



The Impact of Residential Location on the Main Outcomes from the Lungwena Antenatal Intervention Study (LAIS) Trial: Birth Outcomes, Child Growth, Mortality, Morbidity, and Development

Basho Poelman, University of Tampere, Finland Per Ashorn, University of Tampere, Finland Ulla Ashorn, University of Tampere, Finland Lotta Hallamaa, University of Tampere, Finland The LAIS Study Team

June 2018

FANTA FHI 360 I825 Connecticut Ave., NW Washington, DC 20009-5721 Tel: 202-884-8000 Fax: 202-884-8432 fantamail@fhi360.org www.fantaproject.org



This report is made possible by the generous support of the American people through the support of the Office of Health, Infectious Diseases, and Nutrition, Bureau for Global Health, U.S. Agency for International Development (USAID), under terms of Cooperative Agreement No. AID-OAA-A-12-00005, through the Food and Nutrition Technical Assistance III Project (FANTA), managed by FHI 360.

The contents are the responsibility of FHI 360 and do not necessarily reflect the views of USAID or the United States Government.

June 2018

Recommended Citation

Poelman, B. et al. 2018. *The Impact of Residential Location on the Main Outcomes from the Lungwena Antenatal Intervention Study (LAIS) Trial: Birth Outcomes, Child Growth, Mortality, Morbidity, and Development.* Washington, DC: FHI 360/Food and Nutrition Technical Assistance III Project (FANTA).

Contact Information

Food and Nutrition Technical Assistance III Project (FANTA) FHI 360 1825 Connecticut Avenue, NW Washington, DC 20009-5721 T 202-884-8000 F 202-884-8432 fantamail@fhi360.org www.fantaproject.org

Acknowledgments

We would like to thank Emma Kortekangas for her input into the early stages of the research process and Megan Deitchler for her comments on the research process and final text.

Contents

Ack	now	ledgments	i
Abb	orevi	ations and Acronyms	iii
Exe	cutiv	e Summary	1
1	Intr	oduction	2
2	Met	hods	3
	2.1	Study Objectives and Hypotheses Tested	4
	2.2	Study Design and Ethics Statement	5
	2.3	Participants, Interventions, and Randomization and Enrollment	5
	2.4	The Study Site	6
	2.5	Methods for Global Positioning System (GPS) Data Collection and GPS Data Handling	7
	2.6	Variable Descriptions	7
		2.6.1 Birth Outcomes	7
		2.6.2 Child Growth	8
		2.6.3 Child Mortality	8
		2.6.4 Non-Scheduled Visits	8
		2.6.5 Child Development	9
		2.6.6 Segmenting the Study Area	9
	2.7	Statistical Analyses	11
	2.8	Software	12
3	Res	ults	13
	3.1	Sample Formation	13
		3.1.1 Changes in Residential Location	13
	3.2	Objective I: Location as a Predictor	15
		3.2.1 Geographic Clustering of Birth Outcomes	15
		3.2.2 Mortality Outcomes	21
		3.2.3 Number of NS Visits	24
		3.2.4 Development	26
	3.3	Objective II: Interactions between Area, Cluster, and Outcome	28
		3.3.1 Preterm Birth	28
		3.3.2 Low Birth Weight	29
		3.3.3 Length-for-Age Z-Score	30
		3.3.4 Postneonatal Mortality	31
		3.3.5 Development	32
	3.4	Objective III: Adjustment of Outcomes for Area and Household-Level Cluster	33
4	Disc	cussion	35
5	Ref	erences	36
App	oendi	x 1. Included and Excluded Participants	37
App	oendi	x 2. Characteristics of Included Women by Intervention Group	38
App	oendi	x 3. Baseline Characteristics of Included Women by Relocation Status during Trial	39

Abbreviations and Acronyms

analysis of variance
body mass index
confidence interval
complete spatial randomness
global positioning system
human immunodeficiency virus
Lungwena Antenatal Intervention Study
length-for-age z-score
non-scheduled visit
standard deviation
sulfadoxine-pyrimethamine
Universal Transverse Mercator
World Health Organization
weight-for-length z-score

Executive Summary

Background. There are many interventions designed to combat child undernutrition. These interventions show mixed results. This heterogeneity could be the result of differences in study design or data collection, or they could be due to the influence of location.

Objective. The objective of this research was to see whether the heterogeneity found with different interventions can also be found within the outcomes of one trial, taking the influence of study design and data collection out of the equation.

Design. This research involved a secondary analysis of existing data from a completed clinical trial that tested the impact of maternal infection treatment in pregnancy on birth outcomes and child health in the first 5 years of life. We segmented the study area into four residential areas and tested for the influence of these areas on the outcomes of the trial. In addition, we conducted spatial clustering analyses using the Getis-Ord Gi* statistic in ArcGis software. We tested whether areas of "good" and "bad" outcomes could be identified, whether residential location modified the intervention effect, and whether considering residential location could improve the estimate of intervention effect.

Results. The results indicate that there are areas of "good" and "bad" child health outcomes within the study area. The results were not consistent in that different outcomes clustered in different areas. There was some evidence that residential location acted as a modifier and that including residential location into the analysis slightly improved the estimate of the intervention effect, although this evidence was also not consistent.

Conclusions. The heterogeneity in results found in child undernutrition interventions can also be found in the results from one trial, which eliminates study design and method of data collection as the cause of that heterogeneity.

1 Introduction

Healthy growth—along with the prevention of stunted growth and development—is a current focus of the global child health agenda (World Health Organization [WHO] 2017a), and malnutrition is the main health challenge globally. In childhood, undernutrition can present as wasting, stunting, or underweight and is associated with increased mortality, impaired development, and several adverse lifelong consequences. Stunting is a key indicator of chronic malnutrition. Globally, an estimated 155 million children under the age of 5 years suffered from stunting in 2016 (WHO 2017b).

To combat child undernutrition, various new interventions have been designed and implemented around the world. Some of those interventions have shown positive results and some negative, indicating heterogeneity in the results. This leads to the question of whether the geographical location influences the intervention effect or whether the differences can be explained by other factors such as variability in study design and data collection.

To investigate this issue, we conducted a secondary analysis on existing data from a completed clinical trial that tested the impact of maternal infection treatment in pregnancy on birth outcomes and child health in the first 5 years of life. We wanted to see if there would be geographical heterogeneity in the outcomes within one study sample, comprising children living within a relatively small area in Malawi's Mangochi District.

We developed three general hypotheses. The first stated that within the study area, there would be clusters of "good" and "bad" child health outcomes. To explore this hypothesis, we used a two-way approach. First, we segmented the study area into four geographic locales based on our knowledge of the area—and assumptions of what local factors might influence outcomes—and investigated whether this kind of segmentation shows variations in outcome clustering. Second, we used the Getis-Ord Gi* cluster statistic to check for clustering in outcomes. The second and third main hypotheses stated that location would modify the intervention effect and that including information on residential location would improve the estimate of the intervention effect.

2 Methods

We used outcomes from the Lungwena Antenatal Intervention Study (LAIS) conducted in Mangochi District in Malawi. This was a single-center, randomized three-arm clinical trial done by the University of Tampere to investigate whether maternal antibiotic treatment with monthly sulfadoxine-pyrimethamine (SP) and azithromycin during pregnancy would improve fetal growth and reduce the incidence of preterm birth. A follow-up to the original study looked at the number of a child's non-scheduled (NS) visits to a health center during the first 3 years of life, child mortality and growth during the first 5 years of life, and child development at 5 years of age.

The outcomes we selected covered five topics and provided information on the healthy growth and development context in Mangochi District. The first topic covered birth outcomes, where we considered duration of pregnancy, preterm delivery, weight-for-gestational-age, and low birth weight as indicators for prenatal maternal health and nutrition. The second topic covered child length/height-for-age and weight-for-height at 1, 12, 24, and 60 months of age as indicators of chronic and acute malnutrition. As a third topic, we looked at four mortality outcomes; death during the neonatal period (first 28 days of life, including abortions and stillbirths), postneonatal mortality, child mortality, and total mortality before 5 years of age. The fourth topic covered the number of a child's non-scheduled visits to a health center during the first 3 years of life as a proxy for morbidity and health-seeking behavior. Finally, as a fifth topic, we looked at child development by considering scores on the performance, language, and locomotor sub-scales and total score on the Griffiths Mental Development Scale.

We opted for a two-way approach to identifying geographic clusters of outcomes. First, we segmented the study area into four geographic locales based on our knowledge of the area and two assumptions on what might influence the outcomes considered. One assumption was that proximity to Lake Malawi could influence the study participants' diet by enriching it with fish, leading to better outcomes. The other assumption was that that the farther a participant lived from a health center, the worse the outcomes would be. We also chose the four locales because we thought that geographic division could be useful for future trials in the area.

Second, we conducted a spatial clustering analysis on the selected outcomes. We used the Hot Spot analysis tool in the ArcGis software to investigate whether there was household-level clustering of the outcomes. This tool uses the Getis-Ord Gi* statistic to locate clusters with a likelihood (confidence range) of having similar features. We compared clustering in different outcomes and tried to identify areas where bad outcomes cluster.

The second objective was to investigate whether the four selected residential areas or residential location of the child household cluster modified the impact of the LAIS intervention by analysing the interaction between residential area, household-level cluster location, and intervention. We selected five of the outcomes that in an earlier analysis had shown an association with the main intervention group. The selected variables were preterm birth, low birth weight, length-for-age z-score (LAZ)¹ at 24 months, total development score on the Griffiths scale, and postneonatal mortality.

For a third objective we used the same outcome variables selected for the second objective. We wanted to determine whether including the information on residential area or on possible household-level clustering would improve the estimate of the impact of the LAIS intervention.

¹ For simplicity and ease of reading, we use the terms weight-for-length z-score and length-for-age z-score for child growth results at all ages.

2.1 Study Objectives and Hypotheses Tested

The main aim of this study was to investigate whether the trial participants' residential location was associated with selected newborn and child health outcomes (birth outcomes, child growth, NS visits to the local health center [as a proxy for morbidity], child mortality, and child development).

To test this idea, we formulated three formal study objectives:

- 1. To determine whether the location of the child's home is a predictor of birth outcomes, child growth, NS visits to the local health center, child mortality, and child development in the LAIS trial
- 2. To determine if and how much the child's residential location modifies the impact of the LAIS intervention on selected newborn and child health outcomes
- 3. To determine if and how much including information on the child's residential location yields a better or more precise estimate of the LAIS intervention's impact on selected newborn and child health outcomes.

For the first objective, we both checked whether an a priori designed residential area yields differences in trial outcomes and investigated statistically significant household-level clustering of outcome values within the trial area.

For the second objective, we investigated whether associations between the study intervention and five selected outcomes from the trial were modified by the participant's residential area or by the participant's residential location as part of a "good" or "bad" cluster (household-level cluster). For the third objective, we examined whether including information on the residential area or household-level cluster improves the estimate of the LAIS intervention's impact.

These objectives led us to formulate the following hypotheses:

Objective I

Hypothesis 1—A predetermined areal segmentation of the study area shows statistically significant differences for selected birth outcomes, child growth, NS visits, mortality, and development at 5 years of age.

Hypothesis 2—Selected birth outcomes, child growth, NS visits, mortality, and development at 5 years of age in the study area would show clustering at the household level.

Objective II

Hypothesis 3—The impact of an intensified preventive infection treatment in pregnancy on preterm birth, low birth weight, length-for-age z-score at age 24 months, postneonatal mortality, and total development score at 5 years of age is modified by a predetermined areal segregation of the trial catchment area.

Hypothesis 4—The impact of an intensified preventive infection treatment in pregnancy on preterm birth, low birth weight, length-for-age z-score at age 24 months, postneonatal mortality, and total development score at 5 years of age is modified by household-level clusters.

Objective III

Hypothesis 5—The precision of the estimated impact of a preventive infection treatment in pregnancy on preterm birth, low birth weight, length-for-age z-score at age 24 months, postneonatal mortality, and total development score at 5 years of age is improved by a predetermined areal segregation of the trial catchment area.

Hypothesis 6—The precision of the estimated impact of a preventive infection treatment in pregnancy on preterm birth, low birth weight, length-for-age z-score at age 24 months, postneonatal mortality, and total development score at 5 years of age is improved by household-level clusters.

2.2 Study Design and Ethics Statement

These analyses were secondary analyses to the LAIS trial, a single-center, randomized, partially placebocontrolled, outcome assessor-blinded, three-arm clinical trial conducted in rural Malawi.

The original LAIS trial hypothesis was that maternal antibiotic treatment with monthly SP alone or in combination with two doses of azithromycin improves fetal growth and decreases the incidence of preterm delivery and that this leads to increased infant size at birth and at 1 month of age. The primary efficacy and safety outcome measures were the incidence of preterm delivery and serious adverse events. The main secondary outcomes were mean newborn size (weight, length, and head circumference) at birth and at 28 days of age as well as the prevalence of underweight, stunting, and wasting at 28 days of age.

A follow-up study of the LAIS trial assessed the number of a child's visits to a health center during the first 3 years of life, child mortality and growth during the first 5 years of life, and child development at 5 years of age. The hypotheses of the follow-up study were that children born to mothers who were treated during pregnancy with monthly SP, with or without two doses of azithromycin, would have fewer deaths during the first 5 years of life, fewer NS visits to the health center during the first 3 years of life, higher length/height at 2 and 5 years of age, and higher total development scores at 5 years of age than children born to mothers in the control group.

Both the original trial and the follow-up were performed according to Good Clinical Practice guidelines and the ethical standards of the Helsinki Declaration. The protocol was approved by the College of Medicine Research and Ethics Committee in Malawi and the Ethical Committee of Pirkanmaa Hospital District in Finland. An independent data safety and monitoring board monitored the incidence of suspected serious adverse events, conducted one site monitoring visit, and performed two interim analyses for safety and efficacy. Only participants who signed or thumb-printed an informed consent form were enrolled in the study. Key details of the protocol were published at the clinical trial registry of the National Library of Medicine in Bethesda, MD, United States (http://www.clinicaltrials.gov, trial identification NCT00131235).

2.3 Participants, Interventions, and Randomization and Enrollment

The study enrolled women with uncomplicated second trimester pregnancies (gestational age 14–26 weeks by ultrasound assessment) who had felt the movements of the fetus, were available during the follow-up period, and started antenatal care between December 2003 and October 2006 at Lungwena Health Center in southern Malawi. Exclusion criteria included severe illness, receipt of azithromycin during the current pregnancy or SP within the preceding 28 days, allergy to study drugs, and any previous serious allergic reaction. The last follow-up visit for children born to the women enrolled in the study was completed in June 2012.

Participants in the control group received standard Malawian antenatal care, which included intermittent preventive treatment in pregnancy with SP (three tablets orally, each containing 500 mg of sulfadoxine and 25 mg of pyrimethamine) twice: at enrollment and between 28–34 weeks of gestation. At these visits, they also received a placebo instead of azithromycin. Participants in the monthly SP intervention group (the SP group) received SP monthly from enrollment until 37 gestational weeks and a placebo instead of azithromycin. Participants in the AZI-SP group) received monthly SP and prophylactic treatment with active azithromycin (two tablets orally, each containing 500 mg of azithromycin) twice: at enrollment and between 28–34 weeks of gestation. All participants received

ferrous sulfate (200 mg/day) and folic acid (0.25 mg/day) throughout pregnancy (supplied by Malawi Central Medical Stores). SP tablets were purchased from Malawi Central Medical Stores, supplied by Pharmanova (Blantyre, Malawi), Ipca Laboratories Ltd. (Mumbai, India), F. Hoffmann-La Roche Ltd. (Basel, Switzerland), and Universal Corporation Kenya Ltd. (Kikuyu, Kenya). Active azithromycin and its placebo were manufactured and donated by Pfizer Inc. (New York, United States).

A researcher not involved in data collection generated a randomized code list. Based on this list, individual code slips with unique identification numbers, but not group allocation, were sealed in individual opaque randomized envelopes. For each identification number, an individual drug box was prepacked with the appropriate study drugs for each planned study visit in opaque drug envelopes labeled with the identification number and visit information

At the enrollment visit, research personnel interviewed pregnant women who expressed interest in the study about their socioeconomic status and obstetric history, provided them with counseling on HIV testing, and performed an antenatal examination. The duration of pregnancy was determined by measuring the fetal biparietal diameter and femur length with an ultrasound. Hadlock tables were used to calculate fetal age. A laboratory assistant with extensive experience assessed all participants for blood hemoglobin concentration, peripheral blood malaria parasitemia, and syphilis reactivity. HIV-positive participants were not excluded from enrollment. HIV testing was offered to all enrolled participants but was optional. At follow-up visits, HIV counseling was offered to HIV-tested participants, and HIV-infected participants received a 200-mg nevirapine tablet to prevent mother-to-child transmission.

Eligible persons who agreed to participate signed or thumb-printed an informed consent form and picked one envelope with an identification number. Each identification number was randomized to one of the three intervention groups. A research assistant who was not involved in outcome assessment gave the corresponding pre-packed study drugs to the participants under direct observation and monitored them for possible adverse reactions.

2.4 The Study Site

The trial area comprised a stretch of rural land on the eastern shore of Lake Malawi. Roughly, the area stretches 22 km from south to north and about 8 km from west to east, narrowing both toward the north and south end. It is bordered by Lake Malawi on the west side and the Namizimu Forest Reserve on the east, where the area's altitude increases. Further to the north, the Malawian territory eventually ends between Lake Malawi and the national border with Mozambique. Lungwena Health Center is located more or less in the middle of the area. There is one main road (S129) leading through the area that goes north. Other roads exist as well, but they start and end at the main road. Directly to the south of the trial area in the town of Malindi is St. Martin's Hospital. Further south is the district center of Mangochi with the district hospital (Figure 2-1).

In Malawi, an estimated 37% of children under 5 years of age were stunted, or too short for their age, in 2015–16. In Mangochi District, this proportion was estimated to be even higher, at 45%. Of all Malawian children under 5 years of age, 3% were wasted and 12% were underweight. Infant and under-5 mortality were 42 and 63 deaths per 1,000 live births, respectively, with children living in rural and poor environments facing the biggest mortality burden. The perinatal mortality rate was 35 deaths per 1,000 pregnancies (National Statistical Office (NSO) [Malawi] and ICF 2017).



Figure 2-1. Map of Study Site in Mangochi District, Malawi

2.5 Methods for Global Positioning System (GPS) Data Collection and GPS Data Handling

The University of Tampere usually records GPS locations when conducting trials in this research area. The area has no street names or house numbers, and the GPS coordinates are used to retrace participants' residential locations.

For the LAIS trial, the GPS location of each participant's home was recorded in November and December 2013 using a structured data collection form. A data collector visited each participant's home and while there recorded the coordinates given by a Garmin eTrex 10 GPS location device (Universal Transverse Mercator [UTM] coordinate system 'arc 1960 UTM Zone 36S') onto a data collection form. These data were later added to an Excel data workbook for storage and usage. UTM coordinates can be entered straight into the ArcGIS Desktop 10.5 package and used for analysis. The GPS dataset was checked for logical and recording errors.

2.6 Variable Descriptions

2.6.1 Birth Outcomes

For birth outcomes, we looked at duration of pregnancy, preterm birth, low birth weight, and weight-forgestational-age z-score. Duration of pregnancy was determined by measuring the fetal biparietal diameter and femur length with an ultrasound. Hadlock tables were used to calculate fetal age. We defined preterm birth as birth before 37 completed gestation weeks. Upon notification of a delivery, a research assistant visited the delivery site. She weighed the newborn with a spring scale (when at a home) or with electronic infant scale (when at a health facility). We included in the analyses birth weights measured within 2 days of delivery. We defined low birth weight as birth weight less than 2,500 grams. We calculated weight-for-gestational-age by entering birth weight and the duration of pregnancy into the INTERGROWTH-21st Neonatal Size Calculator for Windows (Villar et al. 2014).

2.6.2 Child Growth

Child growth was monitored by measuring length/height and weight at the study clinic visit at the ages of 3, 6, 9, 12, 15, 18, 24, 30, 36, 48, and 60 months. In this secondary analysis, we reported child growth only at ages 1, 12, 24, and 60 months. Anthropometrists measured length (for children \leq 24 months) with a Kiddimetre (Raven Equipment Ltd; reading increment of 1 mm) or height (for children > 24 months) with a Harpenden stadiometer (Holtain Limited; reading increment of 1 mm).² Weight was normally measured with a SECA scale (SECA 834, Chasmors Ltd; reading increment of 10 g). At age 36 months and older, weight was occasionally measured erroneously with a bathroom scale (1 kg increments). For the main analysis, the comparison of weight between the intervention groups, we did a sensitivity analysis where multiple imputation was used to replace weight measurements rounded to the full kilogram. These results were consistent with those from the primary analyses. Hence, we think the recording error is unlikely to have biased the findings of these results, either.

For measurements that were to be taken before 24 months of age, we considered the data missing if the actual measurement date was off by ± 4 weeks from the target date. For measurements that were to be taken at or after 24 months of age, we considered the data missing if the actual measurement date was off by ± 8 weeks from the target date.

We calculated age- and sex-standardized anthropometric indices (LAZ and weight-for-length z-score [WLZ]) by using World Health Organization (WHO) Child Growth Standards (WHO Multicentre Growth Reference Study Group 2006; de Onis et al. 2007). Only length-for-age z-score could be derived for children measured between 60–62 months given that there is no weight-for-length WHO reference for children 60 months of age and older (de Onis et al. 2007).

2.6.3 Child Mortality

Child mortality was assessed during the first 5 years of life. Information about child mortality was collected as soon as the study team heard about the death of the child or if the child did not show up for the scheduled study clinic visit for anthropometric measurements between 1 and 60 months of age. The study team traced the caretaker and collected mortality information if the child had died. This information was collected using a structured verbal autopsy questionnaire administered to the mother or another primary caretaker of the child.

For this study, we defined miscarriage as non-induced loss of pregnancy before 22.0 completed gestation weeks and stillbirth as fetal death at or after 22.0 gestation weeks. We included abortions and stillbirths in mortality during the child's first 28 days. Postneonatal death was defined as death between 29 days to 1 year of age, and child mortality was defined as death between 1–5 years of age. Total number of deaths was calculated as the sum of miscarriages, stillbirths, neonatal deaths, and child deaths occurring between the time of enrollment and 5 years after delivery.

2.6.4 Non-Scheduled Visits

We recorded the number of each participant's NS visits during the first 3 years of life both in total and for each year separately. We recorded information on child morbidity in real time when the child made a NS

visit to a health center. The information on NS visits was recorded on structured data collection forms by the clinician doing the examination and did not include normal under-5 clinic visits. We used the number of NS visits as a proxy for morbidity because we did not have direct data on child morbidity.

2.6.5 Child Development

When the children were 60 months of age, two trained and certified research assistants assessed their development with the *Griffiths Mental Development Scales—Extended Revised: 2–8 Years* (Luiz et al. 2006). These scales assess six areas of development:

- Locomotor—gross motor skills, including the ability to balance and to coordinate and control movements
- Personal-social—activities of daily living, level of independence, and interaction with other children
- Language—receptive and expressive language
- Eye and hand coordination—fine motor skills, manual dexterity, and visual monitoring skills
- Performance—visuospatial skills, including speed of working and precision
- Practical reasoning—ability to solve practical problems, understanding of basic mathematical concepts, and understanding of moral issues

For each child, we summed the scores of each assessment passed and used this as our main outcome. We calculated the sum of the passing scores of each sub-scale separately and used these as our secondary outcomes. In this report, we chose to present results only for the performance, language, and locomotor sub-scales. We selected the performance and language sub-scales for this analysis because earlier analyses had shown association between intervention and performance and possibly language sub-scales (Hallamaa et al. 2017). The locomotor sub-scale was selected because it measures a different dimension of development than performance or language.

2.6.6 Segmenting the Study Area

Areal Location

To get an idea of whether the *residential area* was associated with the trial's main outcomes, we segmented the study area into four geographic locales. Based on previous knowledge of the area, and its size, we felt that a four-way segmentation of the area could be practical and useful for future trials in the area. Four areas would also yield areal units large enough to understand the variability in the area. The area's main road was used as a natural border dividing the area into western and eastern parts. Areas 1, 2, and 3 are along the lake, and Area 4 is east of the road. We hypothesized that proximity to the lake might influence the local diet, type of employment, and possible household assets. The participants living near the main road were allocated to the areas that border the lake.

The area between the lake and the road was cut into three with Area 1 being the farthest north and farthest from the Lungwena Health Center. The Lungwena Health Center is in Area 2, and Area 3 is closer to St. Martin's Hospital, located in Malindi just south of the study area. Area 4 seems more isolated and has a higher altitude.

The four areas were drawn as polygon features, and the main road going through the trial area was a line feature in the ArcGIS software based on the World view areal satellite picture of the area.

Residential Location

After investigating the outcomes by residential area, we examined the outcomes by residential location, referred to as *household-level clustering* in this report. For this, we used the GPS coordinates collected for each participant's home. These were imported into the ArcMAP software and a file geodatabase was constructed.

We conducted a cluster analysis for each selected outcome by using the Hot Spot Analysis (Getis-Ord Gi*) tool³ in the ArcMap Desktop software. For the results reported here, we used the following tool settings: "Conceptualization of Spatial Relationships" based on "Spatial Weights From File," "Distance Method" set to "Euclidean," "Standardization" set to "None," and no false discovery rate correction applied. The environment settings were left at default. Other settings were also tried and essentially presented similar cluster outcomes, although the size of the clusters varied.

The Getis-Ord Gi* tool analyzes each residential location (of a participant) within the context of neighboring locations within the specified distance band. The Hot Spot Analysis tool returns a feature layer that includes a confidence level variable indicating whether the feature [participant] is part of a statistically significant cluster of high (hot spot) or low (cold spot) variable values. The value of a feature and its neighboring features within the specified distance band is compared to the average for the variable within the whole dataset. The null hypothesis states that the values are spread through the area in accordance with complete spatial randomness (CSR). Significant differences from CSR are reported as hot and cold spots, with an accompanying confidence level (at 90%, 95%, and 99% confidence). For the household-level cluster analyses in this report, we used the 90% confidence level as an indication that a participant's residential location is part of a cluster. We reported clusters of hot and cold spots based on a visual inspection of the output feature layer from the Hot Spot tool; the tool itself does not identify separate clusters.

Spatial Weights Matrix

An important setting for this tool is how the spatial relationships among the residential locations are defined. For this, we created a spatial weights matrix using the "Generate Spatial Weights Matrix" tool in the ArcMap Desktop 10.5 software. A spatial weights matrix is a representation of the spatial structure of the data and consists of an NxN (N is the number of features [participants] in the dataset) table with one row and one column for each feature. The cell value for any given row/column combination is the weight that quantifies the spatial relationship between those row and column features. We used the following settings while constructing the spatial weights matrix: "Fixed Distance Band" set to 1,000 meters, a minimum of eight neighbors, and "Row" set to "standardized."⁴

The distances between participants' residential locations were calculated using the Geographic Distance Matrix Generator software.⁵ Three participants in the set were located more than 1 km from their nearest neighbor. These participants were considered spatial outliers, and their influence was reduced by setting the distance band at 1,000 meters. The average distance to the nearest neighbor was 52 meters.

The output of the cluster analysis was exported to maps that show the four predefined residential areas. In the study area maps in the next section, clusters of outcomes that can be considered "good" are indicated

http://desktop.arcgis.com/en/arcmap/latest/tools/spatial-statistics-toolbox/generate-spatial-weights-matrix.htm.

⁵ Software v.1.2.3, downloaded on 5/18/2016 from http://biodiversityinformatics.amnh.org/open_source/gdmg/.

³ For background, see: <u>http://desktop.arcgis.com/en/arcmap/latest/tools/spatial-statistics-toolbox/hot-spot-analysis.htm</u>.

⁴ These settings mean that during the cluster analysis, each participant is compared to neighbors within a distance circle of 1,000 meters. If there are fewer than eight neighbors within the circle, the distance restriction is relaxed until there are eight neighbors. "Row standardized" means that the spatial weights are standardized by row—each weight is divided by its row sum (1/number of neighbors included in the cluster analysis for a particular participant). For further background, see:

in green while clusters of "bad" outcomes are in red. The meaning of low and high outcome values is ambiguous as, for example, a low value on a growth outcome indicates a "bad" outcome, while a low value on a mortality outcome indicates a "good" outcome.



Figure 2-2. Segmentation of the Study Area and Household Locations

2.7 Statistical Analyses

For Objective I, we calculated group means and used analysis of variance (ANOVA) to calculate differences among the four areas for continuous outcomes. We adjusted for multiple comparisons using Tukey's method. For dichotomous outcomes, we calculated percentages and used log binomial regression to estimate risk ratios among the four areas. We rejected the hypotheses of the groups being equivalent if P < 0.05.

For Objectives II and III, we used preterm birth, low birth weight, LAZ at 24 months, total development score, and postneonatal deaths as outcomes. These outcomes were selected because earlier analyses had shown association between them and the intervention groups (Luntamo et al. 2010; Hallamaa et al. 2017). In these analyses, we included only the control and AZI-SP groups because previous analyses showed differences between these two groups.

To determine whether the child's residential location modified the intervention's impact, we performed likelihood ratio tests for the interaction between the intervention and residential area or household-level cluster. When testing for interaction effects, we included the main effect of each term in one model and the main effect of each term and the interaction term in the other model. We considered the interaction statistically significant if P < 0.1; however, we performed the analyses stratified by area and household-level cluster regardless of the result of the likelihood ratio test.

To determine whether including the child's residential location yields a better or more precise estimate of the intervention's impact, we adjusted the analyses to account for area or household-level cluster. We calculated the difference in means and risk ratios unadjusted and adjusted for area or household-level cluster.

We used least squares regression to calculate differences in means for continuous outcomes and log binomial regression models to estimate risk ratios for binary endpoints. In case there were no positive observations in one of the groups for dichotomous outcomes, we used exact logistic regression to estimate odds ratios. We also calculated 95% confidence intervals (CI) for differences in means and risk ratios and tested the null hypothesis that the groups were equivalent. We rejected the hypotheses of the groups being equivalent if P < 0.05. All analyses were done without covariate adjustments.

2.8 Software

Geospatial analysis, map creation, and data storage in a file geodatabase were done with ArcGIS Desktop v.10.5 (Environmental Systems Research Institute [ESRI] 2017, Redlands, CA). A distance matrix was created using the Geographic Distance Matrix Generator (v. 1.2.3, downloaded on 18.05.2016 from http://biodiversityinformatics.amnh.org/open_source/gdmg/). Statistical analyses, other than spatial statistical analyses, were done with Stata 13.1 (StataCorp, College Station, TX, USA).

3 Results

3.1 Sample Formation

Of the 3,358 approached pregnant women, 1,320 were eligible and randomized into three intervention groups, with 436 in the control group, 441 in the monthly SP group, and 443 in the AZI-SP group. From this baseline sample, 1,241 participants were included into this analysis: 409 in the control group, 415 in the monthly SP group, and 417 in the AZI-SP group. The most common reasons for excluding participants from this analysis were duplicate enrollment in the study and missing GPS location (Figure 3-1). The included participants and the 79 excluded participants were similar in baseline characteristics except for a higher number of primiparous women among those included (24.1% vs. 8.9%, P < 0.001) (Appendix Table A1). At enrollment, the three intervention groups were comparable, except for small differences in the proportion of primiparous women (26.2%, 25.1%, and 21.1% in the control, monthly SP, and AZI-SP groups, respectively) and prevalence of malaria parasitemia (10.8%, 9.4%, and 6.0% in the control, monthly SP, and AZI-SP groups, respectively) (Appendix Table B1).

We obtained data on duration of pregnancy from 100.0% of participants and data on birth weight from 91.1%. On the number of NS visits during first 3 years of life, child mortality, child growth by 5 years of age, and child development at 5 years, we obtained data from 95.2%, 100.0%, 71.6%, and 71.7% of participants, respectively (Figure 3-1).

3.1.1 Changes in Residential Location

During the trial period, 69 mothers changed residential location. There were no differences in background characteristics between the group of mothers that changed residential location during the trial and those that did not, except that those who moved had a somewhat lower mean body mass index (BMI) at enrollment (21.3 vs. 21.8; P=0.039) (Appendix Table C1).



Figure 3-1. Participant Flow and Data Availability

3.2 Objective I: Location as a Predictor

3.2.1 Geographic Clustering of Birth Outcomes

For all participants, the mean (SD) duration of pregnancy was 38.6 (2.2) weeks, the mean (SD) weight-for-gestational-age z-score was -0.39 (1.00), the number (%) of preterm deliveries was 176/1,241 (14.2%), and the number (%) of infants with low birth weight was 101/1,130 (8.9%).

Within the four geographic areas, Area 4 had the longest mean duration of pregnancy and the lowest incidence of preterm birth and low birth weight. Area 3 showed the worst values for those outcomes but 3 had the highest weight-for-gestational-age z-score. Area 1 had mostly second "worst" birth outcomes. However, the differences were not statistically significant (Table 3-1).

Table 3-1.	Birth Outcomes, All Participants and By Area (Duration of Pregnancy, Weight-for-
	Gestational-Age, Preterm Delivery, and Low Birth Weight)

	Area 1	Area 2	Area 3	Area 4	P value*
Duration of pregnancy (weeks), mean (SD)	38.5 (2.5)	38.6 (2.2)	38.4 (2.0)	38.8 (2.0)	0.182
Weight-for-gestational-age z-score, mean (SD)	-0.41 (1.01)	-0.42 (0.98)	-0.28 (1.01)	-0.39 (1.02)	0.514
Preterm delivery (<37 weeks), n/N (%)	42/230 (15.4)	64/399 (13.8)	28/142 (16.5)	42/294 (12.5)	0.576
Low birth weight (<2,500 g), n/N (%)	20/227 (8.1)	40/382 (9.5)	17/133 (11.3)	24/287 (7.7)	0.572

* For continuous outcomes, *P* values were derived by ANOVA; for dichotomous outcomes, *P* values were derived by log binomial regression.

Notes: Green indicates "best" values; yellow second "best"; orange second "worst"; red "worst."; "n" refers to the number of participants with the condition in the noted Area, and "N" refers to the total number of participants in the noted Area

For the duration of pregnancy, there were a few small household-level clusters of "worse" outcomes in Areas 1, 2, and 3. Area 1 also showed a small cluster of "better" outcomes, and Area 4 had several clusters of "better" outcomes (Figure 3-2, panel A). For weight-for-gestational-age z-score, there was a clear household cluster of "worse" outcomes and a smaller cluster of "better" outcomes in Area 2, a cluster of "better" outcomes in Area 3, and clusters of both types in Area 4. There was a small cluster of "better" outcomes in Area 1. For preterm births, there were household clusters of "worse" outcomes in Area 1, small clusters of both "worse" and "better" values in Area 2, some very small clusters of "better" outcomes in Area 3, and a small cluster of "better" outcomes in Area 4 (Figure 3-2, panel C). For low birth weight, there was a cluster of "better" outcomes in Area 2, some clusters of "better" outcomes and a small cluster of "better" outcomes and a small cluster of "better" outcomes in Area 4 (Figure 3-2, panel C). For low birth weight, there was a cluster of "better" outcomes in Area 2, some clusters of "better" outcomes and a small cluster of "better" outcomes in Area 4 (Figure 3-2, panel C). For low birth weight, there was a cluster of "better" outcomes in Area 1, a clear cluster of "better" outcomes and a small cluster of "worse" outcomes in Area 2, some clusters of "better" outcomes and a small cluster of "worse" outcomes in Area 4 (Figure 3-2, panel C). For low birth weight, there was a cluster of "better" outcomes in Area 1, a clear cluster of "better" outcomes and a small cluster of "better" outcomes in Area 2, some clusters of "better" outcomes and a small cluster of "worse" outcomes in Area 2, some clusters of "worse" outcomes and a small cluster of "worse" outcomes in Area 2, some clusters of "worse" outcomes in Area 3, and clusters of both "better" and "worse" outcomes in Area 4 (Figure 3-2, panel D).





Child Length

The mean (SD) LAZ of children in the study cohort decreased from -1.28 (1.13) at 1 month of age to -1.69 (1.12) at 12 months and -2.10 (1.04) at 24 months and then increased to -1.67 (0.92) at 60 months.

Within the four areas, the mean LAZ was highest in Area 2 at all ages. The lowest values were typically observed in Area 3, second lowest in Area 1, and second highest in Area 4. At each age, the intergroup differences were statistically significant (Table 3-2).

	Area 1	Area 2	Area 3	Area 4	P value*
LAZ at 1 month, mean (SD)	-1.46 (1.13) ^a	-1.16 (1.11) ^b	-1.40 (1.19) ^{ab}	-1.26 (1.10) ^{ab}	0.007
LAZ at 12 months, mean (SD)	-1.79 (1.10) ^b	-1.54 (1.11)ª	-1.84 (1.12) ^b	-1.73 (1.12) ^{ab}	0.010
LAZ at 24 months, mean (SD)	-2.13 (1.03) ^{ab}	-1.95 (1.02)ª	-2.34 (1.13) ^b	-2.17 (0.99) ^b	0.001
LAZ at 60 months, mean (SD)	-1.78 (0.89)ª	-1.48 (0.93) ^b	-1.95 (1.02)ª	-1.69 (0.84)ª	<0.001

Table 3-2. LAZ by Area at 1, 12, 24, and 60 Months of Age

**P* values derived by ANOVA.

Notes: Groups that do not share a common superscript differ significantly from each other (p < 0.05).

Green indicates "best" values; yellow second "best"; orange second "worst"; red "worst."

For LAZ at 1 month, there was a clear household-level cluster of low values in Area 1 and clear cluster of high values in Area 2. There were also smaller household-level clusters of low values in Areas 1, 3, and 4, and clusters of high values in Areas 2 and 4 (Figure 3-3, panel A). For LAZ at 12 months, there were two clusters of high values in Area 2 and one household-level cluster of low values in Area 1. Areas 3 and 4 did not show clear clusters either way (Figure 3-3, panel B). For LAZ at 24 months, Area 1 showed small clusters of both high and low values. Area 2 had three household-level clusters of high values, and Area 3 had two clusters of low values. Area 4 did not have clusters of either (Figure 3-3, panel C). For LAZ at 60 months, Areas 1 and 3 had small clusters of low values, and Area 3 also had one bigger cluster of low values. Area 2 had three clusters of both types (Figure 3-3, panel D).





Weight-for-Length Z-Score

The mean (SD) WLZ among all participants was 0.57 (1.16), -0.13 (1.08), -0.25 (1.07), and -0.45 (1.04) at 1 month, 12 months, 24 months, and 60 months of child age, respectively.

There were no strong patterns observed for WLZ across areas over time. Area 1 had the highest mean WLZ at 1 month but had the lowest or second-lowest at later time points. While Area 2 had the lowest WLZ at 1 month, it had the highest or second-highest mean scores at the other time points. The second-lowest values were typically observed in Area 4. Area 3 had the second-highest WLZ values at 1 and 12 months and the highest at 24 and 60 months. The differences among the groups were statistically significant except at 12 months (Table 3-3).

	Area 1	Area 2	Area 3	Area 4	P value*
WLZ at 1 month, mean (SD)	0.70 (1.20) ^a	0.44 (1.19) ^b	0.69 (1.11) ^{ab}	0.59 (1.10) ^{ab}	0.017
WLZ at 12 months, mean (SD)	-0.15 (1.13)	-0.04 (1.04)	-0.13 (1.09)	-0.23 (1.09)	0.190
WLZ at 24 months, mean (SD)	-0.45 (1.16)ª	-0.20 (1.03) ^{bc}	0.01 (1.03) ^c	-0.29 (1.04) ^{ab}	0.001
WLZ at 60 months, mean (SD)	-0.52 (1.03) ^{ab}	-0.44 (1.02) ^{ab}	-0.22 (1.12) ^b	-0.52 (1.01)ª	0.048

Table 3-3. WLZ by Area at 1, 12, 24, and 60 Months of Child Age

**P* values derived by ANOVA.

Notes: Groups that do not share a common superscript differ significantly from each other (p < 0.05). Green indicates "best" values; yellow second "best"; orange second "worst"; red "worst."

For WLZ at 1 month, there was a clear household-level cluster of high WLZ values in Area 1 and low WLZ values in Area 2. There were small clusters of high WLZ in Areas 2, 3 and 4 and one small cluster of low values in Area 4 (Figure 3-4, panel A). At 12 months, all four areas had small household-level clusters of low values and Area 2 had two clusters of high values (Figure 3-4, panel B). At 24 months, there was one clear cluster and one smaller cluster of low values in Area 1 and smaller clusters of low values in Areas 2, 3, and 4. There were small clusters of high values in Areas 1, 2, and 3 (Figure 3-4, panel C). At 60 months, there was one bigger cluster of low WLZ values in Area 2 and smaller clusters in Areas 1 and 4. There were small clusters of high WLZ scores in Areas 3 and 4 (Figure 3-4, panel D).





Outcome calculated with ArcMap 10.5 Hot Spot An (Getis-Ord Gi* statistic).

4 Kilo

0 1 2







3.2.2 Mortality Outcomes

Among all included participants, the proportion of deaths during pregnancy and the first 5 years was 14.2%. During pregnancy and the first 28 days of age, the proportion was 6.4%. The proportion of children who died between the 29^{th} day and the end of their first year of life was 3.5%, and 4.7% died between their second and fifth year.

Total mortality during the first 5 years was lowest in Area 1, with the lowest proportion of deaths during the first 28 days and between 2–5 years; however, the area had the second-highest proportion of postneonatal mortality. Area 2 had the second-lowest total mortality, with the second-highest proportion of deaths during the first 28 days, the lowest proportion of postneonatal mortality, and the highest proportion of child mortality. Total mortality was second-highest in Area 4, with the highest proportion of deaths during the first 28 days and the second-lowest proportions of postneonatal and child mortality. Area 3 had the highest proportion of total mortality, with the second-lowest proportion of deaths during the first 28 days, the lowest proportion, with the second-lowest proportion of deaths during the first 28 days and the second-lowest proportions of postneonatal and child mortality. Area 3 had the highest proportion of total mortality, and the second-lowest proportion of deaths during the first 28 days, the lowest protection of total mortality. The differences among the areas were not statistically significant (Table 3-4).

Table 3-4.	Mortality Outcomes by Area (Total Mortality before 5 Years, Deaths before 28 Days,
	Postneonatal Mortality, Child Mortality)

	Area 1	Area 2	Area 3	Area 4	P value*
Total mortality before 5 years, n/N (%)	36/236 (13.2)	63/400 (13.6)	28/142 (16.5)	49/287 (14.6)	0.768
Death during first 28 days, n/N (%)**	15/257 (5.5)	31/432 (6.7)	10/160 (5.9)	23/313 (6.9)	0.911
Postneonatal mortality, n/N (%)	10/250 (3.9)	10/434 (2.3)	10/155 (6.1)	11/307 (3.5)	0.140
Child mortality, n/N (%)	11/249 (4.2)	22/422 (5.0)	8/157 (4.9)	15/303 (4.7)	0.978

*P values derived by log binomial regression.

**Including abortions and stillbirths.

Notes: Green indicates "best" values; yellow second "best"; orange second "worst"; red "worst."; "n" refers to the number of participants with the condition in the noted Area, and "N" refers to the total number of participants in the noted Area

There were some small household-level clusters of higher occurrence of death during the first 5 years in Areas 1, 2, and 3, and a cluster of lower occurrence in Area 4 (Figure 3-5, panel A). There were smaller clusters of higher mortality during the first 28 days in all four areas and two small clusters of lower mortality in Area 1 (Figure 3-5, panel B). There were clusters of higher occurrence of postneonatal mortality in Areas 1 and 3 and a small cluster of lower postneonatal mortality in Area 4 (Figure 3-5, panel C). Child mortality showed a small cluster of higher occurrence in Area 2, larger clusters of higher occurrence in Areas 3 and 4, and two small clusters of lower mortality in Area 4 (Figure 3-5, panel D).





3.2.3 Number of NS Visits

Among all participants, the mean (SD) number of NS visits to a health center during the first 3 years of life was 4.1 (3.6). The mean (SD) number of visits was 2.9 (2.7) during the first year of life, 0.9 (1.5) during the second year, and 0.4 (0.8) during the third year.

Area 1 had the fewest NS visits while Area 4 typically had the most. Area 2 generally had the secondhighest number of NS visits and Area 3 the second-lowest. Differences among groups were statistically significant, except for the number of NS visits during the third year (Table 3-5).

Table 3-5. Number of NS Visits to the Health Center in the First 3 Years and during the First, Second,and Third Years Separately

	Area 1	Area 2	Area 3	Area 4	P value*
NS visits during first 3 years, mean (SD)	2.7 (2.6)	4.6 (3.4) ^a	4.1 (3.7)ª	4.6 (4.1) ^a	<0.001
NS visits during first year, mean (SD)	1.8 (2.0)	3.3 (2.6) ^a	3.0 (3.0)ª	3.1 (3.0)ª	<0.001
NS visits during second year, mean (SD)	0.6 (1.0) ^a	1.0 (1.4) ^b	0.9 (1.4) ^{ab}	1.2 (1.8) ^b	<0.001
NS visits during third year, mean (SD)	0.3 (0.6)	0.4 (0.8)	0.4 (0.8)	0.5 (1.0)	0.055

**P* values derived by ANOVA.

Notes: Groups that do not share a common superscript differ significantly from each other (p < 0.05). Green indicates "best" values; yellow second "best"; orange second "worst"; red "worst."

There was a clear household-level cluster of fewer NS visits during the 3 years in Area 1, as well as smaller clusters in Areas 1 and 4. Areas 2, 3, and 4 had several clusters of more NS visits (Figure 3-6, panel A). For the number of NS visits during the first year, the clusters remained practically the same as the combined visits over 3 years (Figure 3-6, panel B). During the second year, there was a clear cluster of fewer of NS visits in Area 1. Area 4 had clear clusters of more NS visits, and Areas 2 and 3 had some smaller clusters (Figure 3-6, panel C). During the third year, Area 1 continued to have fewer NS visits, but the clusters were less clear and more scattered than earlier. Areas 2, 3, and 4 had small clusters of more NS visits (Figure 3-6, panel D).





3.2.4 Development

Among all participants, the mean (SD) total score on the Griffiths Mental Development Scales was 109.5 (17.6). The mean (SD) score on the performance, language, and locomotor scales was 15.3, (6.2), 14.9 (5.0), and 27.7 (3.8), respectively.

Area 4 had the highest mean total development, performance, language, and locomotor scores. The second-highest total score was in Area 2, which had the same high performance score as Area 4. Area 2 had the lowest language score and the second-highest locomotor score. The total score was second-lowest in Area 3, with the second-lowest score on performance, the second-highest score on language, and the lowest score on locomotor. The lowest total score was in Area 1, where the performance score was second-highest but the language and locomotor scores were the second-lowest. Except for locomotor scores, the differences among areas were not statistically significant (Table 3-6).

	Area 1	Area 2	Area 3	Area 4	P value*
Total score, mean (SD)	107.4 (18.9)	109.1 (17.9)	108.1 (20.2)	112.3 (14.3)	0.075
Performance score, mean (SD)	15.2 (6.6)	15.4 (5.9)	14.6 (6.4)	15.4 (5.9)	0.637
Language score, mean (SD)	14.8 (5.3)	14.6 (5.1)	14.9 (4.8)	15.5 (4.8)	0.209
Locomotor score, mean (SD)	24.8 (3.7)ª	24.8 (3.5)ª	23.6 (4.7)	25.1 (3.7) ^a	0.008

Table 3-6. Total Development, Performance, Language, and Locomotor Scores at 60 Months of Child Age

*P values derived by ANOVA.

Notes: Groups that do not share a common superscript differ significantly from each other (p < 0.05); Green indicates "best" values; yellow second "best"; orange second "worst"; red "worst."

The total score showed clustering of high values in all areas and clustering of low values in Areas 1, 2, and 3 (Figure 3-7, panel A). On the performance scale, there was some clustering of low values and a clearer clustering of high values in Area 2. Area 3 showed small clustering of low values, and Area 4 a small clustering of high values (Figure 3-7, panel B). On the language scale, there were high and low values in all areas except Area 4, where there were only two clusters of high values (Figure 3-7, panel C). On the locomotor scale, there were clusters of low values in Areas 1, 3, and 4 and clusters of high values in Areas 1, 2, and 4 (Figure 3-7, panel D).





Note: Based on the Griffiths Mental Development Scales

3.3 Objective II: Interactions between Area, Cluster, and Outcome

3.3.1 Preterm Birth

In the control group, the lowest incidence of preterm birth was in Area 1 and the highest incidence in Area 3. In the AZI-SP group⁶, the highest incidence of preterm birth was also in Area 3 whereas the lowest was in Area 4. The difference between the two groups was largest in Area 3 and lowest in Area 2. The differences between the groups were not statistically significant (Table 3-7).

Table 3-7. Proportion of Preterm Births (<37 Weeks) by Intervention Group for All Children and Stratified by Area and Household-Level Cluster

				Comparison between and control group	AZI-SP
Preterm birth	Interaction test <i>P</i> value	Control, n/N (%)	AZI-SP, n/N (%)	Risk ratio (95% CI)	P value*
All participants	NA	67/409 (16.4)	47/417 (11.3)	0.67 (0.49 to 0.97)	0.035
Area 1		12/84 (14.3)	9/92 (9.8)	0.68 (0.30 to 1.54)	0.361
Area 2	0.624	24/152 (15.6)	21/162 (13.0)	0. 82 (0.48 to 1.41)	0.476
Area 3	0.024	13/50 (26.0)	9/67 (13.4)	0.52 (0.24 to 1.11)	0.091
Area 4		18/123 (14.6)	8/96 (8.3)	0.57 (0.26 to 1.25)	0.162
Not part of any cluster		61/392 (15.6)	45/389 (12.6)	0.74 (0.52 to 1.06)	0.105
Household-level cluster of preterm births	0.039	6/14 (42.9)	2/23 (8.7)	0.20 (0.05 to 0.87)	0.032
Household-level cluster of full-term births		0/3 (0.0)	0/5 (0.0)	NA	NA

**P* values derived by log binomial regression.

Note: n/N: n=number of preterm births; N=number of participants living in area or part of cluster.

The incidence of preterm birth for participants not part of any household-level cluster was close to that of all participants, and the difference was not statistically significant. The AZI-SP group had a statistically significantly lower incidence of preterm birth compared to the control group in the household-level cluster of preterm births. There were no differences between the groups in household-level cluster of full-term births, with both groups having no preterm births and very few participants in this cluster (Table 3-7).

⁶ As stated in section 2.7, for analyses in Objective III, we included only the control and AZI-SP groups because previous analyses showed differences between these two groups.

3.3.2 Low Birth Weight

In the control group, the lowest incidence of low birth weight was in Area 1 and the highest incidence in Area 3. In the AZI-SP group, the lowest incidence was in Area 4 and the highest in Area 3. The difference between intervention groups was smallest in Areas 1 and 3; the difference was largest-and statistically significant- in Area 4 (Table 3-8).

The incidence of low birth weight for participants not part of any household-level cluster was close to that of all participants, with a statistically significantly lower incidence in the AZI-SP group than in the control group. The incidence of low birth weight was almost half as small in the AZI-SP group compared to the control group in household-level cluster of babies with low birth weight, although not statistically significant. There were a few participants in the household-level cluster of babies with not low birth weight and only one baby with low birth weight in the control group compared to none in the AZI-SP group (Table 3-8).

	Interaction			Comparison between AZI-SP and control group		
Low birth weight	test <i>P</i> value	(%)	AZI-SP, n/N (%)	Risk ratio (95% CI)	P value*	
All participants	NA	44/377 (11.7)	25/385 (6.5)	0.57 (0.35 to 0.89)	0.014	
Area 1		7/77 (9.1)	6/85 (7.1)	0.78 (0.27 to 0.21)	0.635	
Area 2	0.429	17/142 (12.0)	11/145 (7.6)	0.63 (0.31 to 1.30)	0.216	
Area 3	0.428	6/42 (14.3)	7/63 (11.1)	0.78 (0.28 to 2.15)	0.629	
Area 4		14/116 (12.1)	1/92 (1.1)	0.09 (0.01 to 0.67)	0.019	
Not part of any cluster		31/312 (9.9)	18/318 (5.7)	0.57 (0.33 to 1.00)	0.049	
Household-level cluster of babies with low birth weight	ousehold-level uster of babies with w birth weight 0.453		7/57 (12.3)	0.51 (0.22 to 1.20)	0.123	
Household-level cluster of babies with not low birth weight		1/15 (6.7)	0/10 (0.0)	NA	NA	

Table 3-8. Incidence of Low Birth Weight (<2,500g) by Intervention Group for All Children and</th> Stratified by Area and Household-Level Cluster

**P* values derived from log binomial regression.

Note: n/N: n=number of babies with low birth weight; N=number of participants living in area or part of cluster.

3.3.3 Length-for-Age Z-Score

LAZ at 24 months was lowest in Area 3 and highest in Area 2 in both the control and AZI-SP groups. Differences between the groups were smallest in Area 3, while Area 4 had the largest—and statistically significant—differences (Table 3-9).

Results of those not part of any cluster were similar to those of all participants, and the difference between intervention groups was not statistically significant. The difference between the AZI-SP group and the control group was larger in the household-level cluster of higher LAZ than in the household-level cluster of lower LAZ, but the difference was not statistically significant.

				Comparison between AZI- and control group	
LAZ at 24 months	Interaction test <i>P</i> value	Control, mean (SD)	AZI-SP, mean (SD)	Difference in means (95% Cl)	P value*
All participants	NA	-2.17 (1.01)	-1.99 (1.04)	0.18 (0.02 to 0.33)	0.031
Area 1		-2.15 (1.07)	-1.98 (1.02)	0.17 (-0.18 to 0.52)	0.339
Area 2	0.607	-1.98 (0.97)	-1.86 (1.05)	0.11 (-0.15 to 0.37)	0.395
Area 3		-2.34 (1.03)	-2.32 (1.13)	0.02 (-0.44 to 0.49)	0.919
Area 4		-2.33 (0.98)	-1.99 (0.98)	0.33 (0.05 to 0.62)	0.022
Not part of any cluster		-2.16 (1.01)	-2.01 (0.99)	0.15 (-0.02 to 0.32)	0.081
Household-level cluster of lower LAZ	0.575	-2.57 (1.10)	-2.50 (1.43)	0.07 (-0.71 to 0.85)	0.859
Household-level cluster of higher LAZ		-2.00 (0.89)	-1.60 (0.89)	0.40 (-0.00 to 0.80)	0.055

Table 3-9.	LAZ at 24 Months by Intervention Group for All Children and Stratified by Area and
	Household-Level Cluster

* *P* values derived by linear regression.

3.3.4 Postneonatal Mortality

In the control group, the proportion of postneonatal mortality was lowest in Area 2 and highest in Area 3. In the AZI-SP group, postneonatal mortality was lowest in Area 1 and highest in Area 2. Postneonatal mortality in Area 2 was lower in the control group than in the AZI-SP group, but the difference was not statistically significant. In Areas 1 and 3, postneonatal mortality was statistically significantly lower in the AZI-SP group than in the control group (Table 3-10).

The proportion of postneonatal mortality in participants not part of any household-level cluster was lower in the AZI-SP group than in the control group, with a statistically significant difference. Postneonatal mortality was five times higher in the control group than in the AZI-SP groups in clusters of higher mortality, but the difference was not statistically significant. There were no differences between groups in clusters of lower mortality. (Table 3-10).

Destroomatel	Interaction	Control n/N		Comparison between AZI-SP and control group		
mortality	test <i>P</i> value	(%)	AZI-SP, n/N (%)	Risk ratio (95% CI)	P value	
All participants	NA	21/394 (5.3)	7/402 (1.7)	0.33 (0.14 to 0.76)	0.009	
Area 1		8/81 (9.9)	0/88 (0.0)	0.28 (0.00 to 0.71)*	0.002*	
Area 2	0.003	3/145 (2.1)	4/156 (2.6)	1.24 (0.28 to 5.44)	0.776	
Area 3		6/48 (12.5)	1/65 (1.5)	0.12 (0.02 to 0.99)	0.049	
Area 4		4/120 (3.3)	2/93 (2.2)	0.65 (0.12 to 3.45)	0.608	
Not part of any cluster		16/361 (4.4)	6/370 (1.6)	0.37 (0.14 to 0.92)	0.034	
Household-level cluster of higher mortality	0.133	5/32 (15.6)	1/31 (3.2)	0.21 (0.03 to 1.67)	0.139	
Household-level cluster of lower mortality		0/1 (0.0)	0/1 (0.0)	NA	NA	

Table 3-10. Proportion of Postneonatal Mortality by Intervention Group for All Children and Stratified by Area and Household-Level Cluster

* *P* values derived from log binomial regression.

** Estimates obtained from exact logistic regression.

Note: n/N: n=number of postneonatal deaths; N=number of participants living in area or part of cluster.

3.3.5 Development

The total development score at 5 years of age was lowest in Area 2 and highest in Area 4 in the control group. In the AZI-SP group, the score was lowest in Area 1 and highest in Area 4. The difference between groups was smallest in Area 1, while Area 2 had the largest—and statistically significant—difference (Table 3-11).

The development score for those not part of any cluster was similar to those of all participants. The total development score in the AZI-SP group was 8.4 points higher than in the control group in the household-level cluster of lower scores. The respective difference in the cluster of higher scores was 4 points. Differences between groups were not statistically significant in any of the clusters (Table 3-11).

				Comparison between and control group	AZI-SP
Total development score at 60 months	Interaction test <i>P</i> value	Control, mean (SD)	AZI-SP, mean (SD)	Difference in means (95% Cl)	P value*
All participants	NA	108.1 (16.3)	111.1 (19.2)	3.0 (-0.4 to 6.5)	0.083
Area 1		108.1 (14.3)	108.3 (22.4)	0.3 (-7.8 to 8.4)	0.947
Area 2	0.668	105.9 (18.1)	111.5 (17.0)	5.7 (0.2 to 11.1)	0.041
Area 3		108.0 (17.0)	109.2 (22.8)	1.2 (-10.2 to 12.6)	0.835
Area 4		111.2 (14.5)	114.8 (16.0)	3.6 (-2.2 to 9.4)	0.224
Not part of any cluster		107.5 (15.2)	109.9 (18.7)	2.4 (-1.2 to 6.1)	0.193
Household-level cluster of lower scores	0.648	98.8 (19.0)	107.3 (27.6)	8.4 (-9.9 to 26.7)	0.352
Household-level cluster of higher scores		117.3 (18.7)	121.3 (12.1)	4.0 (-4.9 to 12.9)	0.369

Table 3-11. Total Development Score at 60 Months by Intervention Group for All Children andStratified by Area and Household-Level Cluster

**P* values derived by linear regression.

3.4 Objective III: Adjustment of Outcomes for Area and Household-Level Cluster

Using area of residence and household-level clusters to adjust risk ratios for preterm birth, low birth weight, and postneonatal deaths in the control and AZI-SP groups reduced those ratios by 0.02 in all cases except for postneonatal deaths adjusted for household-level clusters; for that outcome, the adjustment did not change the risk ratio. In addition, the adjustment slightly reduced confidence intervals for all outcomes, except for postneonatal deaths adjusted for household-level clusters (Table 3-12).

The differences in means and confidence interval for LAZ at 24 months between the control group and the AZI-SP group remained the same when adjusted for residential area or household-level cluster. The difference in means for total development score increased 0.37 points after adjustment for residential area but remained nearly the same after adjustment for household-level cluster. The confidence intervals were slightly reduced by each adjustment (Table 3-13).

Table 3-12. Proportion of Preterm Birth, Low Birth Weight, and Postneonatal Deaths by Intervention Group without and with Adjustment for Area and Household-Level Cluster

		Comparison between AZI-SP and control			l group	Absolute difference	
	Control, n/N (%), unadjusted	AZI-SP, n/N (%), unadjusted	Unadjusted risk ratio (95% CI)	<i>P</i> value [*]	Adjusted risk ratio (95% CI)	P value [*]	in confidence interval
Preterm birth, adjusted for area					0.67 (0.47 to 0.95)	0.023	-0.01
Preterm birth, adjusted for household-level cluster	67/409 (16.4)	47/417 (11.3)	0.69 (0.49 to 0.97)	0.035	0.67 (0.48 to 0.95)	0.026	-0.01
Low birth weight, adjusted for area					0.54 (0.34 to 0.86)	0.010	-0.02
Low birth weight, adjusted for household-level cluster	44/377 (11.7)	25/385 (6.5)	0.56 (0.35 to 0.89)	0.014	0.54 (0.34 to 0.86)	0.010	-0.02
Postneonatal deaths, adjusted for area					0.31 (0.13 to 0.71)	0.006	-0.04
Postneonatal deaths, adjusted for household-level cluster	21/394 (5.3)	7/402 (1.7)	0.33 (0.14 to 0.76)	0.009	0.33 (0.14 to 0.76)	0.009	0.00

**P* values derived by log binomial regression.

Note: For each analysis, "n" refers to the number of participants in each group with the noted condition and "N" refers to the total number of participants in each group.

Table 3-13. LAZ at 24 Months and Total Development Score by Intervention Group without and with Adjustment for Area and Household-Level Cluster

			Comparison between AZI-SP and control group				Absolute
	Control, mean (SD), unadjusted	AZI-SP, mean (SD), unadjusted	Unadjusted difference in means (95% CI)	P value *	Adjusted difference in means (95% CI)	<i>P</i> value*	in confidence interval
LAZ at 24 months, adjusted for area					0. 18 (0.02 to 0.33)	0.029	0.00
LAZ at 24 months, adjusted for household-level cluster	-2.17 (1.01)	-1.99 (1.04)	0.18 (0.02 to 0.33)	0.030	0.18 (0.02 to 0.33)	0.028	0.00
Total development score at 60 months, adjusted for area					3.40 (-0.05 to 6.85)	0.053	-0.04
Total development score at 60 months, adjusted for household-level cluster	108.12 (16.26)	111.15 (19.18)	3.03 (-0.40 to 6.46)	0.083	3.04 (-0.32 to 6.39)	0.076	-0.18

**P* values derived by linear regression.

4 **Discussion**

Interventions in various locations around the world have had mixed successes addressing the pressing problem of child undernutrition. The heterogeneity in the results from different interventions could be due to differences in study design or data collection or due to the influence of geographical location on the intervention effect. This study was done to investigate whether there is geographical clustering of "good" or "bad" outcomes within one study sample, thereby eliminating the possible effect of study design or data collection. We did not find other literature reporting on similar research. There is a lot of literature on the spatial analysis of certain phenomena in medical research, but we are not aware of any conducted in Malawi covering a broad range of topics as done in this study.

Our study has shown that the outcomes from the LAIS trial did cluster geographically, both by residential area and household cluster, but not consistently. Similarly, there was some evidence that residential area or the household-level cluster modified the intervention effect, but these results were not consistent across the area or topics. In addition, there was some evidence that including information on residential location into regression models did affect the precision of the effect size estimate for some outcomes, although the effect was rather small.

The areal clustering of the outcomes was not consistent. There was heterogeneity in the clustering depending on the outcome topic, meaning that there was not one clear cluster or area that universally ranked the best or worst on all outcomes. In some cases, there were clusters of both "good" and "bad" outcomes in the same residential area (e.g., weight-for-gestational-age z-score in Area 4), but in some other cases there was a consistent structure to the areal scores (e.g., child length where the ranking of the areas was statistically significant and persisted at all child ages).

Several factors in this study could have affected the results. We used the same areal segmentation for all outcomes. This four-way division was partly designed for potential use in future research in the area, but it is possible that a different number of areas might lead to a different result. It is feasible that processes or background variables that cause clustering of outcomes in some areas are not the same for all outcomes. For example, distance to the health center might have a more direct influence on the number of NS visits to a health center than on certain developmental scores at 5 years of age. In addition, to look for consistent areas of outcome clustering in the analysis of the household level, we used the same setting for the spatial relationship on the Hot Spot tool for all outcomes. It could be that the spatial processes causing clustering in some outcomes operate on different spatial levels within the area and therefore do not show up as spatial clusters. Also, the data we used were not collected with the present purpose in mind. This was a post-hoc analysis. Data preferably would have been collected over a larger area and with more equal representation over the whole trial area. For example, some villages were not covered by our data or included very few participants.

Our research has shown that even after eliminating the effect of study design and data collection, there is heterogeneity in child health outcomes within one study done in rural Malawi. Geographical location possibly has its own influence on the results from intervention trials, and it is possible to understand this influence using analyses that look at participants' residential location, either at the household level or by aggregation into residential areas. Further study could indicate whether refinement of analysis settings or a different segmentation of the area would yield essentially different results from the ones reported here.

5 References

Cheung, Y.B. 2014. *Statistical Analysis of Human Growth and Development*. Boca Raton, FL: CRC Press.

de Onis, M. et al. 2007. "Development of a WHO Growth Reference for School-Aged Children and Adolescents." *Bulletin of the World Health Organization*. 85:660–667.

Hallamaa, L. et al. 2018. "Child Health Outcomes After Presumptive Infection Treatment in Pregnant Women: A Randomized Trial." Pediatrics. 141(3):e20172459

Luiz, D. et al. 2006. *Griffiths Mental Development Scales—Extended Revised, 2–8 Years*. Oxford: Hogrefe.

Luntamo, M. et al. 2010. "Effect of Repeated Treatment of Pregnant Women with Sulfadoxine-Pyrimethamine and Azithromycin on Preterm Delivery in Malawi: A Randomized Controlled Trial." *American Journal of Tropical Medicine and Hygiene*. 83: 1212–1220.

Luntamo, M. et al. 2012. "The Effect of Monthly Sulfadoxine-Pyrimethamine, Alone or with Azithromycin, on PCR-Diagnosed Malaria at Delivery: A Randomized Controlled Trial." *PLoS ONE*. 7: e41123.

Luntamo, M. et al. 2013. "The Effect of Antenatal Monthly Sulphadoxine–Pyrimethamine, Alone or with Azithromycin, on Foetal and Neonatal Growth Faltering in Malawi: A Randomised Controlled Trial." *Tropical Medicine and International Health.* Vol. 18, No. 4: 386–397.

National Statistical Office (NSO) [Malawi] and ICF. 2017. *Malawi Demographic and Health Survey* 2015–16. Zomba, Malawi, and Rockville, MD, USA: NSO and ICF.

Villar, J. et al. 2014. "International Standards for Newborn Weight, Length, and Head Circumference by Gestational Age and Sex: The Newborn Cross-Sectional Study of the INTERGROWTH-21st Project." *Lancet*. 384(9946):857–68.

World Health Organization (WHO) Multicentre Growth Reference Study Group. 2006. "WHO Child Growth Standards Based on Length/Height, Weight, and Age." *Acta Paediatrica*. 450:76–85.

WHO. 2017a. "The Healthy Growth Project." Available at: http://www.who.int/nutrition/healthygrowthproj/en/.

WHO. 2017b. "Malnutrition Fact Sheet—May 2017." Available at: <u>http://www.who.int/mediacentre/factsheets/malnutrition/en/</u>.

Appendix 1. Included and Excluded Participants

Table A1. Baseline Characteristics of LAIS Mothers Who Were Either Included or Excluded from the
Residential Location Analyses

Characteristic	Included	Excluded	P value ¹
Participants	1,241	79	
Mean (SD) age (years)	24.9 (6.5)	23.9 (5.4)	0.170
Mean (SD) height (cm)	155 (5.5)	155.3 (5.6)	0.609
Mean (SD) BMI (kg/m ²)	21.8 (2.2)	21.8 (1.8)	0.835
Mean (SD) gestational age at enrollment (weeks)	20.1 (3.1)	19.9 (2.9)	0.511
Proportion (%) who were primiparous	299/1,241 (24.1)	7/79 (8.9)	<0.001
Mothers with HIV (%)	148/1,241 (11.9)	13/79 (16.5)	0.218
Mean (SD) blood Hb concentration	110.3 (18.7)	108.9 (18.8)	0.498
Proportion with microscopic peripheral blood malaria parasitemia (%)	108/1,240 ² (8.7)	9/79 (11.4)	0.413
Proportion of literate participants (%)	357/1,241 (28.8)	27/79 (34.2)	0.308
Mean (SD) years of schooling completed	2.2 (2.7)	2.5 (2.6)	0.406
Proportion of those owning a bed net (%)	904/1,241 (72.8)	64/79 (81)	0.117
Proportion who used bed net during previous night (%)	745/903 (82.5)	52/64 (81.3)	0.737

¹*P* values derived by t-test (comparison of means) or Fisher's exact test (comparison of proportions).

²Value missing for one participant.

Note: BMI = body mass index; HIV = human immunodeficiency virus; Hb = hemoglobin.

Appendix 2. Characteristics of Included Women by Intervention Group

Characteristic	Control (SP twice)	Monthly SP	AZI-SP
No. enrolled women	409	415	417
Mean (SD) age (years)	24.9 (6.9)	25 (6.6)	24.8 (5.9)
Mean (SD) height (cm)	154.9 (5.5)	154.7 (5.4)	155.4 (5.5)
Mean (SD) BMI (kg/m ²)	21.7 (2.2)	21.7 (2.1)	21.9 (2.2)
Mean (SD) gestational age at enrollment (weeks)	20.2 (3)	20.1 (3.2)	20.0 (3.0)
Proportion (%) who were primiparous	107/409 (26.2)	104/415 (25.1)	88/417 (21.1)
Mothers with HIV (%)	44/409 (10.8)	58/415 (14)	46/417 (11)
Mean (SD) blood Hb concentration	110.3 (19)	111 (17.3)	109.7 (19.6)
Proportion with microscopic peripheral blood malaria parasitemia (%)	44/408 ¹ (10.8)	39/415 (9.4)	25/417 (6.0)
Proportion of literate participants (%)	108/409 (26.4)	120/415 (28.9)	129/417 (30.9)
Mean (SD) years of schooling completed	2.1 (2.7)	2.1 (2.6)	2.4 (2.8)
Proportion of those owning a bed net (%)	300/409 (73.4)	296/415 (71.3)	308/417 (73.9)
Proportion who used bed net during previous night (%)	251/300 (83.7)	246/295 (83.4)	248/308 (80.5)

Table B1. Baseline Characteristics of Included Women by Intervention Group, at Enrollment

¹Value missing for one participant.

Note: SP = sulfadoxine-pyrimethamine; AZI = azithromycin; BMI = body mass index; HIV = human immunodeficiency virus; Hb = hemoglobin.

Appendix 3. Baseline Characteristics of Included Women by Relocation Status during Trial

Table C1. Baseline Characteristics of Included Mothers, Tabulated by Whether They Had Moved during Trial

Characteristic	Mothers who did not move during trial period	Mothers who moved during trial period	P value ¹
No. enrolled women	1,166	69	
Mean (SD) age (years)	25 (6.5)	24.1 (6.3)	0.255
Mean (SD) height (cm)	155 (5.5)	155 (5.3)	0.999
Mean (SD) BMI (kg/m ²)	21.8 (2.2)	21.3 (2.3)	0.039
Mean (SD) gestational age at enrollment (weeks)	20.1 (3.1)	20.3 (2.8)	0.622
Proportion (%) who were primiparous	278/1,166 (23.8)	20/69 (29.0)	0.315
Mothers with HIV (%)	139/1,166 (11.9)	8/69 (11.6)	1.000
Mean (SD) blood Hb concentration	110.2 (18.7)	110.8 (17.3)	0.794
Proportion with microscopic peripheral blood malaria parasitemia (%)	102/1,165² (8.8)	6/69 (8.7)	1.000
Proportion of literate participants (%)	337/1,166 (28.9)	17/69 (24.6)	0.496
Mean (SD) years of schooling completed	2.2 (2.7)	2 (2.7)	0.583
Proportion of those owning a bed net (%)	849/1,166 (72.8)	50/69 (72.5)	1.000
Proportion who used bed net during previous night (%)	696/848 (82.1)	44/50 (88.0)	0.343

¹*P* values derived by t-test (comparison of means) or Fisher's exact test (comparison of proportions).

²Value missing for one participant.