assaf, and Benjamin K. Mayala.

DHS Spatial Analysis Reports No. 23

Community Improved Sanitation Coverage and Childhood Stunting

Rose E. Donohue[1,2]
Rachael Church[1,3]
Shireen Assaf[1,2]
Benjamin K. Mayala[1,2]

The DHS Program
ICF
Rockville, Maryland, USA

September 2023

[1] The DHS Program
[2] ICF
[3] Avenir Health

Corresponding author: Rose E. Donohue, The DHS Program, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: 301-572-0507; fax: 301-407-6501; email: rose.donohue@icf.com

Acknowledgments: The authors wish to thank Sara Riese, Ryan Mahoney, and James Winter for their review of this report.

Editor: Diane Stoy
Document Production: Joan Wardell

This study was implemented with support from 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. Additional information about The DHS Program can be obtained from ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850 USA; telephone: +1 301-407-6500, fax: +1-301-407-6501, email: info@DHSprogram.com, internet: www.DHSprogram.com.

Recommended citation:

Donohue, Rose E., Rachael Church, Shireen Assaf, and Benjamin K. Mayala. 2023. *Community Improved Sanitation Coverage and Childhood Stunting*. DHS Spatial Analysis Reports No. 23. Rockville, Maryland, USA: ICF.

CONTENTS

TABLES

FIGURES

PREFACE

The Demographic and Health Surveys (DHS) Program is one of the principal sources of international data on fertility, family planning, maternal and child health, nutrition, mortality, environmental health, HIV/AIDS, malaria, and provision of health services.

The DHS Spatial Analysis Reports supplement the other series of DHS reports that respond to the increasing interest in a spatial perspective on demographic and health data. The principal objectives of all the DHS report series are to provide information for policy formulation at the international level and to examine individual country results in an international context.

The topics in this series are selected by The DHS Program in consultation with the U.S. Agency for International Development. A range of methodologies are used, including geostatistical and multivariate statistical techniques.

It is hoped that the DHS Spatial Analysis Reports series will be useful to researchers, policymakers, and survey specialists, particularly those engaged in work in low- and middle-income countries, and will be used to enhance the quality and analysis of survey data.

Sunita Kishor
Director, The DHS Program

ABSTRACT

Lack of access to improved sanitation facilities remains a global problem, with an estimated 1.1 billion individuals estimated to have no access to an improved sanitation facility in 2020. Poor sanitation has been linked to a variety of deleterious health outcomes that include malnutrition, diarrheal diseases, acute respiratory infections, and many neglected tropical diseases such as schistosomiasis, soil-transmitted helminthiases, and trachoma. Although individual access to improved sanitation in households can prevent exposure to fecal contamination in an individual's household, there is growing awareness that the shared community environment can be contaminated when other households lack access to improved sanitation facilities. Thus, there is growing interest in exploring and evaluating the impact of community-level sanitation coverage on individual health outcomes. In this report, we present maps to visualize community improved sanitation coverage and use geospatial modeling to identify areas at risk. We evaluate the association between childhood stunting and household-level and community-level improved sanitation by using multilevel logistic regression on 2018 DHS surveys from Nigeria and Zambia.

Our findings suggest that community improved sanitation coverage should be considered when evaluating sanitation-related health outcomes. We find that although household sanitation access was not associated with childhood stunting in either of the adjusted regression models, community improved sanitation coverage was significantly and inversely associated with childhood stunting in Nigeria. However, we note the limitations of estimating community sanitation coverage with DHS cluster-level data. All households in a community do not have sanitation data and the lack of household GPS coordinates precludes any analyses that consider the proximity to other households in a community. We present different visualizations and tools that use DHS data which may be useful to researchers and policymakers. We create maps that depict the variation in cluster-level estimates in different regions and offer suggestions for using these maps. We offer guidance on how the modeled surfaces in this report and those available in the Spatial Data Repository can be used by policymakers to aid in decisionmaking and by researchers to provide estimates of community sanitation measures in locations that are not sampled by a DHS survey.

Key words: community sanitation, improved sanitation, stunting, sanitation coverage

ACRONYMS AND ABBREVIATIONS

DHS	Demographic and Health Survey
EA	enumeration area
GAM	generalized additive model
GPS	global positioning system
INLA	integrated nested Laplace approximation
JMP	Joint Monitoring Programme
LASSO	least absolute shrinkage and selection operator
LST	land surface temperature
NTD	neglected tropical disease
SDR	Spatial Data Repository
SPDE	stochastic partial differential equations
STH	soil-transmitted helminth
UNICEF	United Nations Children's Fund
WHO	World Health Organization
XGBOOST	gradient boosting

1 BACKGROUND AND OBJECTIVES

1.1 Background

In 2020, approximately 1.1 billion individuals globally lacked access to an improved sanitation facility, defined as a facility which hygienically separates human waste from human contact. Among this population, 494 million practiced open defecation, while 616 million had access only to an unimproved sanitation facility.[1] Lack of access to improved sanitation has been linked to a variety of deleterious health outcomes that include malnutrition, diarrheal diseases, acute respiratory infections, and many neglected tropical diseases (NTDs) such as schistosomiasis, soil-transmitted helminthiases (STH), and trachoma.[2-4] Improved sanitation access betters health outcomes by preventing exposure to feces, the most dangerous of human excreta, and the fecal-oral transmission pathways of diarrheal diseases and many NTDs.[2] Repeated episodes of diarrhea and STH infections are two of the major biological mechanisms that underlie stunting, along with environmental enteric dysfunction.[5] The health burden imposed by poor sanitation is substantial. A recent study attributed 564,308 diarrheal deaths and 29,548,404 disability-adjusted life years (DALYs) to poor sanitation globally in 2019.[6]

Studies that explored the association between poor sanitation and sanitation-related outcomes at the household-level have found inconsistent results, although the totality of the evidence generally suggests that there is an association between poor household sanitation and health outcomes. A meta-analysis found that access to household sanitation was protective against diarrhea, NTDs that included some STH infections, trachoma, schistosomiasis, and height-for-age. However, this study noted high heterogeneity and poor-quality evidence.[7] Another recent meta-analysis found that 72% of studies showed a significant association between childhood stunting and lack of sanitation.[8] Some studies that evaluated multiple DHS datasets in Ethiopia and Nepal found that household access to improved sanitation was a significant predictor of childhood stunting for some but not all survey years.[9,10]

There has been growing interest in considering community-level sanitation in research and intervention programs related to sanitation-related health outcomes. Although household sanitation may protect household members from exposure to their own waste at home, other community members who lack improved sanitation facilities can contaminate the shared environment. In addition, fecal pathogens in the community can enter the household via humans, animals, and flies.[11,12] Some studies have found that community-level sanitation coverage was more important than household access to sanitation for childhood stunting outcomes.[13,14] Researchers have also posited that one reason interventions that can improve sanitation have not resulted in the expected improvements in health outcomes is the low community-level sanitation coverage in the intervention communities.[11,15]

Although many studies with DHS data have evaluated the association between individual or household-level predictors and different health outcomes, few have explored the community-level predictors. The DHS Program conducts surveys at the household-level, and samples groups of adjacent households known as clusters. Clusters are preexisting, geographic groupings within the population, and in most surveys, census enumeration areas (EA) become the survey clusters. An EA is usually a city block in an urban area. In rural areas, an EA is usually a village, part of a village, or a group of small villages.[16] In this report, we use clusters as a proxy for community, which allowed us to use the DHS data to explore community-level improved sanitation.

1.2 Objectives

This report uses data from the 2018 Nigeria DHS and the 2018 Zambia DHS to explore various aspects of community improved sanitation coverage. The report uses geospatial data products from The DHS Program's Spatial Data Repository (SDR) that include survey boundaries, indicator data, and geospatial covariates. This provides data users with examples of how they can use DHS geospatial data products in practice.

First, we produce maps to visualize community improved sanitation coverage using DHS data. We then leverage the community improved sanitation coverage data with geospatial modeling methods to estimate the community improved sanitation coverage in non-surveyed locations. We further combine the modeled estimates with population count data to estimate the population count without access to improved sanitation. Finally, we seek to add to existing knowledge by using multilevel logistic regression to evaluate how improved access to household sanitation and community improved sanitation coverage are associated with one sanitation-related health outcome, childhood stunting.

2 DATA AND METHODS

2.1 Data

This report used data from two recent DHS surveys, the Nigeria 2018 DHS and the Zambia 2018 DHS. These countries have poor sanitation outcomes. The estimated percentage of the de jure population with access to an improved sanitation facility was 53.4% for Nigeria and 54.0% for Zambia in 2018. Both countries demonstrated substantial geographic variation in improved sanitation access. In Nigeria, DHS subnational estimates ranged from 17.2% in Ebonyi State to 93.1% in Abia State. In Zambia, improved sanitation subnational estimates ranged from 6.2% in the Western Province to 80.0% in Lusaka Province. In Nigeria, 22.9% of the population had no access to a toilet, while 23.7% of the population had access to an unimproved sanitation facility. 34.3% of the population was classified as having access to basic sanitation service, defined as having an improved sanitation facility that is not shared with other households. In Zambia, 9.8% of the population had no access to a toilet, 36.2% of the population had access to an unimproved sanitation facility, and 32.9% had access to basic sanitation service.

Both countries also had stunting prevalence that exceeded the 'very high' threshold established for public health significance of ≥30%.[17] Based on the 2018 DHS data, an estimated 36.8% of children under-5 in Nigeria and 34.6% of children under-5 in Zambia were stunted.

2.2 Visualization

We developed two maps to visualize community improved sanitation coverage for DHS data. We first classified the clusters by the percentage of the de jure population living in the cluster with access to improved sanitation, which we refer to as the community improved sanitation coverage. The clusters were overlaid on a choropleth map that showed the regional estimate for the percentage of the de jure population with access to improved sanitation. For Nigeria, the regional estimates represent the six geopolitical zones, while for Zambia, the regional estimates represent the ten provinces. The color scheme of the cluster and subnational estimates are the same for ease of interpretation.

The second map uses the "pie chart" symbology in ArcGIS Pro to visualize the variation in clusters within and between regions. After calculating the community improved sanitation coverage for each cluster, we classified each cluster into a group based on the cluster's percentage: 0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100% community improved sanitation coverage. We calculated the number of clusters in each of these groups for every region. We then used the chart symbology in ArcGIS Pro to create a pie chart for each region that showed the proportion of clusters in the 0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100% community improved sanitation coverage groups.

2.3 Geospatial Modeling

DHS surveys are generally designed to provide estimates of indicators at the national and largest subnational (Admin 1) levels. However, these estimates can conceal geographic variation that exists at lower levels and higher spatial resolutions, which is important for the effective implementation of health programs and interventions. In this analysis, we employ a geospatial modeling approach to provide estimates of the percentage of the de jure population with access to an improved sanitation facilities across the entire country, with a spatial resolution of 5 x 5 km resolution. Each 5 x 5 km area is referred to as a pixel. Our primary objective was to provide prevalence predictions at a fine spatial resolution for Zambia and Nigeria. After creating a modeled surface that estimated the percentage of the population with improved sanitation access in each pixel, we further extend these estimates by combining them with population count data to calculate the population at risk.

The methodology used in this analysis is similar to previous analyses of water and sanitation,[18] childhood diarrhea disease,[19] vaccination coverage,[20] and HIV.[21] This method was utilized because it has been shown to improve the predictive accuracy of geospatial modeling. We provide details of this approach in the next sections.

2.3.1 Geospatial covariates

In this analysis, we included the following socioeconomic, environmental, and health-related covariates to improve the predictions of improved sanitation (Table 1).

Table 1 Geospatial covariates used to develop the models in this study

Covariates	Spatial resolution	Source
Travel time to nearest settlement >50,000 inhabitants	5 x 5 km	Malaria Atlas Project
Aridity	10 x 10 km	Climatic Research Unit gridded Time Series (CRUTS)
Diurnal temperature range	10 x 10 km	CRUTS
Potential evapotranspiration (PET)	10 x 10 km	CRUTS
Daily maximum temperature	10 x 10 km	CRUTS
Elevation	1 x 1 km	NOAA
Enhanced vegetation index (EVI)	5 x 5 km	NASA
Daytime land surface temperature (LST)	5 x 5 km	NASA
Diurnal difference in LST	5 x 5 km	NASA
Nighttime LST	5 x 5 km	NASA
Population distribution	1 x 1 km	WorldPop

The geospatial covariates were selected because they present factors or proxies for factors that previous studies have identified as associated with various DHS indicators, including the population with access to an improved sanitation facilities.[18,22,23,24]

The covariates were obtained from a variety of data sources, and have different spatial references, projections, extents, and dimensions. Therefore, a series of essential spatial processing steps were undertaken, which involved:

- Reprojection: All data layers were reprojected to align with a consistent coordinate reference system, which was the widely used World Geodetic System 1984.
- Masking: Done to an extent that encompassed the boundaries of the study area; and
- Resampling: To facilitate uniform analysis, the input covariate rasters were resampled to align the spatial resolution of the covariate to the 5 km × 5 km resolution used in modeling.

For the population covariates, a summation method was employed for resampling, while those originally available at a 5 × 5 km resolution required no further adjustments. Other covariates underwent resampling with bilinear interpolation. These technical processing steps were executed using the 'raster' and 'shapefiles' packages within the R software environment,[25] which ensured rigorous data preparation for the subsequent analysis.

2.3.2 Overview of the modeling approach

Figure 1 provides a conceptual outline of the geospatial modeling framework employed to model DHS indicators and produce the pixel level estimates. The approach involved the following steps:

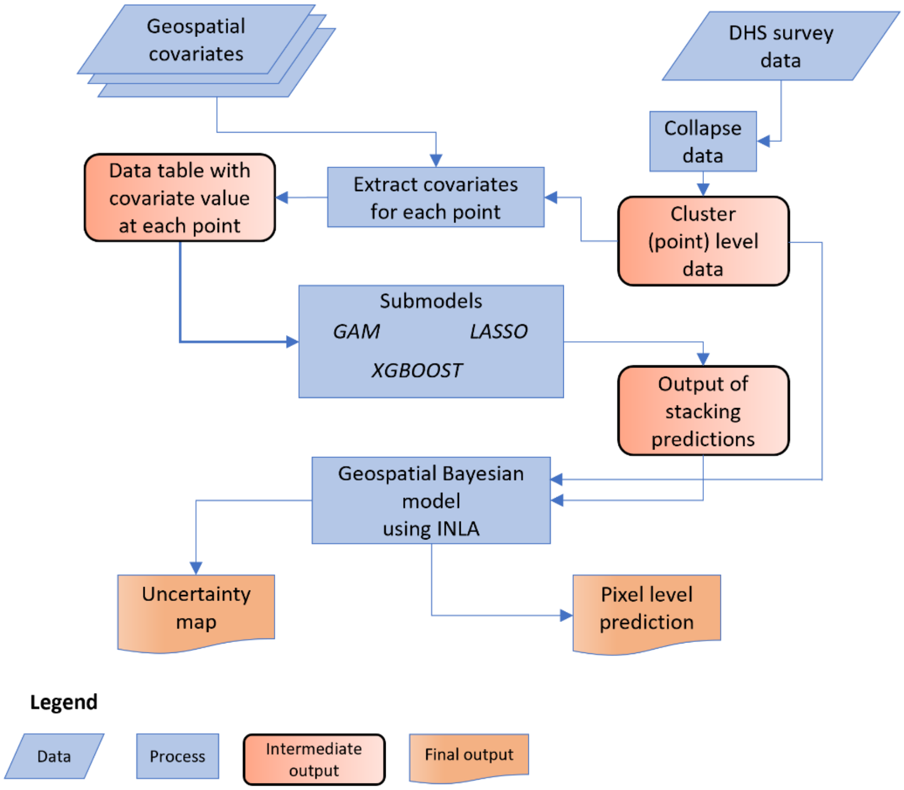
Step 1 We aggregated individual-level DHS survey data to the finest spatial resolution (latitude and longitude) that represented the location of the survey cluster.

Step 2 Both covariates and the cluster (point) level data were imported into the R environment for statistical computing. We then applied the "raster" package to extract the corresponding covariate pixel values at each survey cluster point.

Step 3 The point level data (from Step 2), along with their associated covariates, were integrated into a stacked generalization ensemble model, elaborated in Section 2.3.2.

Step 4 Prediction surfaces produced by the ensemble model were then utilized as covariates to calibrate the final geospatial Bayesian model. The outputs of the final model are pixel-level mean estimates with associated uncertainty at a 5 x 5 km resolution.

Figure 1   Geospatial modeling process flowchart



2.3.3 Covariate modeling with stacked generalization

In many applications, generic geospatial modeling is sufficient to produce a highly predictive model. However, when modeling outcomes in which the underlying process is linked to the covariates and demographic parameters through complex non-linearities and interactions, a simple linear mean of the form βX can be insufficient. To address this, we use a stacked generalization approach, which preprocesses covariates by using highly predictive machine learning techniques.

Stacked generalization is a general ensemble modeling approach that combines multiple model algorithmic methods to produce a meta-model that has equal or better predictive performance relative to a single modeling approach. We adopted this approach to capture the potential complex interactions and non-linear effects among the geospatial covariates.[26]

Our selection of machine learning (submodels) includes three approaches: (1) GAM: generalized additive model;[27] (2) LASSO: least absolute shrinkage and selection operator regression;[28] and (3) XGBOOST: gradient boosting.[29] These submodels were fitted to the improved sanitation survey data with geospatial covariates (described in Table 1) as explanatory predictors. Submodels were fit in R using the mgcv, xgboost, and glmne packages.[25]

2.3.4 Model specification and development

We modeled improved sanitation as Y_i, the number of 'positive' individuals among those sampled at cluster location $s_i, i = 1, \ldots n$, using a binomial spatial regression with a logit link function.[30,31] If N_i is the total number of individuals sampled at cluster s_i, the model can be written as:

$$Y_i \sim Binomial(N_i, p_i)$$

$$logit(p_i) = \beta_0 + \beta \mathrm{X}_i + \omega_i + \varepsilon_i$$

$$\omega_i \sim GP(0, \Sigma)$$

Where:

- β_0 denotes the intercept,
- p_i is the probability, representing the underlying prevalence at cluster s_i,
- $X_i = (X_{i1,} X_{i2}, \ldots X_{im})$ is the vector of logit-transformed covariates for location s_i obtained from the submodels (GAM, LASSO, and XGBOOST) generated from the stacked generalization modeling (as described in Section 3.3.2),
- $\beta = (\beta_1, \beta_2, \ldots \beta_m)$ vector of regression coefficients on the submodels represent their respective predictive weighting and are constrained to the sum of one,[26]
- ω_i is a correlated spatial error term, accounting for spatial autocorrelation between data points, and
- $\varepsilon_i \sim N(0, \sigma_{nug}^2)$ is an independent error term known as the nugget effect.

The spatial error term ω_i is modeled as a Gaussian process with a zero-mean and spatially structured covariance matrix $\sum$. The spatial covariance $\sum$ was modeled using a stationary and isotropic Matérn function.[30]

The Bayesian geostatistical model analysis was implemented through a stochastic partial differential equations (SPDE) approach in the integrated nested Laplace approximation (INLA) algorithm as applied in the R-INLA package (Rue, Martino, and Chopin 2009).[32]

2.3.5 Pixel-level model estimates

The prediction surfaces generated from the submodels (described in Section 2.3.2) were used as input covariates in the geostatistical models implemented in INLA. The final estimates (and uncertainty) for each indicator were generated by taking $k = 1, \ldots 1000$ samples from the posterior predictive distribution. Pixel level estimates that covered the modeling country were produced at a high spatial resolution of 5 x 5 km.

2.3.6 Population at risk estimates

While the modeled surfaces depict the percentage of the population with access to improved sanitation facilities for each pixel, it is also important to consider how many individuals are affected in each pixel. In this report, we utilized the population count data from the survey year estimated at a 1 x 1 km resolution by

WorldPop (https://www.worldpop.org/). We resampled this raster layer to match the 5 x 5 km resolution of the modeled surfaces. Since policymakers are interested in reducing both the percentage and population count without access to an improved sanitation facility, we calculated the percentage without access to an improved sanitation facility for each pixel by subtracting the modeled estimate of the population percentage with access to an improved sanitation facility from 100%. We then multiply this value for each pixel by the population count residing in each pixel to compute the population count in each pixel without access to an improved sanitation facility.

2.4 Multilevel Regression

We used multilevel, multivariable logistic regression to evaluate the association between community-level improved sanitation coverage and childhood stunting, as well as the association between household access to improved sanitation and childhood stunting. We also conducted multilevel, multivariable regression for community-level improved sanitation coverage and childhood stunting by region in Nigeria and Zambia.

2.4.1 Individual-level measures

Outcome

The outcome of interest was stunting among children under age 5. Stunting was defined by the World Health Organization (WHO) definition of having a height-for-age score two standard deviations below the mean on the WHO Child Growth Standards.

Covariates of interest

We evaluated the following covariates of interest that have been shown in the literature to be associated with childhood stunting.[33]

Age of child. Child's age was grouped into four categories: <6 months, 6–11 months, 12–23 months, and 24–59 months.

Sex of child. Child's sex was grouped into two categories: male and female.

Stature of child's mother. The stature of the mother was grouped into two categories: short and not short. For adult mothers (age 20 or older), women with heights below 145cm were classified as short. For adolescent mothers (age 15 to 19), we calculated the height-for-age score. Adolescent mothers with a height-for-age score below two standard deviations of the mean were classified as short. The *zanthro* package in Stata was used for this calculation.

Household wealth quintile. The DHS Program constructs a household wealth index that represents the relative household wealth based on household asset ownership and household characteristics. The de jure population is classified as belonging to a wealth quintile, which ranges from the lowest, or poorest, quintile, to the highest, or wealthiest, quintile. Further information on the wealth index construction can be found in previous DHS publications.[34,35]

Household crowding. Household crowding was defined as when a household's de jure population divided by the number of sleeping rooms was greater than or equal to three.

Improved source of drinking water. The status of the drinking water source of each household was classified as either improved or unimproved based on the definitions established by the WHO/United Nations Children's Fund (UNICEF) Joint Monitoring Programme (JMP) for Water Supply and Sanitation.[36]

2.4.2 Community-level measures

We evaluated five covariates at the cluster-level, two of which were from The DHS Program's Geospatial Covariates (GC) datasets, which are available through The DHS Program's website and on the SDR (https://spatialdata.dhsprogram.com/covariates/).

Community improved sanitation coverage. Community improved sanitation coverage was defined as the percentage of households within a cluster that have access to an improved sanitation facility, based on the classifications developed by the WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation.[36] Shared facilities were classified as improved facilities, which follows the JMP definition change in 2017. We divided the community improved sanitation coverage by 10 so the model odds ratios represent the percentage increase in odds for every 10 percentage point increase in community improved sanitation coverage

Nightlights. The nightlights covariate measures the luminosity of an area during the nighttime hours as measured by the Visible Infrared Imaging Radiometer Suite (VIIRS). This indicator is used as a proxy for economic development. The value for each cluster represents the average nighttime luminosity of the area within a 2km (urban) or 10km (rural) buffer. Detailed information on the extraction process can be found in the second edition of *The Geospatial Covariate Datasets Manual*.[37] Studies have observed associations between nightlights and childhood stunting.[38-40]

Travel times. The value for each cluster represents the average time in minutes that are required to reach a settlement of 50,000 or more people from the area within a 2km (urban) or 10km (rural) buffer of the cluster's displaced location based on 2015 data.[37] Travel times represents the accessibility of each cluster to reach the opportunities and services that are available in urban centers. Studies have found associations between childhood stunting and travel times to urban centers as well as health facilities.[38,41]

Residence. Each cluster is defined as primarily urban or rural based on the designation by the country's statistical office at the time of the survey.

Region. Each cluster belongs to a particular region of the country. The GPS coordinates of each cluster are validated to these regions, as well as lower subnational administrative levels. For Nigeria, we used the six geopolitical zones, which represent groups of states. For Zambia, we used the ten provinces.

2.4.3 Regression analysis

We evaluated the association between the covariates of interest and childhood stunting using multilevel, multivariable logistic regression, as has previously been done in AS82.[42] The analysis was run separately for the Nigeria and the Zambia datasets.

We used the *melogit* command in Stata with the cluster as the grouping structure and *svyset* multilevel weights. The multilevel models used cluster-level weights which were constructed with the methods recently developed to estimate level weights for complex surveys.[43]

We evaluated the association between each variable and the stunting outcome. We assessed the pairwise correlation between variables to assess any collinearity issues, and then added all indicators that did not exhibit high correlation with the multivariable model.

3 RESULTS

3.1 Visualization

3.1.1 Nigeria

Figure 2a depicts the cluster estimates that represent the percentage of the population in each cluster with access to improved sanitation, overlaid on the regional estimates of the percentage of the de jure population with access to improved sanitation. There is considerable variation within regions, and the density of clusters in the southern part of the country makes it challenging to visualize all clusters.

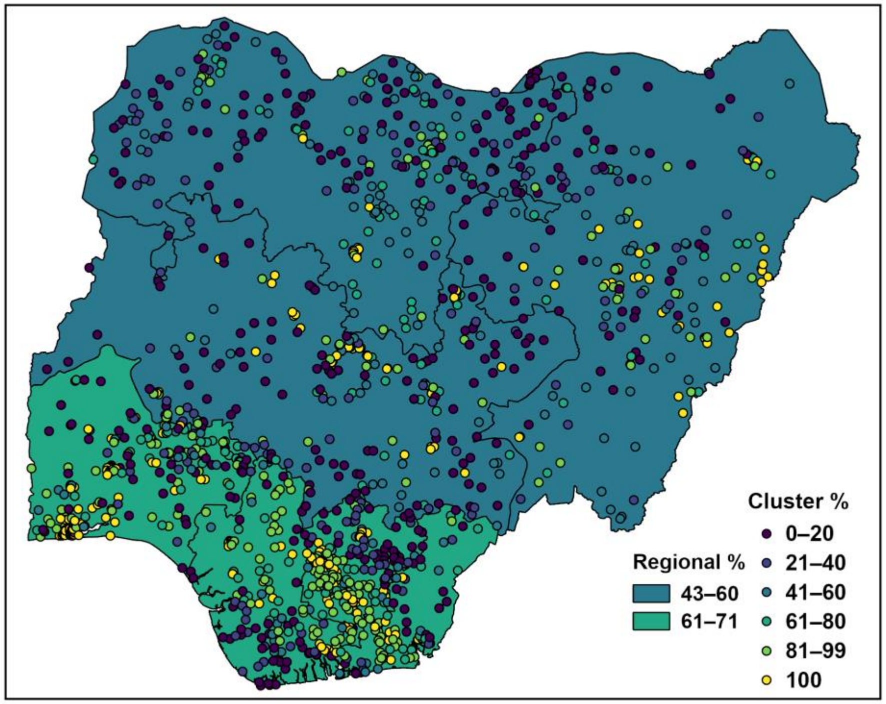
Figure 2b depicts a pie chart for each region that represents the proportion of clusters in each category of improved sanitation coverage, with the categories 0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100%. The data used to produce this map are shown in Table 2. This symbology allows users to visualize how cluster estimates vary within a region, and how regions are performing in relation to other regions. Here, we see that the southern regions, in particular the South West region, have higher percentages of cluster with 100% improved sanitation coverage, as represented by the yellow color in the pie chart. From Table 2, we see that 31% of clusters in South West (74 out of 238 clusters) have 100% improved sanitation coverage. North West has the lowest percentage of clusters with 100% improved sanitation coverage, with only 4% (10 out of 275 clusters). While South West has the highest percentage of cluster with 100% improved sanitation coverage, we also see that each region has a sizeable percentage of clusters with 0–20% improved sanitation access, ranging from 19% in South West (46 out of 238 clusters) to 39% in North Central (99 out of 252 clusters).

Table 2 Number of clusters in each region grouped by the percentage of the de jure population with access to improved sanitation, Nigeria

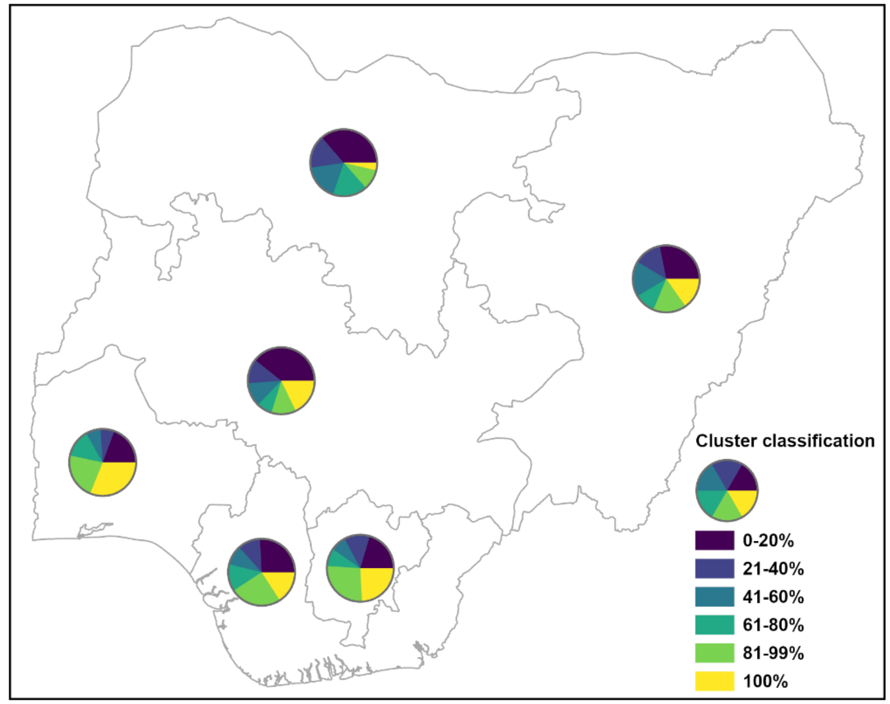
Region	0–20%	21–40%	41–60%	61–80%	81–99%	100%	Total
North Central	99	30	29	19	30	45	**252**
North East	61	29	37	22	35	33	**217**
North West	100	44	48	46	27	10	**275**
South East	38	23	14	16	50	45	**186**
South South	57	24	21	29	55	35	**221**
South West	46	16	17	32	53	74	**238**
Grand Total	**401**	**166**	**166**	**164**	**250**	**242**	**1,389**

Figure 2 Maps that depict community sanitation coverage for Nigeria 2018 DHS: (a) cluster-level improved sanitation coverage estimates overlaid on the regional estimates that represent the percentage of the de jure population with access to improved sanitation facilities; and (b) the proportion of clusters in each region classified by their community improved sanitation coverage (0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100%)

(a)



(b)



3.1.1 Zambia

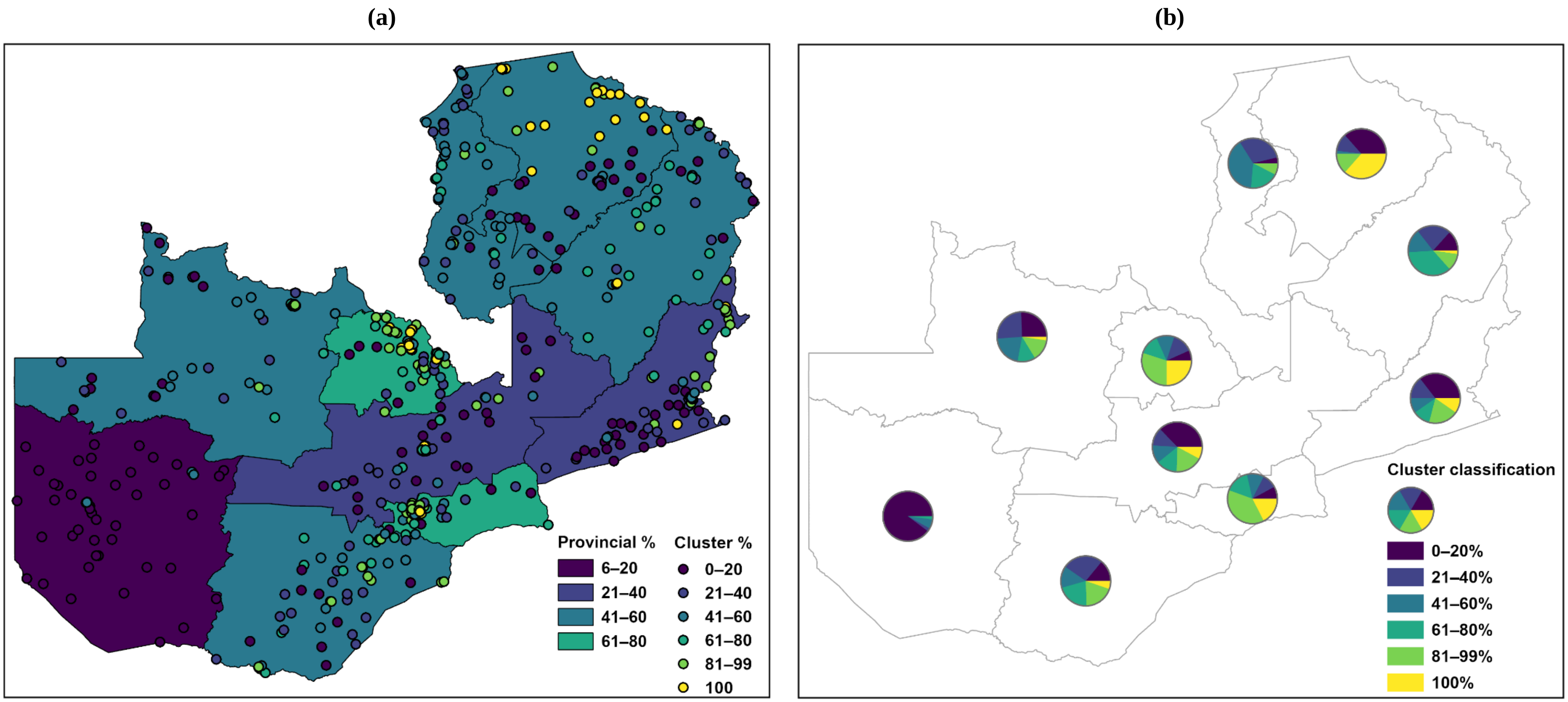
Figure 3a depicts the cluster estimates that represent the percentage of the population in each cluster with access to improved sanitation, overlaid on the regional estimates of the percentage of the de jure population with access to improved sanitation. Figure 3b depicts a pie chart for each region that represents the number of clusters belonging to each group of improved sanitation coverage, including the groups 0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100%. The data used to produce this map are shown in Table 3.

Variation exists within the regions. For example, in the Northern Province, 19 of its 53 (36%) clusters had an improved sanitation coverage of 0–20%, while another 19 had 100% improved sanitation coverage. Other regions show less variation. In the Western Province, 44 of the 49 (90%) clusters were classified as having between 0–20% improved sanitation coverage. This can be visualized in Figure 3b.

Table 3 Number of clusters in each region grouped by the percentage of the de jure population with access to improved sanitation, Zambia

	Community improved sanitation coverage						
Region	**0–20%**	**21–40%**	**41–60%**	**61–80%**	**81–99%**	**100%**	**Total**
Central	20	8	7	7	9	4	**55**
Copperbelt	4	9	7	8	18	15	**61**
Eastern	22	9	6	7	12	6	**62**
Luapula	2	16	22	10	4	0	**54**
Lusaka	6	7	8	10	24	11	**66**
Muchinga	6	10	7	16	5	1	**45**
North Western	11	11	9	5	6	1	**43**
Northern	19	6	1	0	8	19	**53**
Southern	8	15	8	12	11	3	**57**
Western	44	1	3	1	0	0	**49**
Total	**142**	**92**	**78**	**76**	**97**	**60**	**545**

Figure 3 **Maps that depict community sanitation coverage for Zambia 2018 DHS: (a) cluster-level improved sanitation coverage estimates overlaid on the regional estimates that represent the percentage of the de jure population with access to improved sanitation facilities; and (b) the proportion of clusters in each region classified by their community improved sanitation coverage (0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100%)**



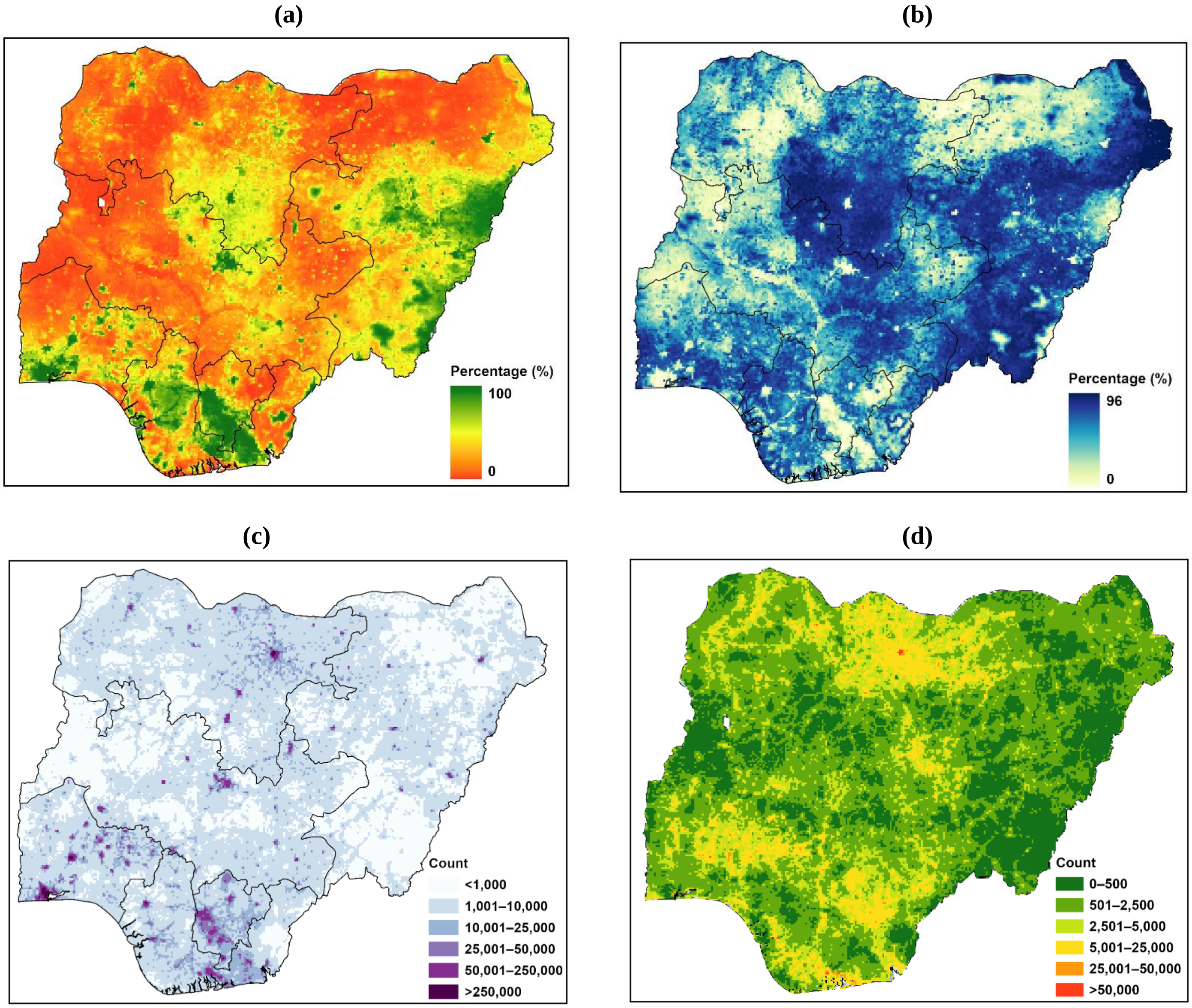
3.2 Geospatial Modeling

3.2.1 Nigeria

The modeled estimates for the percentage of the de jure population using an improved sanitation facility at a 5 x 5 km resolution for Nigeria are presented in Figure 4a. Considerable variation is observed in each region. The green color represents high improved sanitation access, which is observed in the southern and eastern areas of the country, with additional smaller areas of high access throughout the country. The northern and western parts of the country appear to have very low access to improved sanitation, as shown by the red color. The uncertainty of these estimates, represented by the width of the 95% uncertainty interval (the difference between the upper bound and lower bound) is depicted in Figure 4b.

The population count, classified into six classes to better visualize the variation, is shown in Figure 4c. The population at risk, which is computed by multiplying the population count raster layer (Fig. 4c) by (100 - the population estimate with access to improved sanitation), is presented in Figure 4d. The population at risk quantifies the number in each pixel without access to improved sanitation. This map varies from the prevalence estimates in Figure 4a. Here we see the highest number of individuals who lack access to improved sanitation facilities in small, concentrated areas of the south, west, and north of the country.

Figure 4 Modeled surfaces for (a) the percentage of the de jure population using an improved sanitation facility; (b) width of the 95% uncertainty interval; (c) the population count in 2018 obtained from WorldPop; and (d) the estimated population at risk

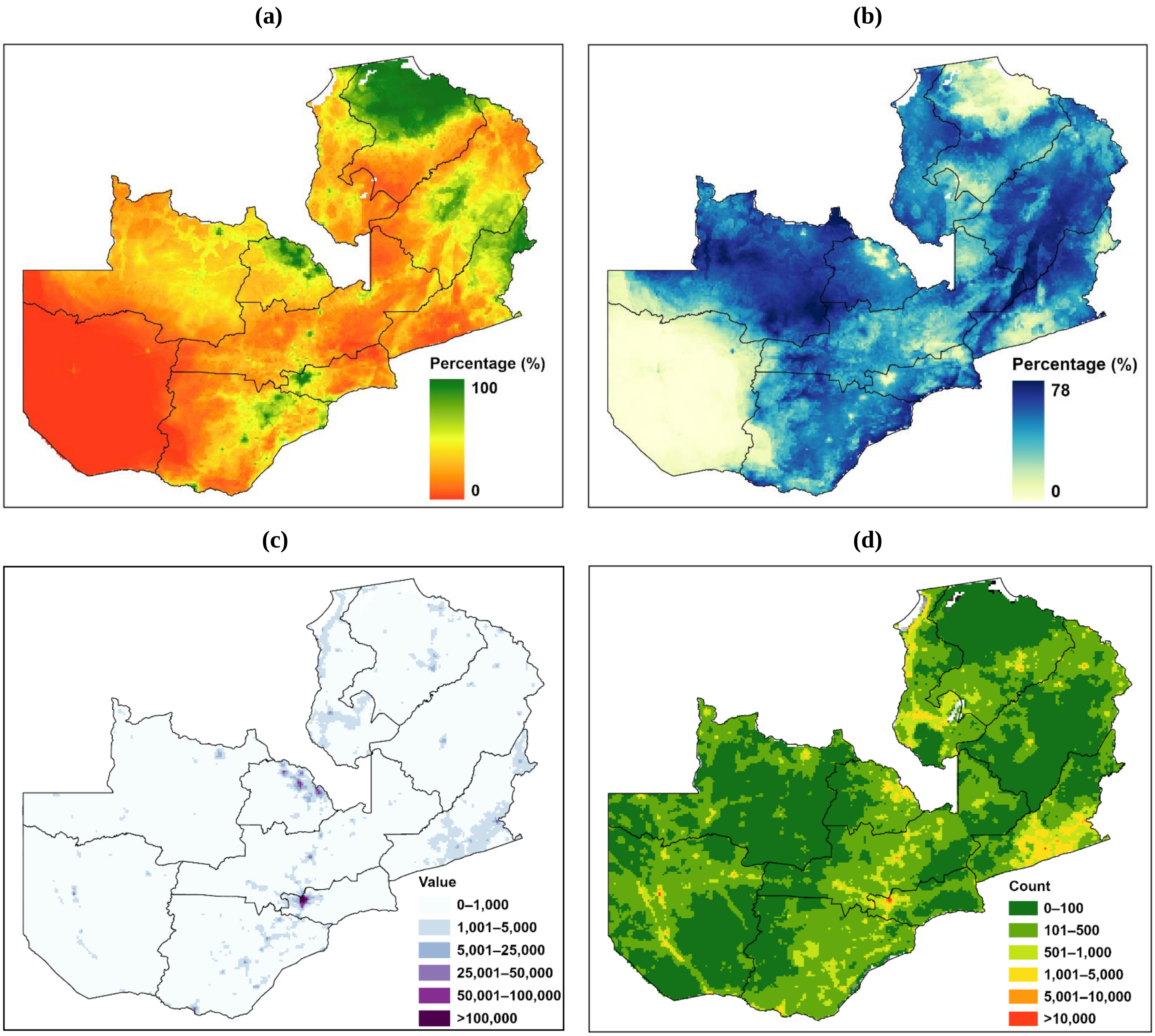


3.2.2 Zambia

Figure 5a depicts the modeled estimates of the percentage of the de jure population using an improved sanitation facility at a 5 x 5 km resolution in Zambia. The estimates of this indicator vary from 0–100%, with low access to improved sanitation observed in the west and center of the country, and high access to improved sanitation observed in the northeast, along with other smaller areas throughout the country. The uncertainty of these estimates is shown in Figure 5b. Again, we see variation, with some areas having very little uncertainty, while other areas have high uncertainty in the modeled estimates.

The population count, classified into six classes, is shown in Figure 5c. Figure 5d shows the population at risk, representing the count in each pixel without access to an improved sanitation facility, which was computed by multiplying the population count raster layer (Fig. 5c) by (100 - the population estimate with access to improved sanitation). Again, we see the patterns in this map vary from the prevalence estimates in Figure 5a.

Figure 5 Modeled surfaces for (a) the percentage of the de jure population using an improved sanitation facility; (b) width of the 95% uncertainty interval; (c) the population count in 2018 obtained from WorldPop; and (d) the estimated population at risk



3.3 Multilevel Regression

3.3.1 Nigeria

The associations between community-level improved sanitation coverage and childhood stunting is presented in Table 4. Increased, community-level improved sanitation coverage was associated with reduced odds of childhood stunting in the unadjusted model (OR = 0.87; 95% CI: [0.85, 0.89]). This association remained significant in the adjusted model (aOR = 0.97; 95% CI: [0.95, 0.99]), although the adjusted odds ratio was less protective after adjusting for other important determinants of stunting, such as the age and sex of the child, the stature of the mother, household wealth, and the region of residence. The adjusted model suggests that for every 10% increase in community-level improved sanitation coverage, the odds of stunting decreases by 3%.

Female children had reduced odds of stunting compared to male children (aOR = 0.74; 95% CI: [0.65, 0.83]. Children with mothers who were considered to be of short stature had significantly higher odds of stunting compared to children whose mothers were not considered short (aOR=2.56; 95% CI: [1.63, 4.01]). A clear gradient was observed among the age groups, with children in higher age groups being increasingly more likely to be stunted when compared to children younger than 6 months. A similar gradient was observed between household wealth and stunting; the OR and AOR were progressively smaller moving from the second wealth quintile to the highest wealth quintile when compared to the lowest quintile. The difference between the second wealth quintile and the lowest quintile was no longer significant in the adjusted model, although children in the higher three quintiles had significantly lower odds of stunting when compared to children in the lowest wealth quintile.

The association between household access to improved sanitation and childhood stunting is shown in Appendix Table 1. While improved household sanitation was significantly associated with childhood stunting in the unadjusted model (OR = 0.74; 95% CI: [0.64, 0.85]), improved household sanitation was no longer significant in the multivariable model.

The results of the multilevel regression models conducted for each region in Nigeria are presented in Appendix Table 3. Of note, community-level improved sanitation coverage was significantly associated with reduced odds of childhood stunting in the multivariable models in two of the six regions: North West and South West. In the other four regions, community-level sanitation coverage was not significantly associated with reduced odds of stunting. In three of the regions, the aOR was less than 1, while in the South South Region, the aOR between community-level improved sanitation coverage and childhood stunting was 1.00.

Table 4 Association of community-improved sanitation coverage and childhood stunting using multilevel logistic regression, Nigeria

Individual-level covariates	Unadjusted OR	[95% CI]	Adjusted OR	[95% CI]
Child's age				
<6 months	Ref.	--	Ref.	--
6–11 months	1.42*	[1.06, 1.90]	1.37*	[1.02, 1.84]
12–23 months	3.87**	[2.99, 5.00]	3.83**	[2.95, 4.97]
24–59 months	4.36**	[3.41, 5.58]	4.31**	[3.36, 5.52]
Child's sex				
Male	Ref.	--	Ref.	--
Female	0.74**	[0.66, 0.83]	0.74**	[0.65, 0.83]
Mother's stature				
Not short	Ref.	--	Ref.	--
Short	2.57**	[1.65, 4.02]	2.56**	1.63, 4.01
Household wealth quintile				
Lowest quintile	Ref.	--	Ref.	--
Second quintile	0.75**	[0.64, 0.89]	0.90	[0.75, 1.08]
Middle quintile	0.52**	[0.43, 0.62]	0.79*	[0.64, 0.97]
Fourth quintile	0.32**	[0.26, 0.39]	0.57**	[0.44, 0.73]
Highest quintile	0.17**	[0.14, 0.22]	0.35**	[0.25, 0.48]
Household crowding				
Not crowded	Ref.	--	Ref.	--
Crowded	1.09	[0.98, 1.22]	1.05	[0.94, 1.18]
Household water source				
Unimproved	Ref.	--	Ref.	--
Improved	0.60**	[0.52, 0.69]	0.92	[0.79, 1.06]
Cluster-level covariates				
Nightlights	0.94**	[0.92, 0.96]	0.99	[0.97, 1.01]
Travel time to city	1.01**	[1.01, 1.01]	1.00	[1.00, 1.00]
Community improved sanitation coverage	0.87**	[0.85, 0.89]	0.97*	[0.95, 0.99]
Residence				
Urban	Ref.	--	Ref.	--
Rural	2.46**	[2.15, 2.82]	1.03	[0.87, 1.22]
Region				
North Central	Ref.	--	Ref.	--
North East	2.75**	[2.24, 3.38]	2.55**	[2.04, 3.21]
North West	4.28**	[3.53, 5.18]	4.14**	[3.41, 5.02]
South East	0.57**	[0.45, 0.73]	0.75*	[0.59, 0.96]
South South	0.63**	[0.50, 0.81]	0.84	[0.64, 1.08]
South West	0.76*	[0.60, 0.96]	1.23	[0.96, 1.57]
Observations			11,364	
Number of groups			1,378	

* Significant at *p* < .05; ** significant at *p* < .01; OR = odds ratio; CI = confidence interval

3.3.2 Zambia

For Zambia, the results of the multilevel regression are presented in Table 5. Community-level improved sanitation coverage was not significantly associated with reduced odds of childhood stunting in the unadjusted (OR = 0.98; 95% CI: [0.96, 1.00]) or adjusted models (aOR = 1.00; 95% CI: [0.98, 1.03]). Significant determinants of childhood stunting in the adjusted model included the age and sex of the child, household wealth, the region of residence, and the community-level variable that measured accessibility, which included travel times to cities.

Female children had reduced odds of stunting compared to male children (aOR = 0.67; 95% CI: [0.60, 0.76]). Similar to the Nigeria results, clear gradients were observed between both age and stunting, and household wealth and stunting. For age, children in higher age groups had increasingly higher odds of stunting when compared to children younger than 6 months of age. For household wealth, the higher the wealth quintile, the lower the odds of stunting, when compared to the reference group of the lowest wealth quintile. Like Nigeria, the comparison between the second quintile and the lowest quintile was not significant in the adjusted model, although children belonging to the highest three wealth quintiles had significantly lower odds of stunting in the adjusted model.

The association between household access to improved sanitation and childhood stunting are shown in Appendix Table 2. While improved household sanitation was significantly associated with childhood stunting in the unadjusted model (OR = 0.83; 95% CI: [0.71, 0.97]), improved household sanitation was no longer significant in the multivariable model.

Similar to the national model, community-level improved sanitation coverage was not significantly associated with childhood stunting in any of the subnational models (Appendix Table 4).

Table 5 **Association of community improved sanitation coverage and childhood stunting using multilevel logistic regression, Zambia**

Individual-level covariates	Unadjusted OR	[95% CI]	Adjusted OR	[95% CI]
Child's age				
<6 months	Ref.	--	Ref.	--
6–11 months	1.42*	[1.06, 1.91]	1.39*	[1.04, 1.87]
12–23 months	3.33**	[2.55, 4.35]	3.38**	[2.57, 4.45]
24–59 months	2.78**	[2.22, 3.48]	2.83**	[2.24, 3.57]
Child's sex				
Male	Ref.	--	Ref.	--
Female	0.68**	[0.61, 0.77]	0.67**	[0.60, 0.76]
Household wealth quintile				
Lowest quintile	Ref.	--	Ref.	--
Second quintile	0.88	[0.74, 1.04]	0.87	[0.73, 1.03]
Middle quintile	0.67**	[0.55, 0.81]	0.62**	[0.50, 0.75]
Fourth quintile	0.61**	[0.49, 0.76]	0.49**	[0.37, 0.65]
Highest quintile	0.36**	[0.28, 0.47]	0.28**	[0.19, 0.40]
Household crowding				
Not crowded	Ref.	--	Ref.	--
Crowded	1.10	[0.95, 1.28]	1.03	[0.88, 1.20]
Household water source				
Unimproved	Ref.	--	Ref.	--
Improved	0.90	[0.77, 1.05]	1.02	[0.87, 1.21]
Cluster-level covariates				
Nightlights	1.00	[0.99, 1.00]	1.01	[1.00, 1.02]
Travel time to city	1.00	[1.00, 1.00]	1.00**	[1.00, 1.00]
Community improved sanitation coverage	0.98	[0.96, 1.00]	1.00	[0.98, 1.03]
Residence				
Urban	Ref.	--	Ref.	--
Rural	1.29**	[1.13, 1.48]	0.86	[0.68, 1.09]
Region				
Central	Ref.	--	Ref.	--
Copperbelt	0.85	[0.65, 1.11]	0.94	[0.72, 1.23]
Eastern	1.03	[0.81, 1.32]	0.86	[0.67, 1.11]
Luapula	1.69**	[1.30, 2.21]	1.47**	[1.13, 1.93]
Lusaka	1.12	[0.87, 1.45]	1.26	[0.94, 1.71]
Muchinga	1.03	[0.77, 1.39]	0.89	[0.66, 1.21]
Northern	1.84**	[1.41, 2.41]	1.56**	[1.18, 2.07]
North Western	1.02	[0.74, 1.41]	1.00	[0.73, 1.37]
Southern	0.80	[0.56, 1.14]	0.90	[0.62, 1.32]
Western	0.82	[0.62, 1.08]	0.77	[0.56, 1.04]
Observations			7,854	
Number of groups			497	

* Significant at $p < .05$; ** significant at $p < .01$; OR = odds ratio; CI = confidence interval

4 DISCUSSION AND CONCLUSIONS

In this report, we presented two ways to visualize community-level improved sanitation coverage using the DHS cluster-level data. We then used the cluster-level data to produce modeled surfaces of community-level improved sanitation coverage for both Nigeria and Zambia. We found that community-level improved sanitation coverage was a significant predictor of childhood stunting for the Nigeria 2018 DHS, but this indicator was not significant for the Zambia 2018 DHS dataset. Access to improved household sanitation was significantly associated with childhood stunting in the unadjusted models, but was not a significant predictor in the adjusted models.

4.1 Visualization

In this report, we demonstrated two ways to visualize community improved sanitation coverage. The first map overlaid the raw cluster-level estimates on the regional estimates. This map is useful in showing the variation that exists between clusters, and highlighting that in some regions, the cluster-level data underlying the regional estimates are highly variable. This variation would not be apparent from the regional estimates alone. Better understanding of the geographic variation in estimates of demographic and health indicators is a major factor in the efforts to produce estimates of DHS indicators at lower subnational levels using geospatial modeling.[44,45]

Using clusters as the unit of measurement, the second map grouped clusters based on their community improved sanitation coverage, and showed the proportion of clusters in each group. In this map, we used the groups 0–20%, 21–40%, 41–60%, 61–80%, 81–99%, and 100%, although these groups could vary to reflect programmatic targets. To sustainably improve health outcomes, achieving complete sanitation coverage is an important target.[46] The maps that use the pie chart symbology allow users to quickly compare how regions are performing in terms of achieving complete, improved sanitation coverage. The goal would be for all pie charts to be filled in yellow, which would mean that every cluster in that particular region had an improved sanitation coverage of 100%. This map also more easily summarizes the variation in cluster-level estimates, which is difficult to see in the cluster-level maps (Figs 2a and 3a), particularly in urban areas where many of the clusters overlap due to the map scale.

From the maps, policymakers can visualize if regions are homogenous or have substantial variation in the cluster-level estimates, which may inform policy responses. In regions with substantial variation, policymakers may want to explore if any geographic patterns exist to help determine if targeting a particular geographic area within a region is necessary. In addition, in cases where there are areas of low improved sanitation crossing borders, national policymakers may find it more cost-effective to target geographic areas rather than administrative areas. Policymakers can also use the modeled surfaces and subnational administrative level 2 estimates as additional decision-making tools. Users interested in replicating these maps can find the shapefiles with the survey boundaries on the Survey Boundaries page of the SDR, the shapefiles with the regional estimates on the Indicator Data page of the SDR, and can register for access to obtain the GPS coordinates of the survey clusters.

4.2 Geospatial Modeling

The modeled surfaces produced through geospatial modeling of the cluster-level improved sanitation data estimate the percentage of the population with access to an improved sanitation facility in locations that were not surveyed (Figures 4a and 5a). The modeled surfaces for both Nigeria and Zambia reveal considerable variation that is concealed when exploring only the regional estimates.

In this report, we combined the modeled surfaces with the estimated population count raster in 2018 from WorldPop (Figures 4c and 5c) to estimate the population at risk in every 5 x 5 km pixel (Figures 4d and 5d). Here, the population at risk represents the count of individuals without access to improved sanitation for every 5 x 5 km pixel in the country. The population at risk estimates provide an additional decision-making tool for policymakers. Policymakers may reach more of the population at risk in certain areas by targeting urban areas with higher access to improved sanitation than targeting rural areas with relatively lower access to improved sanitation. However, the modeled surface of the population at risk is limited by estimating for every 5 x 5 km pixel, which may not be a policy-relevant area. Aggregating these estimates to more policy-relevant boundaries may be useful to better understand the total population at risk for different communities or local administrative divisions. This is a major driver underlying The DHS Program's recent focus on producing modeled estimates at subnational administrative level 2, referred to as admin 2 estimates.[44,45]

4.3 Multilevel Regression

In this report, we found that while household access to improved sanitation was significant in the unadjusted models for Nigeria and Zambia, this variable was not significant in the adjusted models for both countries. These findings are not surprising given the inconsistent findings in the literature, although a recent meta-analysis suggested that 72% of studies found an association between household sanitation and childhood stunting.[8]

Although the biological mechanisms between sanitation and stunting are established,[5] stunting is multifactorial.[33] Other determinants, such as child's age, child's sex, and mother's stature may be more important drivers of childhood stunting.[47] In addition, the impact of household sanitation on childhood stunting may be reduced when children are exposed to fecal contamination in other locations, or in their own home by fecal pathogens brought in from humans, animals, and flies.[11]

This study found that community-level improved sanitation coverage was significant associated with stunting for Nigeria in the adjusted model (AOR=0.97; 95% CI [0.95, 0.99]), although not for Zambia. While many studies consider household sanitation, fewer studies have explored the association between stunting and community sanitation. Most of these studies have explored any sanitation access, unlike this study that evaluated improved access to sanitation. A meta-analysis of DHS and MICS data conducted by Larsen and colleagues found that access to community-level sanitation (either improved or unimproved) significantly reduced the odds of stunting and anemia in children.[48] Other studies have also found that community-level open defecation rates are important. One study found that community-level open defecation rates were inversely associated with child HAZ *z* scores in three of the four age groups studied for boys, and two of the four age groups studied for girls.[49] Another study found that children living in communities without open defecation are less likely to be stunted than children living in communities where all households practice open defecation.[50]

Our finding that community improved sanitation coverage was significantly associated with childhood stunting in Nigeria, but household improved sanitation access was not, suggests that in Nigeria, community-level improved sanitation coverage is a more important predictor than access to household improved sanitation in this context. Of the few studies that have evaluated both household and community sanitation, other studies have found that community sanitation coverage was a more important predictor than household sanitation access.[13,14] However, we only found that community improved sanitation coverage was significant associated with stunting in two of the six regions in Nigeria, which suggests that there may be regional variations in this pattern. However, since the adjusted odds ratios and confidence intervals were relatively similar, regional variation in these estimates may not be particularly strong (Appendix Table 3).

One limitation of this study which may explain the mixed results is that DHS data is only georeferenced to one location per cluster. Thus, the individual household coordinates are not available. Therefore, there is no information available on proximity to households that lack access to improved sanitation facilities. Studies have found that sanitation coverage of nearby households is important, and that the distance to households without sanitation may moderate the effect. One study conducted in rural Ecuador found that improved sanitation coverage in surrounding households was significantly associated with stunting, and that children with 100% improved sanitation coverage nearby were significantly less likely to be stunted than children with 0% sanitation coverage nearby.[13] Another study from rural Mali found that increased sanitation coverage within a 200 mile radius was associated with significant reductions in HAZ *z* scores.[14] A longitudinal study conducted in rural Bangladesh found that 100% sanitation coverage in all neighboring households within 50 miles was marginally associated with reduced diarrheal disease and acute respiratory infection prevalence, and that the effects were attenuated when evaluating neighboring households within 100 miles.[11] While proximity to households that practice open defecation may affect the association between community sanitation coverage and health outcomes related to poor sanitation, this study and others suggest that evaluating community sanitation coverage from survey data may still show an effect.

Another factor that may explain the inconsistent results in this study is our use of DHS cluster-level data as a proxy for community-level sanitation. All households are not surveyed in a cluster, and clusters may not always approximate communities. For an indicator like community sanitation coverage, in which few individuals without sanitation access can continue to contaminate the shared environment and maintain disease transmission,[52] ideally all households within specific distances would be surveyed. However, given that studies using DHS data have found that community sanitation was associated with poor health outcomes, including childhood stunting, anemia,[48] and several maternal health outcomes,[51] DHS data may still be useful in approximating community-level sanitation.

Additional analyses using this approach may help to better elucidate the association between community sanitation and health outcomes using DHS data. While this study focused on improved sanitation, there are more nuanced sanitation classifications that should be explored. The current sanitation classification ladder used by the JMP comprises five service levels, ranging from open defecation to safely managed sanitation services. Improved sanitation is split into three categories: limited, basic, and safely managed.[1] Additional analyses evaluating the association between community open defecation, as well as community basic sanitation coverage, and childhood stunting may be useful in understanding how the different sanitation levels affect childhood stunting. Additionally, evaluating the association between community sanitation and other health outcomes like diarrhea and anemia would help to better understand the interplay between sanitation and health outcomes in DHS surveys. Additionally, we only evaluated a single DHS survey from two countries in this study, while many additional survey countries and survey years are available. One study evaluating the impact of various indicators on nutritional changes over multiple rounds of DHS surveys found that sanitation improvements explained nutritional improvements in South Asian countries, but not in the African countries evaluated.[53]

4.4 Conclusions

While The DHS Program has long provided reliable estimates of demographic and health estimates at the national and subnational levels, DHS data has been further extended with geospatial modeling of cluster-level data, including the routine production of modeled surfaces at a 5 x 5km resolution starting in 2016 to the more recent production of subnational administrative level 2 estimates. This report demonstrates the value of considering cluster-level estimates of DHS data through the multilevel regression analysis of improved sanitation coverage and provides examples of how the geospatial modeled surfaces that estimate community-level prevalence can be used in practice. The maps and modeled surfaces presented in this report, along with the already available modeled surfaces on the SDR and any future subnational administrative level 2 estimates in DHS reports or on the SDR, can be used as decision-making tools to better understand how indicators vary geographically.

The users of DHS datasets who are interested in analyzing a health topic related to sanitation should consider exploring both household-level sanitation and community-level sanitation in their analyses. As demonstrated in this report, users can compute the community improved sanitation coverage by collapsing the data for each cluster and calculating the percentage of households or the percentage of the de jure population that is using an improved sanitation facility in each cluster. Although DHS surveys do not survey every household in a community and there is no information on household locations which make distance analyses or thresholds unavailable, the findings from this study and other studies that use DHS data,[48,51] suggest that DHS is still valuable in approximating community improved sanitation coverage.

Researchers who are analyzing communities where DHS surveys were not conducted may consider using estimates extracted from The DHS Program modeled surfaces, which are available on the SDR. The sanitation variable currently available as a standard indicator on the modeled surfaces page is the percentage of the de jure population that is practicing open defecation. Multiple reports have found an association between open defecation rates in a community and health outcomes linked to poor,[48-51] which suggests that this may be an important indicator of interest. Users of the modeled surfaces should be aware of the uncertainty that accompanies many of these modeled estimates. As seen in Figures 4b and 5b, the uncertainty in these modeled surfaces varies considerably, with some pixels having very little uncertainty and other pixels with very high uncertainty. Users should refer to Spatial Analysis Report 14 for additional guidance on the use of the modeled surfaces geospatial data.[54]

REFERENCES

1. World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). Progress on household drinking water, sanitation and hygiene 2000-2020: Five years into the SDGs. Geneva: World Health Organization (WHO) and the United Nations Children's Fund (UNICEF): 2021. https://data.unicef.org/resources/progress-on-household-drinking-water-sanitation-and-hygiene-2000-2020/

2. Mara D, Lane J, Scott B, Trouba D. Sanitation and health. *PLoS Medicine.* 2010;7(11):e1000363. https://doi.org/10.1371/journal.pmed.1000363

3. He Z, Bishwajit G, Zou D, Yaya S, Cheng, Z, Zhou Y. Burden of common childhood diseases in relation to improved water, sanitation, and hygiene (WASH) among Nigerian children. *Int J Environ Res Public Health.* 2018;5(6):1241. https://doi.org/10.3390/ijerph15061241

4. Shrestha, A, Six J, Dahal D, Marks S, Meierhofer R. Association of nutrition, water, sanitation and hygiene practices with children's nutritional status, intestinal parasitic infections and diarrhoea in rural Nepal: A cross-sectional study. *BMC Public Health*. 2020;20(1):1–21. https://doi.org/10.1186/s12889-020-09302-3

5. Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting? Current evidence and policy implications. *Matern Child Nutrit.* 2016;12:91–105. https://doi.org/10.1111/mcn.12258

6. Wolf J, Johnston RB, Ambelu A, Arnold BF, Bain R, Brauer M, Brown J, Caruso BA, Clasen T, Colford JM, Mills JE, et al. Burden of disease attributable to unsafe drinking water, sanitation, and hygiene in domestic settings: A global analysis for selected adverse health outcomes. *Lancet*. 2023; 401(10393):2060–71. https://doi.org/10.1016/S0140-6736(23)00458-0

7. Freeman MC, Garn JV, Sclar GD, Boisson S, Medlicott K, Alexander KT, Penakalapati G, Anderson D, Mahtani AG, Grimes JE, Rehfuess EA. The impact of sanitation on infectious disease and nutritional status: A systematic review and meta-analysis. *Int J Hyg Envir Health*. 2017; 220(6):928–949.

8. Mudadu Silva JR, Vieira LL, Murta Abreu AR, de Souza Fernandes E, Moreira TR, Dias da Costa G, Mitre Cotta RM. Water, sanitation, and hygiene vulnerability in child stunting in developing countries: A systematic review with meta-analysis. *Public Health*. 2023; 219:117–123. https://doi.org/10.1016/j.puhe.2023.03.024

9. Adhikari RP, Shrestha ML, Acharya A. et al. Determinants of stunting among children aged 0–59 months in Nepal: Findings from Nepal Demographic and Health Survey, 2006, 2011, and 2016. *BMC Nutr*. 2019;5:37. https://doi.org/10.1186/s40795-019-0300-0

10. Bekele T, Rawstorne P, Rahman B. Trends in child growth failure among children under five years of age in Ethiopia: Evidence from the 2000 to 2016 Demographic and Health Surveys. *PLOS One*. 2021;16(8):e0254768. https://doi.org/10.1371/journal.pone.0254768

11. Contreras JD, Islam M, Mertens A, Pickering AJ, Kwong LH, Arnold BF, Benjamin-Chung J, et al. Influence of community-level sanitation coverage and population density on environmental fecal contamination and child health in a longitudinal cohort in rural Bangladesh. *Int J Hyg Envir Health.* 2022;245:114031. https://doi.org/10.1016/j.ijheh.2022.114031

12. Julian TR. Environmental transmission of diarrheal pathogens in low and middle income countries. *Environ Sci: Process Impacts*. 2016;18(8):944–55. https://doi.org/10.1039/C6EM00222F

13. Fuller JA, Villamor E, Cevallos W, Trostle J, Eisenberg JN. I get height with a little help from my friends: Herd protection from sanitation on child growth in rural Ecuador. *Int J Epidemiol*. 2016: 45(2):460–469. https://doi.org/10.1093/ije/dyv368

14. Harris M, Alzua ML, Osbert N, Pickering A. Community-level sanitation coverage more strongly associated with child growth and household drinking water quality than access to a private toilet in rural Mali. *Environ Sci Technol.* 2017;51(12):7219–27. https://doi.org/10.1021/acs.est.7b00178

15. Pickering AJ, Null C, Winch PJ, Mangwadu G, Arnold BF, Prendergast AJ, Njenga SM, Rahman M, Ntozini R, Benjamin-Chung J, Stewart CP et al. The WASH Benefits and SHINE trials: Interpretation of WASH intervention effects on linear growth and diarrhoea. *Lancet Glob Health*. 2019;7(8):e1139–46. https://doi.org/10.1016/S2214-109X(19)30268-2

16. ICF International. *Demographic and Health Survey Sampling and Household Listing Manual.* ICF:2012.

17. de Onis M, Borghi E, Arimond M, Webb P, Croft T, Saha K, De-Regil LM. et al. Prevalence thresholds for wasting, overweight and stunting in children under 5 years. *Public Health Nutr.* 2019;22(1):175–179. https://doi.org/10.1017/S1368980018002434

18. Deshpande A, Miller-Petrie MK, Lindstedt PA, Baumann MM, Johnson KB, Blacker BF, Abbastabar H, et al. Mapping geographical inequalities in access to drinking water and sanitation facilities in low-income and middle-income countries, 2000–17. *Lancet Glob Health*. 2020;8(9):e1162-e1185. https://doi.org/10.1016/S2214-109X(20)30278-3

19. Reiner RC, Wiens KE, Deshpande A, Baumann MM, Lindstedt PA, Blacker BF, Troeger CE, Earl L, Munro SB, Abate D, Abbastabar H. Mapping geographical inequalities in childhood diarrhoeal morbidity and mortality in low-income and middle-income countries, 2000–17: Analysis for the Global Burden of Disease Study 2017. *Lancet*. 2020;396(10246):238. https://doi.org/10.1016/S0140-6736(20)31248-4

20. Mosser JF, Gagne-Maynard W, Rao PC, Osgood-Zimmerman A, Fullman N, Graetz N, Burstein R, et al. Mapping Diphtheria-pertussis-tetanus Vaccine Coverage in Africa, 2000–2016: A Spatial and Temporal Modelling Study. *Lancet*. 2019; 393(10183):1843–1855. https://doi.org/10.1016/S0140-6736(19)30226-0

21. Sartorius B, VanderHeide JD, Yang M, Goosmann EA, Hon J, Haeuser E, Cork MA, Perkins S, Jahagirdar D, Schaeffer LE, Serfes AL, et al. Subnational mapping of HIV incidence and mortality among individuals aged 15–49 years in sub-Saharan Africa, 2000–18: A modelling study. *Lancet HIV*. 2021; 8(6)e363-e375. https://doi.org/10.1016/S2352-3018(21)00051-5

22. Alegana VA, Atkinson PM, Pezzulo C, Sorichetta A, Weiss D, Bird T, Erbach-Schoenberg E, Tatem AJ. Fine Resolution Mapping of Population Age-structures for Health and Development Applications. *J R Soc Interface*. 2015; 12(105): 20150073. https://doi.org/10.1098/rsif.2015.0073.

23. Gething P, Tatem A, Bird T, Burgert-Brucker C. *Creating Spatial Interpolation Surfaces with DHS Data*. DHS Spatial Analysis Reports No. 11. ICF:2015. https://dhsprogram.com/pubs/pdf/SAR11/SAR11.pdf.

24. Osgood-Zimmerman A, Millear AI, Stubbs RW, Shields C, Pickering BV, Earl L, Graetz N, et al. Mapping Child Growth Failure in Africa between 2000 and 2015. *Nature*. 2018; 555(7684):41–47. http://dx.doi.org/10.1038/nature25760.

25. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: 2023. https://www.R-project.org/

26. Bhatt S, Cameron E, Flaxman SR, Weiss DJ, Smith DL, Gething PW. Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J R Soc Interface.* 2017; 14(134). https://doi.org/10.1098/rsif.2017.0520

27. Wood SN. *Generalized Additive Models: An Introduction with R*. 2nd Edition. Chapman and Hall/CRC; 2017. https://doi.org/10.1201/9781315370279

28. Zou H, Hastie T. Regularization and variable Selection via the elastic net. *J R Stat Soc Series B Stat Methodol*. 2005; 67(2):301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

29. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat.* 2011;29(5):1189–1232. https://doi.org/10.1214/aos/1013203451.

30. Banerjee S, Carlin, BP, Gelfand AE. *Hierarchical Modeling and Analysis for Spatial Data*. 2nd ed. Chapman and Hall/CRC; 2014. https://doi.org/10.1111/biom.12290

31. Diggle PJ, Giorgi E. *Model-based Geostatistics for Global Public Health: Methods and Applications*. Chapman and Hall/CRC; 2019. https://doi.org/10.1201/9781315188492

32. Rue H, Martino S, Chopin N. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *J R Stat Soc Series B Stat Methodol*. 2009;71(2):319. https://doi.org/10.1111/j.1467-9868.2008.00700.x.

33. Stewart CP, Iannotti L, Dewey KG, Michaelsen KF, Onyango AW. Contextualising complementary feeding in a broader framework for stunting prevention. *Matern Child Nutrition*. 2013;9:27–45. https://doi.org/10.1111/mcn.12088

34. Rutstein SO, Johnson K. *The DHS Wealth Index*. DHS Comparative Reports No. 6. ORC Macro: 2004. http://dhsprogram.com/pubs/pdf/CR6/CR6.pdf

35. Rutstein, SO. *The DHS Wealth Index: Approaches for Rural and Urban Areas*. DHS Working Papers No. 60. Macro International; 2008. http://dhsprogram.com/pubs/pdf/WP60/WP60.pdf

36. World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). *JMP Methodology: 2017 Update & SDG Baselines*. Geneva: WHO and UNICEF: 2018. https://washdata.org/sites/default/files/documents/reports/2018-04/JMP-2017-update-methodology.pdf

37. Mayala BK, Fish TD, Eitelberg, D, Dontamsetti T. *The DHS Program Geospatial Covariate Datasets Manual*. 2nd ed. ICF:2018. https://spatialdata.dhsprogram.com/references/DHS%20Covariates%20Extract%20Data%20Description%202.pdf

38. Amoako Johnson, F. Spatiotemporal clustering and correlates of childhood stunting in Ghana: Analysis of the fixed and nonlinear associative effects of socio-demographic and socio-ecological factors. *PLoS One*. 2022;17(2):e0263726. https://doi.org/10.1371/journal.pone.0263726

39. Amare, M, Arndt C, Abay KA, Benson T. Urbanization and child nutritional outcomes. *World Bank Econ Rev*. 2020;34(1):63–74. https://doi.org/10.1093/wber/lhy015

40. Islam MR, Alam M, Afzal MN, Alam S. Nighttime light intensity and child health outcomes in Bangladesh. *SN Bus Econ*. 2023;3(9):177. https://doi.org/10.1007/s43546-023-00556-8

41. Aoun N, Matsuda H, Sekiyama M. Geographical accessibility to healthcare and malnutrition in Rwanda. *Soc Sci Med*. 2015;130:135–45. https://doi.org/10.1016/j.socscimed.2015.02.004

42. Riese S, Assaf, S, Edmeades J. *Collective Gender and Fertility Norms and Modern Contraceptive Use*. DHS Analytical Studies No. 82. ICF:2022. https://dhsprogram.com/pubs/pdf/AS82/AS82.pdf

43. Elkasabi M, Ren R, Pullum TW. Multilevel modeling using DHS surveys: A framework to approximate level-weights. DHS Methodological Reports No. 27. ICF:2020. https://www.dhsprogram.com/pubs/pdf/MR27/MR27.pdf

44. Mayala BK, Dontamsetti T, Fish TD, Croft TN. *Interpolation of DHS Survey Data at Subnational Administrative Level 2*. DHS Spatial Analysis Reports No. 17. ICF: 2019. https://dhsprogram.com/pubs/pdf/SAR17/SAR17.pdf

45. Janocha B, Donohue RE, Fish TD, Mayala BK, Croft TN. *Guidance and recommendations for the use of indicator estimates at subnational administrative level 2.* DHS Spatial Analysis Reports No. 20. ICF: 2021. https://dhsprogram.com/pubs/pdf/SAR20/SAR20.pdf

46. Andres L, Briceño B, Chase C, Echenique JA. Sanitation and externalities: evidence from early childhood health in rural India. *J Water Sanit Hyg Dev*. 2017;7(2):272–89. https://doi.org/10.2166/washdev.2017.143

47. Akombi BJ, Agho KE, Hall JJ, Wali N, Renzaho AM, Merom D. Stunting, wasting and underweight in sub-Saharan Africa: a systematic review. *Int J Environ Res Public Health*. 2017;14(8):863 https://doi.org/10.3390/ijerph14080863.

48. Larsen DA, Grisham T, Slawsky E, Narine L. An individual-level meta-analysis assessing the impact of community-level sanitation access on child stunting, anemia, and diarrhea: Evidence from DHS and MICS surveys. *PLoS Negl Trop Dis*. 2017;11(6):e0005591. https://doi.org/10.1371/journal.pntd.0005591

49. Chakrabarti S, Singh P, Bruckner T. Association of poor sanitation with growth measurements among children in India. *JAMA Netw Open.* 2020;3(4):e202791. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2764400

50. Cameron L, Chase C, Haque S, Joseph G, Pinto R, Wang Q. Childhood stunting and cognitive effects of water and sanitation in Indonesia. *Econ Hum Biol.* 2021;40:100944. https://doi.org/10.1016/j.ehb.2020.100944

51. Kmush BL, Walia B, Neupane A, Frances C, Mohamed IA, Iqbal M, Larsen DA. Community-level impacts of sanitation coverage on maternal and neonatal health: A retrospective cohort of survey data. *BMJ Glob Health.* 2021;6(10):e005674. http://dx.doi.org/10.1136/bmjgh-2021-005674

52. Ziegelbauer K, Speich B, Mäusezahl D, Bos R, Keiser J, Utzinger J. Effect of sanitation on soil-transmitted helminth infection: systematic review and meta-analysis. *PLoS Med*. 2012;9(1):e1001162. https://doi.org/10.1371/journal.pmed.1001162

53. Headey D, Hoddinott J, Park S. Accounting for nutritional changes in six success stories: a regression-decomposition approach. *Glob Food Sec*. 2017;13:12–20. https://doi.org/10.1016/j.gfs.2017.02.003

54. Burgert-Brucker, CR, Dontamsetti T, Marshall A, and Gething P. *Guidance for Use of The DHS Program Modeled Map Surfaces*. DHS Spatial Analysis Reports No. 14. ICF:2016.

APPENDIX

Appendix Table 1 Association of household access to improved sanitation and childhood stunting using multilevel logistic regression, Nigeria

Individual-level covariates	Unadjusted OR	[95% CI]	Adjusted OR	[95% CI]
Child's age				
<6 months			Ref.	--
6–11 months			1.36*	[1.01, 1.83]
12–23 months			3.83**	[2.95, 4.97]
24–59 months			4.31**	[3.36, 5.52]
Child's sex				
Male			Ref.	--
Female			0.73**	[0.65, 0.82]
Household wealth quintile				
Lowest quintile			Ref.	--
Second Quintile			0.86	[0.72, 1.03]
Middle quintile			0.72**	[0.58, 0.89]
Fourth Quintile			0.50**	[0.39, 0.64]
Highest quintile			0.29**	[0.21, 0.41]
Household crowding				
Not crowded			Ref.	--
Crowded			1.05	[0.94, 1.17]
Household water source				
Unimproved			Ref.	--
Improved			0.90	[0.78, 1.04]
Household sanitation				
Unimproved	Ref.	--	Ref.	--
Improved	0.74**	[0.64, 0.85]	1.09	[0.94, 1.26]
Cluster-level covariates				
Nightlights			0.99	[0.97, 1.01]
Travel time to city			1.00	[1.00, 1.00]
Residence				
Urban			Ref.	--
Rural			1.06	[0.90, 1.25]
Region				
North Central			Ref.	--
North East			2.55**	[2.04, 3.21]
North West			4.14**	[3.41, 5.02]
South East			0.75*	[0.59, 0.96]
South South			0.84	[0.64, 1.08]
South West			1.23	[0.96, 1.57]
Observations			11,023	
Number of groups			1,348	

* Significant at *p* < .05; ** significant at *p* < .01; OR = odds ratio; CI = confidence interval

Appendix Table 2 Association of household access to improved sanitation and childhood stunting using multilevel logistic regression, Zambia

Individual-level covariates	Unadjusted OR	[95% CI]	Adjusted OR	[95% CI]
Child's age				
<6 months			Ref.	--
6–11 months			1.40*	[1.04, 1.88]
12–23 months			3.38**	[2.57, 4.45]
24–59 months			2.83**	[2.24, 3.57]
Child's sex				
Male			Ref.	--
Female			0.67**	[0.60, 0.76]
Household wealth quintile				
Lowest quintile			Ref.	--
Second quintile			0.87	[0.73, 1.04]
Middle quintile			0.62**	[0.51, 0.76]
Fourth quintile			0.50**	[0.37, 0.67]
Highest quintile			0.29**	[0.20, 0.41]
Household crowding				
Not crowded			Ref.	--
Crowded			1.03	[0.88, 1.20]
Household water source				
Unimproved			Ref.	--
Improved			1.03	[0.87, 1.21]
Household sanitation				
Unimproved	Ref.	--	Ref.	--
Improved	0.83*	[0.71, 0.97]	0.94	[0.79, 1.12]
Cluster-level covariates				
Nightlights			1.01	[1.00, 1.02]
Travel time to city			1.00**	[1.00, 1.00]
Residence				
Urban			Ref.	--
Rural			0.85	[0.67, 1.08]
Region				
Central			Ref.	--
Copperbelt			0.96	[0.73, 1.24]
Eastern			0.87	[0.68, 1.12]
Luapula			1.50**	[1.15, 1.96]
Lusaka			1.28	[0.95, 1.73]
Muchinga			0.91	[0.68, 1.23]
Northern			1.61**	[1.22, 2.12]
North Western			1.01	[0.74, 1.38]
Southern			0.92	[0.63, 1.35]
Western			0.75	[0.56, 1.01]
Observations			7,854	
Number of groups			497	

* Significant at $p < .05$; ** significant at $p < .01$; OR = odds ratio; CI = confidence interval

Appendix Table 3 **Association of community improved sanitation coverage and childhood stunting using multilevel logistic regression by region, Nigeria**

	North Central		North East		North West		South East		South South		South West	
	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]
Child's age												
<6 months	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
6–11 months	1.01	[0.55, 1.86]	2.25**	[1.26, 4.04]	1.71*	[1.09, 2.66]	1.30	[0.52, 3.25]	0.58	[0.24, 1.42]	0.77	[0.30, 1.95]
12–23 months	1.80*	[1.05, 3.06]	6.48**	[3.62, 11.63]	5.85**	[3.83, 8.92]	2.37*	[1.13, 4.97]	1.16	[0.56, 2.43]	2.37*	[1.17, 4.80]
24–59 months	2.43**	[1.48, 3.98]	6.84**	[3.94, 11.88]	8.16**	[5.62, 11.85]	2.03*	[1.02, 4.05]	1.17	[0.61, 2.26]	1.97	[0.99, 3.94]
Child's sex												
Male	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Female	0.71**	[0.57, 0.89]	0.91	[0.69, 1.19]	0.67**	[0.54, 0.82]	0.72*	[0.54, 0.97]	0.83	[0.57, 1.21]	0.72	[0.48, 1.09]
Mother's stature												
Not short	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Short	3.40*	[1.09, 10.61]	1.01	[0.41, 2.49]	6.04**	[2.82, 12.94]	6.15*	[1.47, 25.72]	0.85	[0.19, 3.72]		
Wealth quintile												
Lowest quintile	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Second quintile	0.77	[0.50, 1.18]	0.93	[0.65, 1.35]	0.97	[0.72, 1.29]	0.48	[0.20, 1.14]	0.73	[0.25, 2.18]	1.30	[0.67, 2.52]
Middle quintile	0.77	[0.48, 1.25]	0.59*	[0.37, 0.96]	0.82	[0.58, 1.15]	0.77	[0.29, 2.08]	0.74	[0.23, 2.41]	1.07	[0.50, 2.28]
Fourth quintile	0.43**	[0.24, 0.78]	0.63	[0.34, 1.16]	0.65	[0.41, 1.02]	0.50	[0.19, 1.32]	0.48	[0.15, 1.52]	0.72	[0.30, 1.70]
Highest quintile	0.36*	[0.16, 0.85]	0.36**	[0.17, 0.78]	0.39*	[0.19, 0.80]	0.34*	[0.12, 0.97]	0.25*	[0.07, 0.94]	0.42	[0.17, 1.04]
Household crowding												
Not crowded	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Crowded	0.80	[0.62, 1.04]	1.00	[0.75, 1.34]	1.20	[0.98, 1.47]	1.27	[0.94, 1.71]	0.89	[0.62, 1.27]	1.09	[0.78, 1.52]
Household water source												
Unimproved	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Improved	0.73	[0.50, 1.06]	1.21	[0.91, 1.60]	0.84	[0.65, 1.09]	0.73	[0.49, 1.09]	0.86	[0.55, 1.37]	1.50	[0.81, 2.75]
Residence												
Urban	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Rural	0.72	[0.41, 1.27]	1.32	[0.73, 2.38]	1.27	[0.93, 1.72]	0.88	[0.62, 1.26]	1.24	[0.77, 2.01]	0.86	[0.53, 1.41]
Cluster-level covariates												
Nightlights	0.97	[0.87, 1.09]	1.02	[0.97, 1.08]	0.98	[0.95, 1.02]	1.01	[0.90, 1.13]	1.01	[0.96, 1.06]	0.99	[0.96, 1.03]
Travel time to city	1.00	[0.99, 1.01]	1.00	[1.00, 1.00]	1.00	[0.99, 1.00]	1.01	[0.99, 1.03]	1.00	[1.00, 1.00]	1.00	[1.00, 1.01]
Improved sanitation coverage	0.98	[0.92, 1.05]	0.97	[0.92, 1.03]	0.95*	[0.90, 1.00]	0.98	[0.91, 1.05]	1.00	[0.94, 1.06]	0.94*	[0.88, 1.00]
Observations	1,976		1,800		2,695		1,643		1,265		1,639	
Number of groups	247		194		275		183		214		235	

* Significant at *p* < .05; ** significant at *p* < .01; aOR = adjusted odds ratio; CI = confidence interval

Appendix Table 4 Association of community improved sanitation coverage and childhood stunting using multilevel logistic regression by region, Zambia

	Central		Copperbelt		Eastern		Luapula		Lusaka	
	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]
Child's age										
<6 months	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
6–11 months	1.47	[0.69, 3.12]	1.62	[0.54, 4.85]	1.91	[0.83, 4.37]	2.99*	[1.31, 6.86]	0.67	[0.27, 1.65]
12–23 months	3.75**	[1.78, 7.86]	4.33**	[1.79, 10.47]	4.56**	[1.93, 10.78]	5.13**	[2.35, 11.22]	1.93	[0.84, 4.42]
24–59 months	2.02**	[1.21, 3.39]	2.39*	[1.16, 4.90]	3.60**	[1.59, 8.13]	4.31**	[2.21, 8.38]	2.09*	[1.11, 3.92]
Child's sex										
Male	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Female	0.74	[0.52, 1.06]	0.72*	[0.53, 0.98]	0.76	[0.55, 1.06]	0.79	[0.57, 1.10]	0.54**	[0.37, 0.79]
Wealth quintile										
Lowest quintile	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Second quintile	0.89	[0.62, 1.29]	0.43	[0.15, 1.25]	1.00	[0.65, 1.53]	0.95	[0.58, 1.55]	1.77	[0.70, 4.49]
Middle quintile	0.98	[0.56, 1.69]	0.47	[0.16, 1.37]	0.63*	[0.41, 0.97]	0.42*	[0.21, 0.86]	0.66	[0.28, 1.57]
Fourth quintile	0.57	[0.30, 1.07]	0.29*	[0.09, 0.95]	0.44	[0.19, 1.03]	0.37*	[0.16, 0.82]	0.70	[0.24, 2.01]
Highest quintile	0.33*	[0.12, 0.86]	0.20*	[0.06, 0.71]	0.23**	[0.09, 0.59]	0.44	[0.12, 1.56]	0.41	[0.13, 1.29]
Household crowding										
Not crowded	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Crowded	0.90	[0.55, 1.48]	1.40	[0.68, 2.88]	1.15	[0.82, 1.61]	0.78	[0.57, 1.08]	1.13	[0.79, 1.63]
Household water source										
Unimproved	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Improved	0.68	[0.41, 1.13]	0.95	[0.59, 1.53]	1.34	[0.80, 2.22]	1.24	[0.81, 1.91]	1.50	[0.50, 4.52]
Residence										
Urban	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Rural	0.41*	[0.20, 0.84]	1.56	[0.80, 3.03]	0.50	[0.23, 1.10]	0.64	[0.30, 1.36]	1.14	[0.60, 2.17]
Cluster-level covariates										
Nightlights	1.01	[0.95, 1.07]	1.00	[0.96, 1.05]	1.02	[0.95, 1.08]	1.04	[0.87, 1.24]	1.00	[0.99, 1.02]
Travel time to city	1.00	[1.00, 1.00]	0.99**	[0.98, 1.00]	1.00	[1.00, 1.00]	1.00	[1.00, 1.00]	1.00	[0.99, 1.00]
Improved sanitation coverage	0.95	[0.85, 1.06]	0.98	[0.88, 1.09]	1.01	[0.96, 1.06]	0.86	[0.72, 1.03]	1.04	[0.95, 1.15]
Observations	782		815		1,042		974		884	
Number of groups	51		60		62		53		63	

Continued...

Appendix Table 4—*Continued*

	Muchinga		Northern		North Western		Southern		Western	
	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]	aOR	[95% CI]
Child's age										
<6 months	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
6–11 months	0.95	[0.36, 2.55]	1.23	[0.57, 2.65]	1.40	[0.63, 3.14]	1.75	[0.13, 23.40]	0.98	[0.43, 2.23]
12–23 months	1.95	[0.96, 3.94]	3.47**	[1.86, 6.45]	2.42	[0.97, 6.02]	7.24	[0.57, 91.15]	2.91*	[1.26, 6.74]
24–59 months	3.02**	[1.56, 5.82]	3.61**	[1.97, 6.61]	2.28*	[1.05, 4.95]	7.19	[0.90, 57.17]	2.09*	[1.04, 4.21]
Child's sex										
Male	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Female	0.55*	[0.35, 0.87]	0.62*	[0.42, 0.90]	0.68	[0.43, 1.08]	0.50*	[0.25, 0.99]	0.75	[0.55, 1.02]
Wealth quintile										
Lowest quintile	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Second quintile	0.94	[0.56, 1.58]	0.78	[0.47, 1.30]	0.71	[0.39, 1.30]			0.59*	[0.39, 0.90]
Middle quintile	0.99	[0.54, 1.82]	0.49**	[0.29, 0.83]	0.36**	[0.19, 0.70]	2.83	[0.52, 15.35]	0.40**	[0.22, 0.70]
Fourth quintile	0.92	[0.33, 2.58]	0.61	[0.20, 1.87]	0.38	[0.12, 1.22]	2.20	[0.62, 7.81]	0.36	[0.11, 1.24]
Highest quintile	0.13**	[0.03, 0.50]	0.36	[0.12, 1.13]	0.09**	[0.03, 0.34]			0.08**	[0.02, 0.32]
Household crowding										
Not crowded	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Crowded	1.10	[0.64, 1.88]	0.91	[0.56, 1.48]	0.73	[0.53, 1.01]	1.16	[0.44, 3.09]	1.07	[0.71, 1.60]
Household water source										
Unimproved	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Improved	1.93*	[1.16, 3.22]	0.81	[0.55, 1.21]	0.92	[0.55, 1.55]	1.52	[0.12, 18.69]	0.68	[0.42, 1.11]
Residence										
Urban	Ref.	--	Ref.	--	Ref.	--	Ref.	--	Ref.	--
Rural	3.17	[0.79, 12.74]	0.84	[0.46, 1.54]	0.83	[0.27, 2.57]			0.96	[0.34, 2.76]
Cluster-level covariates										
Nightlights	1.20*	[1.02, 1.41]	0.80	[0.59, 1.07]	1.01	[0.86, 1.17]	0.99	[0.90, 1.09]	0.98	[0.81, 1.19]
Travel time to city	1.00	[0.99, 1.00]	1.00	[1.00, 1.00]	1.00**	[0.99, 1.00]	1.01	[0.99, 1.02]	1.00	[1.00, 1.00]
Improved sanitation coverage	0.93	[0.84, 1.04]	1.03	[0.98, 1.08]	1.05	[0.91, 1.22]	1.05	[0.82, 1.34]	1.39	[0.89, 2.16]
Observations	796		871		731		217		741	
Number of groups	45		52		43		19		49	

* Significant at $p < .05$; ** significant at $p < .01$; aOR = adjusted odds ratio; CI = confidence interval