

Zambia - Water, Sanitation and Drainage Infrastructure

Report generated on: May 30, 2018

Visit our data catalog at: <https://data.mcc.gov/evaluations/index.php>

Overview

Identification

COUNTRY

Zambia

EVALUATION TITLE

Water, Sanitation and Drainage Infrastructure

EVALUATION TYPE

Independent Impact Evaluation

ID NUMBER

DDI-MCC-ZAM-LWSSD-CDC-2018-v01

Version

VERSION DESCRIPTION

- v01: Edited, anonymous dataset for public distribution.

Overview

ABSTRACT

The Zambia Compact is comprised of interventions to the water, sanitation, and drainage infrastructure in the city of Lusaka. The goals of these interventions include reducing poverty through economic development, decreasing the incidence of waterborne and water-related diseases, and generating time savings for businesses and households. To evaluate the impact of the Compact interventions the following evaluation questions were developed by MCC, MCA Zambia, and CDC.

Health

- What are the health benefits attributable to each type of Compact activity?

Safe Water Supply/Consumption

- What are the current consumption rates of safe versus un-safe water consumption and usage?
- Do Compact activities lead to an increase in safe water consumption?

Economic and Social

- Do households experience an increase in income due to Compact activities?
- Are households able to afford household connections, toilets, and water bills?
- Were subsidy provisions adequate for sanitation connections?
- What is the probability of finding work for beneficiaries?
- What are the time and cost savings/use attributable to each Compact activity?

Flooding

- Is there a decrease in the frequency, intensity and duration of flooding?
- Is there a decrease in property damage and loss of business caused by flooding?
- Is there a decrease in travel time due to reduced flooding?

Given the cause-and-effect nature of the Compact as described in the logic model, CDC recommended an impact evaluation as the most appropriate method to evaluate the effect of the interventions. The impact of interventions to extend water supply and sanitation networks to residential areas will be evaluated by comparing pre and post outcome measures in both intervention and control (counterfactual) areas. Control areas are peri-urban areas that are similar to intervention areas with respect to water and sanitation characteristics, but that are not receiving any interventions.

A suitable control group for the households and businesses that will be affected by the drainage intervention in central Lusaka does not exist. Characteristics that make the Bombay Drain system unique include its mix of commercial and residential development and their propensity to flood due to their co-location along a primary drainage system (the Bombay Drain). The Bombay Drain flows through the major business and commercial districts of Lusaka and this commercial density is not present to the same degree in other drainage systems. The Bombay Drain is also distinguished by its length and the unique nature of its flooding. As a result of these characteristics, we were unable to identify a drainage system in Lusaka with characteristics similar enough to the Bombay Drain to serve as an adequate control area. Therefore, the impact of the drainage intervention on households and businesses will be estimated using pre-post comparisons with no control group.

The primary quantitative study design will be a prospective, cross-sectional intervention-control and pre-post impact evaluation of new water supply and sanitation interventions and drainage improvement interventions. We will evaluate the changes in outcome measures between baseline and post-intervention time periods within the intervention areas and compare these measures to any differences observed in the control areas during the same time period.

EVALUATION METHODOLOGY

pre-post, randomization, difference-in-difference

UNITS OF ANALYSIS

individuals, households, community

KIND OF DATA

Sample survey data [ssd]

TOPICS

Topic	Vocabulary	URI
Water, Sanitation and Hygiene	MCC Sector	

KEYWORDS

Lusaka, Zambia, Compact, Peri-urban, Infrastructure, Monitoring, Laboratory, Cross-sectional, WASH, Water, Sanitation, Hygiene, Health impact, Economic impact, Expenditures, Drainage, Traffic evaluation, Business evaluation, Diarrhea, Respiratory illness, Flooding, Impact evaluation, Random sampling, Time savings, Water quality, Stored water, Source water, Point-of-consumption water, Household water, E. coli, Escherichia coli, Total coliforms, Nitrates, Free residual chlorine, Turbidity, Field sampling

Coverage

GEOGRAPHIC COVERAGE

Urban and peri-urban areas throughout the capital city Lusaka, Zambia.

UNIVERSE

Individuals and households in urban and peri-urban areas of Lusaka, Zambia.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Centers for Disease Control and Prevention	

FUNDING

Name	Abbreviation	Role
------	--------------	------

Name	Abbreviation	Role
Millennium Challenge Corporation	MCC	

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Millennium Challenge Corporation	MCC		Review of Metadata
Centers for Disease Control and Prevention	CDC		Independent Evaluator

DATE OF METADATA PRODUCTION

2018-05-30

DDI DOCUMENT VERSION

Version 1.1 (May 2018)

DDI DOCUMENT ID

DDI-MCC-ZAM-LWSSD-CDC-2018-v01

MCC Compact and Program

COMPACT OR THRESHOLD

Zambia Compact

PROGRAM

The CDC baseline evaluation of the Lusaka Water Supply Sanitation and Drainage project (LWSSD) examined key health and economic indicators in peri-urban and urban areas around Lusaka, Zambia that received water supply interventions (municipal water network extensions and new and rehabilitated water kiosks): Chipata/SOS East, Ng'ombe, and Kamanga (CP3 and CP5). The evaluation also examined the combined sanitation and water supply interventions in Mtendere. Both areas of Mtendere (West and East) received an extension of the sewerage network and water supply (CP3 and CP4). The evaluation also examined the urban and peri-urban areas within the 250 meter catchment of the Bombay Drain improvements (CP 7 and 8) to study flooding impacts on the surrounding community and businesses as well as health and economic indicators.

MCC SECTOR

Capacity Building and Institutional Development (Cap Bldg & Inst Dev)

PROGRAM LOGIC

MCC in partnership with the Government of Zambia (Millennium Challenge Account Zambia [MCA Zambia]) is implementing a large-scale, \$350 million upgrade and extension of the water, sanitation, and drainage infrastructure in Lusaka to increase population access to potable water, sanitation, and flood protection. The Lusaka Water Supply, Sanitation and Drainage (LWSSD) project (the Compact) will strengthen and upgrade the main surface water treatment plant for Lusaka, extend water supply and sanitation networks into areas with limited household water connections and toilets, rehabilitate existing water kiosks, and improve the drainage network for the primary business district and surrounding residential communities along the Bombay drainage catchment area in Lusaka. The location, type, and scope of the Compact interventions was determined by MCC, the Government of Zambia (GRZ), MCA Zambia, and key stakeholders such as the Lusaka Water and Sewerage Company and the Lusaka City Council. Compact activities included: Core Water Network Rehabilitation (Water Supply): Rehabilitation of the Iolanda treatment plant to restore production to 110,000 m³/day from 98,000 m³/day, rehabilitation of transmission mains and distribution centers, and strengthening of the primary distribution system. This component is expected to benefit approximately 860,000 people by providing more reliable water service and increased water supply and coverage. MCA Zambia refers to these sub-projects as LP-1, LP-6, LS-1. These sub-projects are part of Contract Packages (CP) (e.g., construction projects) 1 and 2. Chelston Distribution Line Rehabilitation and Expansion (Water Supply): Extension and rehabilitation of secondary and tertiary networks into the Central and Chelston Branch district metering areas, including the extension of distribution pipes into residential areas to facilitate new household connections, construction of new water kiosks, and rehabilitation of existing kiosks. Over 568,000 beneficiaries are expected from these activities: 416,000 from rehabilitation and 152,000 from network expansion. Beneficiaries are located in Chipata/SOS East, Ng'ombe, Kamanga, Mtendere, Kwamwena, and Ndeke/Vorna Valley. MCA Zambia refers to these sub-projects as LS-2 and LS-3. These sub-projects are part of CPs 3 and 5. Chelston and Kaunda Square Sewersheds Rehabilitation and Expansion (Sanitation): Expand sewer network to facilitate new household and business connections in Mtendere, upgrade Chelston

sewage pumping station and Kaunda Square interceptor, and upgrade and expand the Kaunda Square stabilization ponds. 98,000 beneficiaries are projected from the expanded sewer network in Mtendere; nearly 57,000 beneficiaries are expected from rehabilitative sanitation interventions. MCA Zambia refers to these sub-projects as CSE-44, CSU-4, CSU-15, TU-5, TE-3. These sub-projects are part of CPs 3 and 4. Bombay and Mazyopa Drain Improvements (Drainage): Extension and rehabilitation of the Bombay drainage system in central Lusaka, and rehabilitation of the Mazyopa Drain in Northern Lusaka to accommodate the expected increased flow from the Bombay Drain. An estimated 188,000 people are expected to benefit from these drainage improvements. These sub-projects are part of CPs 7 and 8.

PROGRAM PARTICIPANTS

The study population for the baseline evaluation consisted of residents living in urban and peri-urban areas throughout Lusaka. Survey participants were any adult (aged 18 and over) male or female head of household. These heads of household also answered questions pertaining to all individuals in their household. The Zambian Central Statistics Office (CSO) defines a household as a group of people that normally cook, eat, and live together who may or may not be related by blood.

Sampling

Study Population

Individuals and households in urban and peri-urban areas of Lusaka, Zambia.

Sampling Procedure

Geocoding: A two stage cluster sampling strategy was employed for both the WaSH survey and the household drainage survey. The primary sampling units were the "Standard Enumeration Areas" (SEAs) provided by the Zambian Central Statistics Office (CSO). Households within each randomly selected SEA are the secondary sampling units (SSU). Google Earth aerial images of Lusaka provided by the CDC GIS Unit showed all building structures within each SEA. However, each structure contained multiple households (SSUs) that needed to be identified and enumerated through the process of "geocoding". At the beginning of each work week, each team of enumerators was assigned 1-2 SEAs for enumeration/identification of all households (geocoding). The team entered the assigned SEAs and used the GPS enabled app "Avenza PDF Maps" on their Android tablets to locate structures. The aerial maps with the boundaries of the SEAs outlined were pre-loaded on the enumerator tablets by CDC staff. Each team divided the SEA into segments and walked structure to structure locating the exact number of households in each SEA and dropping "pins" that allow the household geo-coordinates ("geocode") to be captured and saved. At the end of each day of geocoding, the teams sent the maps with the captured household coordinates as CSV files via bluetooth to CDC staff who then created SEA maps complete with all enumerated households in ArcGIS. CDC then used ArcGIS to randomly select 75 households in each SEA (1-42 for surveying and the remaining 43-75 as potential replacements in case of refusals, vacant households, ineligible households, etc.). These maps were then reloaded back onto the same enumeration team's tablets and used to locate the randomly selected households for interviews (this also aided in familiarity and ease of location of selected households). This process was used for the drainage and WaSH surveys.

WaSH Sampling Frame: For the WaSH survey, 50 SEAs from the water supply and sanitation intervention areas (Mtendere East and Mtendere West) were selected using probability proportional to size (PPS) sampling. In the water supply intervention areas (Chipata/SOS, Kamanga, and Ngombe), 75 SEAs were selected using the same method. And for the control areas, 150 SEAs were selected using PPS sampling. Thus, the total sampling frame for the survey consisted of 275 SEAs in peri-urban areas around Lusaka. Due to the lower number of SEAs available in Mtendere East and Mtendere West, only 50 SEAs were selected in the primary sampling unit stage instead of 75 SEAs. In order to have the same number of households per treatment area (~3,000) and a 1:1 intervention to control household interview ratio across all treatments areas, the number of households being randomly selected and interviewed in Mtendere East and West was increased from 42 to 62 households. In all other areas, 42 households were interviewed per SEA. Overall, 3,186 households were sampled in the water supply only intervention areas; 3,097 households were sampled in the water supply and sanitation intervention areas; and 6,229 households were sampled from control areas. Additionally, within each selected SEA, 11 stored water samples, 3 source water samples and 3 point of consumption water samples were randomly selected from the 42 or 62 households that had been selected for a household interview.

WaSH Household Replacement Strategy:

Households that were selected for an interview could be replaced for the following reasons:

- A household was selected for an interview and no one was home after three visits.
- A head of household or their spouse did not consent to be interviewed.
- A head of household or their spouse were under the age of 18.
- A household had not been living in that structure for 2 or more months.
- A head of household or their spouse could only be found on Sunday (enumerators did not work on Sundays)
- If the head of household or their spouse were only found outside of field hours (8:30am to 3:30pm).
- If the household was selected for water collection but the right type(s) of water was not available, that household was still interviewed, but another was selected to take its place as an interview with water collection.

If a household needed to be replaced for one of these reasons, supervisors would assign the enumerator the nearest replacement household from the replacement pool (households randomly selected as 43-75 or 63-85 for Mtendere East and West) and the enumerator would note the replacement in their field control sheet.

Household Drainage Sampling Frame: For the household drainage survey, the sampling frame was generated by overlaying census SEA boundaries and the boundary of the Bombay drainage catchment area using ArcView GIS. SEAs that were located within approximately 250 meters on either side of the Bombay Drain, where the greatest impact of flooding and flooding reduction was expected, were eligible for random selection. Once a list of SEAs that fell in this catchment area was generated, PPS sampling was used to select 75 SEAs for the sampling frame. Forty-two households located in each of these

selected SEAs were then randomly selected for inclusion in the evaluation. Overall, 3,142 households were sampled in the drainage intervention catchment area. This part of the evaluation is a pre/post comparison with no control group.

Household Drainage Replacement Strategy: The household drainage survey used the same replacement protocol as the WaSH survey without water replacements (as there was no water sampling in the household drainage survey).

Business Drainage Sampling Frame:

The market area was defined and divided into three general locales:

1. Kamwala Interior: Interior of Kamwala Market
2. Kamwala Exterior: Exterior of Kamwala Market i.e. outer perimeter: shops facing outward from market.
3. Along Bombay drain: Business along Bombay Road and connecting roads, south of Independence Ave, and north of Chongwe Road.

The enumerators went to all the existing businesses and created a sampling frame. In Kamwala interior and exterior regions they used the shop numbers already given by Kamwala committee and Chinese Management Company. Detailed sketch maps of all the shops were done by hand and shop numbers marked. Along the Bombay road shops, the enumerators gave their own numbers and wrote the business names in the map. During this census, they also sensitized the owner/manager about the upcoming survey. This created sampling frame was then transferred to Microsoft Excel (Microsoft Inc., Redwood, WA). A simple random sample of around 24% (9% to 45%) of enlisted businesses was drawn by CDC using MS Excel. A sufficient pool of replacement businesses was also sampled in case of refusal and other circumstances. Since Kamwala Interior and Exterior were in close proximity and the number of businesses along Bombay drain were scattered over a larger area, the sampling strategy was shifted to draw more business from the larger area. In addition, the proportion varied due to the fact that the numbers proposed in the EDR were based on incomplete sampling frame and we wanted each block, shopping complex and market represented. Overall, 280 businesses inside the Kamwala Market were surveyed as well 60 businesses outside of the Kamwala Market. 247 businesses along the Bombay Drain (transient vendors) were also surveyed for a total of 587 business surveyed all together.

Business Drainage Replacement Strategy:

A sufficient pool of replacement businesses (25% of sample size) were also sampled in case of refusal and other circumstances. All businesses that were replaced either refused or exhausted three visits on three different days at different times.

Overall, a total of 15,656 households and 587 business were sampled in the baseline impact evaluation of the Zambia Compact.

Deviations from Sample Design

N/A

Response Rate

For the WaSH baseline survey, enumerators approached 16,059 households in total and completed interviews at 12,514 of those households for a response rate of 77.9%. For the household drainage baseline survey, enumerators approached 4,078 households in total and completed interviews at 3,142 of those households for a response rate of 77.0%.

Business Drainage Response Rate:

Kamwala Interior: $280/290 = 96\%$

Kamwala Exterior: $60/100 = 60\%$

Along Bombay drain= $247/275 = 89\%$

Weighting

Households were selected using probability proportionate to size (PPS) sampling. A two-stage cluster sampling strategy was used in which Standard Enumeration Areas (SEAs) served as the primary sampling units (PSUs) and households within selected SEAs served as secondary sampling units (SSUs). SEAs are the smallest enumeration units in Lusaka representing approximately 175 households on average. All SEAs that were located within intervention and control areas were listed in a sampling frame with cumulative population sizes (number of households). For each intervention area, 75 or 50 SEAs were selected using a PPS sampling strategy with emphasis on selecting at least half the number of SEAs in a given intervention area in order to reduce design effects. 150 control SEAs were selected using PPS. Households located in these selected SEAs

were then randomly selected using simple random sampling. Since PSUs were selected using PPS sampling in the first stage and the same number of households were then selected per PSU regardless of its size in the second stage, each household had the same probability of being selected in the population. PPS is “self-weighting,” eliminating the need for weights.

Questionnaires

Overview

The questionnaires for the Impact Evaluation of the Lusaka Water Supply, Sanitation and Drainage Project were designed to evaluate MCC key indicators and answer research questions in the Evaluation Design Report. The WASH household questionnaire included 12 modules: household demographics; sickness (prevalence of diarrhea and influenza-like illness) and associated costs; caretaker time loss; water collection; water storage and treatment; sanitation; hygiene; household garbage disposal; time use and expenditures; observations; and field control information for case management. The WASH household questionnaire was administered to 12,514 randomly selected households over a 12 month period.

The household drainage survey included the above 12 modules (with abbreviated WASH modules) and an additional module on flooding and the impact of flooding on household time use. The household drainage questionnaire was administered to 3,150 randomly selected households over a 4 month period. The business drainage survey primarily measured the impact of flooding on business activity and contained sections on business characteristics; flooding; illness; and revenue loss and property damage due to flooding. The business survey was administered to 587 systematically selected businesses in Lusaka during a 3 week period.

All three questionnaires were written in English and translated into Nyanja and Bemba, the two most prominently spoken tribal languages in Lusaka. For the drainage surveys (household and business), the English questionnaire was translated into Nyanja and Bemba by a local translation firm, RangeView Enterprises. The translated questionnaires were then back-translated into English by the same firm. For the WaSH household survey, the RuralNet Field Manager led the survey translation task. The draft English survey was translated into Nyanja and Bemba prior to the pre-test. As edits were made to the English version, the translated surveys were also updated. Feedback on the translated surveys was implemented based on comments from the pre-test, during training and the pilot.

Data Collection

Data Collection Dates

Start	End	Cycle
2016-02-01	2016-05-06	Drainage Baseline Evaluation
2016-10-17	2017-10-26	WaSH Baseline Evaluation

Data Collection Notes

Household Drainage Pretest:

The initial draft of the questionnaire was pre-tested by Joan Brunkard during the second CDC visit to Zambia in 2014.

Household Drainage Pilot:

The pilot for the household drainage survey took place January 25th- 29th, 2016 in a randomly selected SEA in the Garden peri-urban area that was within the 250 meter drainage catchment area but not selected for the sampling frame. All 24 enumerators participated in the pilot.

Household Drainage Average Interview Time: 60 minutes

Household Drainage Schedule:

Field staff arrived at the MCAZ office at 8am Wednesday-Friday and were expected to be on the buses heading to the field by 8:30. They arrived at MCAZ Saturdays at 8:30 departing at 9am and Sundays at 11:45 a.m. departing at 12:00pm. Weekends were incorporated into the schedule in order to locate household respondents who were only available on weekends.

Household Drainage Questionnaires:

Interviews for the household drainage survey were recorded on paper surveys. At the end of each day, field supervisors would collect surveys from their team, check for completeness and then store all completed surveys in locked cabinets at the MCAZ office for later data entry.

Business Drainage Pretest:

The initial draft of the questionnaire was pretested by Martin Meltzer during the second CDC visit to Zambia in 2014. The length of the questionnaire was reduced when business owners cited being too busy and felt the questions would mean loss of business.

Business Drainage Pilot:

The business drainage pilot occurred March 17th-18th, 2016. Businesses located north of Independence Road near Kamwala Market (not in the sampling frame) were used. Minor changes and adjustments were added to the questionnaires like adding income categories, correcting skip patterns and local language corrections.

Business Drainage Average Interview Time: 15 minutes

Business Drainage Schedule:

Data collection occurred from March 21st, 2016 to April 15th, 2016. Each working day (Monday to Friday) the enumerators would arrive at the market between 9:30 am to 10:00 am. The enumerators would then collect data until 3:30pm and arrive back in the MCAZ head office at approximately 4:00pm. From 4:00pm to 5:00pm, field level problems were resolved and the paper questionnaires were checked for accuracy. Then completed questionnaires were securely stored in file cabinets.

Business Drainage Questionnaires:

Interviews for the business drainage survey were recorded on paper surveys. At the end of each day, field supervisors would collect surveys from their team, check for completeness and then store all completed surveys in locked cabinets at the MCAZ office for later data entry.

Four staff members from CDC-Atlanta and one from CDC-Zambia provided overall fieldwork coordination and supervision for the drainage surveys.

WaSH Pre-test:

The pre-test for the WaSH survey began on Monday, August 29, 2016 and concluded on Friday, September 2, 2016 in the Linda community, approximately 20 kilometers north-west of the city of Lusaka, in a SEA that was not selected for the sampling frame. The pre-test was conducted by NORC and RuralNet to test the WaSH questionnaire for clarity and flow and

to obtain feedback on translations.

WaSH Pilot:

The WaSH survey pilot also took place in the Linda community on October 6, 2016. The field team visited and geo-coded 142 households in the selected SEA. A subset of the geo-coded households was then randomly selected for the pilot interviews. The pilot helped the field team practice the geo-coding process and become familiar with the questionnaires. After receiving feedback on the questionnaires from the field team, CDC and NORC updated the survey and extending the piloting period for another week from October 10th- 14th, 2016.

WaSH Average Interview Time: 60 minutes

WaSH schedule:

The WaSH survey enumerators worked from 7:30am to around 4:00pm from Monday to Thursday and Saturday between October 2016 and January 2017. Based on records indicating that response rates were not as high on Saturday, the enumerators shifted to working on Monday through Thursday and alternated between working Friday or Saturday for the rest of the field period.

WaSH electronic data capture platform:

For the WaSH survey, the CommCare Open Data Kit (ODK) platform was used for electronic data capture in the field on Android tablets. The completed, encrypted forms were sent to the server at RuralNet at the end of each day and the data was stored in its encrypted form. The data on the server was only available to a user with security clearance (username and password). A back up of the data was created on a daily basis to a secondary drive.

Three staff members from CDC-Atlanta, one staff member from CDC-Zambia, and two staff members from NORC provided training and supervision for the WaSH survey. The overall field coordinator was Kelvin Munjile from RuralNet Associates.

Questionnaires

The questionnaires for the Impact Evaluation of the Lusaka Water Supply, Sanitation and Drainage Project were designed to evaluate MCC key indicators and answer research questions in the Evaluation Design Report. The WASH household questionnaire included 12 modules: household demographics; sickness (prevalence of diarrhea and influenza-like illness) and associated costs; caretaker time loss; water collection; water storage and treatment; sanitation; hygiene; household garbage disposal; time use and expenditures; observations; and field control information for case management. The WASH household questionnaire was administered to 12,514 randomly selected households over a 12 month period.

The household drainage survey included the above 12 modules (with abbreviated WASH modules) and an additional module on flooding and the impact of flooding on household time use. The household drainage questionnaire was administered to 3,150 randomly selected households over a 4 month period. The business drainage survey primarily measured the impact of flooding on business activity and contained sections on business characteristics; flooding; illness; and revenue loss and property damage due to flooding. The business survey was administered to 587 systematically selected businesses in Lusaka during a 3 week period.

All three questionnaires were written in English and translated into Nyanja and Bemba, the two most prominently spoken tribal languages in Lusaka. For the drainage surveys (household and business), the English questionnaire was translated into Nyanja and Bemba by a local translation firm, RangeView Enterprises. The translated questionnaires were then back-translated into English by the same firm. For the WaSH household survey, the RuralNet Field Manager led the survey translation task. The draft English survey was translated into Nyanja and Bemba prior to the pre-test. As edits were made to the English version, the translated surveys were also updated. Feedback on the translated surveys was implemented based on comments from the pre-test, during training and the pilot.

Data Collectors

Name	Abbreviation	Affiliation
National Opinion Research Center at the University of Chicago	NORC	
RuralNet Associates, Ltd.	RuralNet	

Supervision

For the baseline evaluation, interviews were conducted by teams of enumerators. The composition of teams varied

depending on the survey (WaSH, household drainage, business drainage, traffic):

Twenty-four enumerators were hired for the household drainage survey. Four of those enumerators were selected to be data entry operators. Out of the remaining 20 enumerators, four were chosen to be field supervisors. Thus, the household drainage survey consisted of four enumeration teams each consisting of one supervisor and four enumerators. The four survey teams shared two buses with two drivers. The supervisors coordinated with each other on drop off and pick up in their various SEAs. In order to finish data entry and double data entry by the end of the contract, two additional data entry operators were hired from the short list of business survey enumerator candidates midway through the drainage evaluation (March 21st, 2016).

A total of eight enumerators were hired for the drainage business survey on March 5th, 2016 and were trained from March 15th-18th, 2016. The enumerators selected amongst themselves one person to be nominated as supervisor. The supervisor had the additional responsibilities of allocating replacement businesses, coordinating with the driver and reporting on problems encountered in the field.

Thirty one enumerators were hired for the traffic study as part of the drainage baseline evaluation and were trained from January 26th-27th, 2016. They were grouped into 3 teams including supervisors. Two teams of 14 each (12 enumerators, 2 supervisors) were assigned to vehicle counts initially for 12 hour periods at all primary and secondary intersections. A separate team of 3 enumerators was assigned for pedestrians count at those intersections and were alternately rotated between the two main teams.

The role of the field supervisors was to assist MCA-Z and CDC in the supervision of the data collection staff (the enumerators). These supervisors played an important role in the continuous training of enumerators and in ensuring the quality of the survey data. Specifically, field supervisors were responsible for reviewing each of their team members' completed questionnaires at the end of each day for completeness and consistency, observing their team members interviews to ensure that questions were being asked appropriately and answers were being recorded correctly, meeting with enumerators on a daily basis for performance reviews, helping enumerators resolve any field issues including locating assigned households, understanding the questionnaire, dealing with difficult respondents, etc., and performing back checks on household geolocation points to ensure that enumerators interviewed the correct households/respondents.

The general responsibilities of the field enumerators were to locate their assigned households and complete the written questionnaire with a male or female head of household, to accurately identify all persons living in the households with careful emphasis on recording sickness episodes accurately, to check their completed interviews for legibility and accuracy for ease and accuracy of data entry, and to make return attempts to household where respondents were not available during the initial visit.

The data entry operators served as the link between the hard copy of data and the digital form. Their responsibilities included accurately entering data from the paper questionnaires into the REDCap database each day, reporting issues with the database, properly filing and safeguarding the paper questionnaires at the MCAZ office, and following up with field staff over illegible responses in the paper questionnaires. Double data entry was performed and all errors were reconciled before the close of the survey to ensure accuracy of the data.

Enumerators for the WaSH baseline survey were recruited in three waves. The first round of recruitment and training took place from 19 September to 6 October, 2016. Twenty staff were selected for the main fielding and two staff were selected as standby enumerators. The second recruitment effort took place from 6 March to 24 March, 2017 and the third recruitment effort took place from 22 May to 5 June, 2017. Five enumerators from the March training and five enumerators from the May training were engaged for a permanent position (two enumerators chose not to stay with the project following the March training). Supervisors were selected based on their leadership qualities and objectivity. Overall, four enumerators were selected to serve as supervisors and lead four enumeration teams of four enumerators per team. After the second and third recruitment efforts, a fifth supervisor was selected and a fifth enumeration team was formed. The additional enumerators from those recruitments were also added onto existing teams giving 5 field teams with 5 supervisors and 4-5 enumerators per team. In summary, there were 4 field teams between months 1-6 of data collection and 5 field teams between months 7-12 of data collection.

Supervisors were responsible for interacting with local officials, ensuring that households were correctly located, making sure enumerators complied with all consent and confidentiality requirements as approved by the IRBs, implementing the household replacement strategy (when necessary), answering enumerator questions, observing interviews, participating in geo-coding, conducting interviews if needed, arranging the transport of water samples to the lab, tracking and reporting field issues, reviewing field control sheets, meeting with enumerators to discuss performance, and verifying the completeness and internal consistency of the questionnaires before transmission to the central office.

Enumerators were responsible for geo-coding households, traveling to selected households, obtaining consent from the respondent, and administering the survey, following consent protocols and securely storing signed consent forms, collecting and/or testing water samples at selected households, distributing chlorine, transmitting electronic surveys to server at the end of each day, and informing the supervisor of any issues, challenges or recommendations. The enumerators were also responsible for reading the consent form and gaining consent in the language the head of household felt most comfortable conversing in (English, Nyanja, or Bemba).

Each team had one driver who transported the field team from RuralNet's Central Office to their daily assignment and transported the team back to the Central Office at the end of the day. The driver also transported water samples to the lab.

For the drainage and the WaSH baseline surveys, CDC-Atlanta and CDC-Zambia had staff in the field observing interviews as a quality control measure.

Data Processing

Data Editing

WaSH Survey:

Data processing took place in the RuralNet central office in Lusaka, and was led by the RuralNet data manager, Mofya Phiri. The tablet program included controls that avoided the input of inconsistent data, both in terms of defining valid ranges for each data item and in logical checks that ensured that the appropriate skip patterns for the questionnaire were followed. Data processing consisted of reconciling the number of households approached and interviews completed with the number of records in the database. NORC wrote program logic that checked for inconsistencies and out of range values. Code was written in SAS and run against the data exported from ODK. CDC also wrote SAS code that was run on a weekly basis to identify interview records that needed additional review and to look for systematic errors. All errors or inconsistencies were logged and the resolutions recorded. Data cleaning was carried out after the survey was completed for detecting and correcting, removing, or flagging incorrect data, errors in format, incomplete data, inconsistent data, etc. NORC's team members also conducted nightly data extracts and methodological analysis of the questionnaire and case management data during the data collection field period to assess the completeness of responses, the consistency of responses, the quality of data captured, the effectiveness of the instruments in capturing the measures of interest, and the performance of the field staff, which will be summarized in the field survey report. MCAZ, CDC, RuralNet and NORC also held weekly conference calls to discuss observations and concerns and share ideas.

During the creation of the database and underlying data schema, NORC focused on file construction, variable names and labels, and the ability of the database to function as required for MCA-Z and CDC evaluation goals. This facilitated the delivery of uncleaned data files, the interim data files and the final data file. All data files were securely delivered to CDC in encrypted format using a secure file platform, NORC's SFTP site. Data were delivered in SAS format per instruction from MCA-Z and CDC. NORC submitted weekly interim data files to CDC. The data codebook for all variables included relevant instructions and documentation on interpreting data files as well as any data cleaning. The codebook detailed variable names and labels, type of variable and values assigned during coding.

Drainage Survey:

All of the drainage surveys were completed in hard copy on paper questionnaires. Therefore, consistency checking was the responsibility of both the enumerator who conducted the interview and their supervisor before handing the completed survey over to the data entry operators for data entry and verification. The data cleaning and consistency checks inherit in the REDCap database and the double data entry process are outlined in "data entry process" section.

Additionally, CDC staff conducted weekly downloads of survey data from REDCap into SAS v.9.4 in which a standardized code was run to look for inconsistencies and standard errors. Anomalies in the data were examined and stratified by enumerator in order to assess whether interview questions were being asked accurately in the field. CDC talked to field staff directly when consistent errors were noted and then brought issues to the group anonymously so those mistakes in data collection would not be repeated by the rest of the staff.

Other Processing

Software: Data were entered into a REDCap online database. REDCap (Vanderbilt University, TN) is a server based online database platform. The data entered were stored on a secure server located at CDC Atlanta. Access was restricted to authorized personnel only. Double data entry reconciliations were performed using SAS v. 9. 4 (SAS, Cary ,NC).

Note: Data entry was only conducted for the drainage baseline surveys since they were conducted on paper instruments and the WaSH baseline surveys were electronic data capture.

Data entry for the household drainage survey was conducted at the MCAZ head office in Lusaka. Four data entry operators were originally hired with two additional operators being hired midway through the survey to assist with completion of all data and double data entry by the end of the contract period. The six data entry operators manually entered hard copy survey data into a REDCap database on six secure, CDC laptops Monday through Friday from 8:00am-4:30am at a run rate of approximately 14 surveys entered per operator per day. All range checks and skips were controlled by REDCap and the operators could not override them. Range and consistency checks were also incorporated into the REDCap database.

A second, identical REDCap database was used for 100% verification of all variables using independent verification, i.e. double data entry. A second data entry operator would enter a survey into this database and then CDC staff would export the files from REDCap to SAS and then compare the first (original) record to the second (duplicate) record using the "proc compare" command to receive a report on inconsistencies. The data entry operators would then receive a print out of the inconsistencies report and work together to rectify those inconsistencies (using the hard copy as the gold standard) until the

records were identical.

Three enumerators from the traffic study were hired as data entry operators for the traffic study data entry task and five enumerators from the business survey were hired as data entry operators for entering the business study data. Data entry and verification were performed in the same manner as outlined above for the household drainage survey.

After the overall completion of data entry, access rights to REDCap were removed for all non-CDC personnel and the hard copies of the completed questionnaires were moved to CDC-Zambia office premises.

Data Appraisal

No content available