



UNIVERSITY OF LEEDS

This is a repository copy of *Learning by design: lessons from a baseline study in the NAMWASH Small Towns Programme, Mozambique*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/123833/>

Version: Accepted Version

Article:

Barrington, DJ orcid.org/0000-0002-1486-9247 and Admiraal, R (2014) Learning by design: lessons from a baseline study in the NAMWASH Small Towns Programme, Mozambique. *Waterlines*, 33 (1). pp. 13-25. ISSN 0262-8104

<https://doi.org/10.3362/1756-3488.2014.003>

(c) Practical Action Publishing, 2014. The definitive, peer-reviewed and edited version of this article is published in *Waterlines*, volume 33, issue 1, 13-25, 2014, <https://doi.org/10.3362/1756-3488.2014.003>.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Learning by Design : Lessons From Designing a Baseline Study in the NAMWASH Small Towns Programme, Mozambique

Dani J. Barrington^{1*}, Ryan Admiraal²

*Indicates corresponding author

Dani Barrington is a Research Fellow at Monash University, Australia, and Ryan Admiraal is a Lecturer at Murdoch University, Australia.

1. School of Engineering and Information Technology

Murdoch University

90 South Street

Murdoch

Western Australia 6150

Australia

E: Dani.Barrington@monash.edu

T: +61 412 528 402

2. School of Engineering and Information Technology

Murdoch University

90 South Street

Murdoch

Western Australia 6150

Australia

E: R.Admiraal@murdoch.edu.au

Abstract

This paper examines the design and application of a baseline study for a comprehensive water, sanitation, and hygiene (WASH) intervention in Mozambique. The study was developed to investigate the relationships among key parameters of interest both for comparison to post-implementation data and to contribute to planning the WASH intervention itself. We use this study to discuss key issues surrounding baseline studies. This includes providing guidelines for designing a WASH baseline survey, determining an appropriate sample size, and highlighting key considerations in analysing the survey data, such as incorporating the study design in statistical analyses, post-stratifying and utilising geospatial data. We also show how statistical analyses from a baseline survey can be used to inform subsequent surveys. For example, results from this study suggest that in future WASH studies, self-reporting by households should be supplemented by observational or population data to remove or quantify reporting bias, and care must be taken to reduce respondent fatigue.

Introduction

This article examines the design of a baseline study, including monitoring of non-target communities, for a comprehensive water, sanitation and hygiene (WASH) intervention program in Mozambique. It explains how the study itself was designed and implemented and discusses some of the potential improvements identified following data collection. It is not intended to report the results of this study but instead provides practical insights into the design and evaluation of future WASH baseline studies.

This baseline study was conducted in Mozambique as part of the Nampula Province Water, Sanitation and Hygiene (NAMWASH) Programme. The Programme is aimed at accelerating the achievement of the WASH Millennium Development Goals (MDGs), particularly Goal 7c: “Halve, by 2015, the proportion of the population without sustainable access to safe drinking water and basic sanitation” (United Nations, 2013) in Nampula Province. The project aims to significantly increase access to safe drinking water and sanitation, the adoption of appropriate hygiene practices, improved water safety management and WASH facilities in schools between 2012 and 2016. The program also aims to strengthen the technical and managerial capacity of government, private sector and civil society in small towns and create a WASH intervention model for small towns that can be further applied throughout Mozambique. A baseline study was required to aid in the design of the WASH interventions as well as to provide information on the efficacy of the program post-intervention. It was also intended that the design and assessment of this baseline study would contribute to methodological knowledge on how to undertake impact evaluation of WASH activities in small towns across the globe.

The baseline study consisted of seven towns, where five target towns are expected to receive WASH interventions as part of the NAMWASH program, and two non-target towns are not expected to receive WASH interventions within 3-5 years following the baseline year. Several key players were involved:

- UNICEF Mozambique, who designed the baseline study as the first stage of a multi-year longitudinal study of WASH interventions in towns in Nampula province,
- A national consultant who carried out data collection,
- Nampula Provincial Directorate of Health, who spearheaded a multi-sectorial team collecting water quality data, and
- Murdoch University, who carried out statistical analysis of data from the baseline study, including post-survey quality control.

In total, the baseline study took five weeks to complete (early September to early October of 2012), and the impact study comprised approximately 2% of the total budget.

Baseline studies and WASH interventions

After several decades of water, sanitation and hygiene (WASH) interventions in developing communities, it has become apparent that the best results are achieved when technical and educational programs relating to each of the WASH subthemes are integrated (e.g. Aziz et al., 1990). It has also become apparent that progress towards individual MDGs is closely linked to, and may assist in the attainment of, other MDGs. As an example, several MDGs, including those related to WASH, have been shown to significantly contribute to the goal of reducing child mortality (Gakidou et al., 2007). As such, integrated WASH interventions are being implemented across the globe.

The collection and analysis of community-level data can assist in the development of appropriate integrated WASH interventions (Huber and Mosler, 2013), and an appropriate baseline study can provide such useful data. A well-designed baseline study allows for the planning of an effective WASH intervention program by identifying areas of concern as well as allowing for the examination of relationships among various WASH indicators. For example, the World Bank's Water and Sanitation Program has collected baseline data to inform designs for more effective hand washing interventions in various communities across the globe (Perez et al., 2011).

It is inherently difficult to design a baseline study for a multi-arm intervention that will include not only aspects of the various WASH subthemes (water, sanitation and hygiene) but also range in methodology from infrastructural improvements to education programs. Such projects require investigators to monitor physical parameters (such as water quality), behavioural changes and attitudes towards WASH. It is thus important to design a baseline study that incorporates methods which will assess all of the indicators that will both contribute to WASH intervention design and determine its efficacy post-implementation. With sound statistical design and analysis, such baseline studies can uncover valuable relationships which can inform design of both the current and future WASH interventions. Guidelines for good WASH baseline studies are included in Table 1, and suggestions for good WASH follow up studies in Table 2.

Table 1: Guidelines for good WASH baseline studies.

Include non-target communities	Despite Blum and Feachem's seminal paper on the weaknesses of inferring health outcomes of WASH interventions from inadequate baseline studies (Blum and Feachem, 1983), many WASH studies do not include baseline sampling of communities not expected to receive interventions (non-target communities). However, monitoring non-target communities is important for post-implementation analysis. Their inclusion in both the baseline study and post-implementation monitoring allows for analysis of the efficacy of WASH interventions by measuring changes in WASH indicators which occurred without WASH interventions. (for a Bangladeshi example of a thorough WASH study including monitoring of non-target communities see Aziz et al., 1990).
Conduct pre-, post- and interim monitoring during similar timeframes or seasons of the year	Many WASH baseline studies suffer the inherent flaw of temporal effects. Where studies cannot be performed over a long period (as is almost always the case), reporting of WASH factors may only represent their value at that point in time. For example, the volume of water collected may differ significantly between wet and dry seasons, and the incidence of diseases, such as the many which cause diarrhea, may vary throughout the year (Alexander et al., 2013). Although it is likely impossible to increase the period of data collection, especially when conducting household surveys where data is collected at only a certain point in time, seasonal effects can be accounted for to some extent by keeping sound records of climate in the weeks or months prior to, and performing post-intervention monitoring during the same time of the year as, the baseline study. More useful would be to collect data on seasonally affected WASH variables in both the target and non-target clusters regularly throughout the intervention period to determine seasonal patterns and the overall change in the variables. Such patterns of variables may not be obvious from collecting data from single points in time pre- and post-intervention.

<p>Pay careful attention to the selection of impact variables</p>	<p>Variables need to be chosen that can realistically be measured in the development context. Those which cannot be measured directly during monitoring and instead rely on government records may prove difficult or impossible to gain access to and may not be of the quality required to perform meaningful analyses. Additionally, some variables may be more accurate than others. For instance, using reported handwashing practices as a measure of hygiene may be a far less reliable measure than observed handwashing practices.</p>
<p>Provide relevant information for sample size calculations for follow-up studies</p>	<p>In determining an appropriate sample size for the baseline study in this work, the percentage of households using improved water supplies was used as the key indicator. Post-intervention studies will almost certainly be interested in a variety of other key indicators (e.g. prevalence of diarrhea, water quality at the source or household, percentage of people using soap when washing hands), and the baseline study can provide current measures of those key indicators that will inform appropriate sample sizes for the next study.</p>

Table 2: Suggestions for good WASH follow-up studies.

<p>Where possible, using a randomized controlled trial</p>	<p>When designed appropriately, these provide the best assessment of the effect of an intervention by essentially eliminating confounding. In the context of a WASH study, a truly randomized controlled trial at the household level would be difficult to implement across all sub-themes, especially water supply, where introduced improved water sources may in many cases be public standposts, not private taps. Additionally, Cairncross and Valdmanis (2006) argue that such a design at the household level poses potential ethical and political dilemmas. Such randomization is possible at the community level, though, provided that the geographic space is sufficient to lead to limited interaction among the various communities. Ethical and political considerations could still come into play in this context, but we note that, based on the NAMWASH baseline study, the proposed control communities actually had greater access to improved water sources than target communities (chi-square test p-value < 0.001). Additionally, as previously mentioned, an external control (as opposed to internal control) can provide a measure of the shifts in the WASH subthemes that would naturally have occurred over time, and that allows us to better understand the true impact of the intervention(s).</p>
<p>Investigate multiple intervention arms.</p>	<p>Designs such as factorial designs provide a means to tease apart the individual effects of the WASH subthemes as well as any interactions among them. Simply providing all household with improved water sources, sanitation, and hygiene does not allow the researcher to understand the individual impacts of these components on health indicators, nor does it allow researcher to understand how these may be linked. For instance, as observed in the NAMWASH baseline study, water quality tends to deteriorate from source to home, so providing improved water sources in the absence of sanitation and/or hygiene education may have negligible (if any) impact on improved health outcomes. A study that included interventions for only individual WASH subthemes could possibly conclude that providing improved water, improved latrines, and better hygiene has no health impacts, while a study that has one intervention incorporating all three may fail to understand that the three work in tandem to improve health.</p>

Study Design

The baseline study was designed by UNICEF Mozambique and the Government of Mozambique based upon the project Terms of Reference provided by the Programme team. The Programme was designed to be a longitudinal study, where both target and non-target towns would be monitored pre-

and post-WASH project implementation. In order to assess the many facets of a WASH intervention, particular indicators were selected which would be investigated during the baseline study. The indicators were chosen based upon existing programme- and donor-specific indicators and AusAID's Performance Assessment Framework criteria. The baseline study was designed to specifically investigate these indicators prior to WASH intervention, including relationships among parameters of interest which may inform the design of the intervention. (For example, is water quality better for households using improved water sources? Is there a difference in male and female use of school toilets?) The baseline study was designed such that similar monitoring could be conducted following the implementation of WASH interventions so as to determine their efficacy, thus meeting the reporting requirements of donors and informing future interventions.

The sampling methodology was designed based upon the Rapid Assessment of Drinking Water Quality Approach (WHO and UNICEF, 2012) and that developed by Bennett et al. (1991).

Baseline tools

Tools were developed based upon the indicators selected for investigating WASH behaviours and practises. The primary tools developed were surveys, which included a mixture of qualitative and quantitative data collection methods. The surveys were designed to be conducted at households, schools and water points and included questions of the respondents and observations by the enumerator. Three different surveys were carried out, covering four different target populations. One survey was designed to be administered to a random selection of households and included not only household-level questions but also individual-level questions, a second survey to a subset of schools, and a third survey covered the primary water sources identified by survey respondents.

Sample size determination

Determining appropriate sample sizes for the four different target populations (households, individuals, schools, primary water sources) could have been done separately if these were distinct populations. The appropriate sample size for each target population depends in part on the particular variables that will be measured through the survey questionnaire as well as the statistical analyses that will be utilized, so a set of key indicators/variables and analyses are typically decided upon for which to calculate minimum required sample sizes. The largest of the resulting minimum required sample sizes is then used as the sample size for that target population.

Such calculations first assume that a simple random sample will be drawn from the population, so typical sample size calculations for simple random samples hold. For example, in the case of point estimates for proportions, the sample size n_{SRS} should satisfy

$$n_{\text{SRS}} \geq \frac{z^2 p(1-p)}{e^2}$$

where p is an estimated proportion for a binary variable of interest and z is the standard normal quantile corresponding to a desired significance level. When considering a proportion, the margin of error e is typically set at $e = 0.1p$. If the expected range of values of p is unknown, values close to 0.5 produce more conservative estimates for the sample size, so, oftentimes, $p = 0.5$ (and $e = 0.05$) is used by default to produce a conservative estimate.

In practice, simple random samples are often rejected in favour of stratified or cluster sampling due to monetary savings, the ability to reduce sample sizes and to address demographic or geographic heterogeneity. In these cases, the sample size is adjusted to account for the more complex sample design, and

$$n \geq n_{\text{SRS}} \times \text{deff}$$

where *deff* is the design effect. The design effect is the ratio of the standard error for a variable under a complex sample design to the standard error for the same variable under simple random sampling. Many software packages now provide built-in functions to easily calculate design effects for variables from existing survey data. When design effects for relevant variables cannot be estimated from existing surveys with similar designs to the proposed survey, design effects can be estimated as

$$\text{deff} = 1 + \rho(n_c - 1)$$

where n_c is the proposed sample size for each cluster, and ρ is a measure of within-cluster correlation for the variable of interest (Lumley, 2010).

Finally, non-response is likely in most surveys. If the minimum required sample size is used, then non-response may affect the ability to estimate key indicators to a certain level of precision or reduce the power of statistical tests, making them less likely to be able to detect significant differences. Consequently, it is important to inflate the minimum sample size to adjust for non-response. For example, if the required minimum sample size is 100 and there is 5% non-response, then the actual number of respondents when surveying 100 respondents would be expected to be 95. If we inflate the number of people to survey by $1/0.95$ and survey 106 people, the resulting expected number of respondents is $106 \times 0.95 = 100.7$ people, ensuring that we achieve the minimum required sample size. This means that to achieve a specific number of respondents, it is necessary to divide the minimum required sample size by the expected response rate *RR*. Thus, if we wanted to estimate the required sample size for a complex sample design for, say, a point estimator for a proportion, this required sample size would be given by

$$n \geq \frac{z^2 p(1-p)}{e^2} \left(\frac{\text{deff}}{RR} \right) \quad (1)$$

For the NAMWASH study a minimum required sample size for households was calculated using (1) and

$z = 1.96$ (corresponding to a 5% significance level),

$p = 0.431$ (percentage using improved water sources, as estimated from the 2008 Multiple Indicators Cluster Survey (MICS) (National Statistics Institute of Mozambique, 2009)),

$e = 0.05$,

$\text{deff} = 4$ (standard for WASH surveys, where *deff* typically ranges from 2 to 10),

$RR = 0.90$ (i.e. 10% non-response).

This produced a minimum sample size of $n = 1,658$, which was rounded to $n = 1,660$. An alternative calculation using a response rate of $RR = 0.97$ and $p = 0.362$ (the percentage of households using improved water sources in the seven towns included in the NAMWASH study as estimated by the 2008 MICS) produced a lower minimum sample size of 1,462 households. Ultimately, 1,610 households were randomly sampled, 7 from each of 230 clusters of approximately equal size. 70% of clusters were from target towns and 30% from non-target towns. This design was approved by INE (National Statistics Bureau).

Such sample size calculations were not carried out for the other target populations. For the water points survey, all primary water sources used by sampled households were included in the survey, and for the schools survey, a total of forty schools were sampled with thirty located in target towns and ten in non-target towns.

Non-target towns were selected based upon their similar size to target towns and because they were not expected to have WASH interventions within the 3-5 years following 2012. However, some clusters that were not expected to receive WASH interventions may do so during the implementation stage and vice versa.

Data collection

Data collection: Methods utilised in the NAMWASH study

Local enumerators were recruited to conduct the surveys. They had each completed a minimum of secondary schooling and were fluent in the local language, Macua. The enumerators were trained over a one week period, including the pre-testing and adjustment of the survey tools. The enumerators were provided with field manuals and a survey protocol document before entering the field. They also carried with them a GPS manual and drawings of water point and latrine designs.

The household and water point surveys were conducted by six enumerators working in teams of two, with the composition of these teams changed each day to minimise the enumerator effect. An additional enumerator completed the school surveys in each district. The household, water point and school surveys took approximately five weeks to complete. Two weeks following the completion of household surveys, water quality testing was conducted at 5% of households and their primary water source. At each of these households and water sources three replicates were collected and tested for microbiological (thermotolerant coliforms and coliform forming units) and physicochemical parameters (turbidity and pH).

Data collection: Discussion

Sources of bias

Practitioners familiar with survey techniques will be aware of the many biases possible due to response effects. Those of particular concern when conducting baseline surveys include:

- Non-response bias: Respondents may not answer some questions due to the nature of the question, and those who choose to respond may be fundamentally different from those who do not. In the case of the NAMWASH baseline study, no one declined to take part in the survey, and the rate of non-response was negligible for variables of interest.
- Response bias: Respondents may answer a question in a certain way because they believe that a certain response is desirable, even if it is not true.
- Respondent fatigue: Long surveys may lead to apathy or decreased attention to later items, increasing the likelihood of non-response or reporting errors.

Response effects were observed during the NAMWASH study, particularly response bias due to self-reporting. For example, nearly all primary caregivers of children reported washing their hands before eating or feeding children, after using the toilet and after cleaning up children's faeces, yet only 17.5% (15.5%, 19.7%) of households had handwashing facilities. At the same time, enumerators observed that, even where respondents identified that they washed their hands with soap or ash, only 25% (11.5%, 43.4%) of these respondents actually used soap or ash when asked to demonstrate their handwashing regime. Both of these are classic examples of response bias where questions are answered based upon how respondents believe they are expected to respond. Another example in the NAMWASH study arose where households were asked to estimate their daily water usage, with some estimating up to 420L per day. Such an amount is very unlikely, given that in most cases this water must be manually collected and carried to households. Such overestimations of behaviour are

commonly seen during WASH studies (e.g. Arnold et al., 2009), although it can be unclear as to whether this is due to respondents answering how they believe is appropriate, or difficulties in performing estimations themselves. Supplementing self-reporting data with observational and health data will assist in removing such effects.

There has been some post-monitoring concern around the potential occurrence of respondent fatigue in the long household survey of the NAMWASH study. This was particularly observed in a series of agree/disagree questions towards the end of the household survey, where a response pattern emerged suggesting that respondents may have initially assumed that the questions were meant to be answered affirmatively before realising that this was not always the case and, subsequently, more carefully considering the questions. For studies where the potential for this is identified in advance, measures can be taken, such as randomising the order of questions between households, reducing the length of surveys, or reducing the difficulty of questions (Lee et al., 2004).

Data analysis

Data analysis: Methods utilised in the NAMWASH study

Point estimation and statistical modelling for complex surveys

In carrying out statistical analyses based on the resulting survey data, it is important that the researcher take into account the sample design. In the case of the NAMWASH household survey, exact cluster sizes were not known, but clusters were based on enumeration areas (provided by INE) and were assumed to be of roughly the same size. Equal sample sizes in each cluster would result in a PPS sample, in which case inclusion probabilities (and, hence, sample weights) would be the same across all households.

For the schools survey, the number of schools randomly sampled from target and non-target towns was known, as was the total number of schools in each town. This latter information allowed us to post-stratify on town, producing differential sample weights depending on the town from which the school was sampled. For the water points survey, sampling of primary water sources was based on those used by sampled households. Without information as to the number of primary water sources and how many people use each water source, it was not possible to compute inclusion probabilities for the surveyed water sources, and sample weights were not incorporated in analyses for the water points survey.

Based on the sample design and resulting sample weights, point estimates and standard errors can be calculated, as well as standard statistical models fit and most standard hypothesis tests performed. Point estimates and standard errors are typically obtained using the Horvitz-Thompson estimator and corresponding variance (Horvitz and Thompson, 1952),

$$\hat{T} = \sum_{i=1}^n \frac{X_i}{\pi_i}$$

$$V(\hat{T}) = \sum_{i=1}^n \sum_{j=1}^n \frac{X_i X_j}{\pi_{ij}} - \frac{X_i}{\pi_i} \frac{X_j}{\pi_j}$$

which gives the point estimate and variance for a population total, where X is the variable of interest, π_k is the inclusion probability for individual k , and π_{ij} is the probability that both individuals i and j are sampled. Means and proportions are simply a function of the total, so they can be easily calculated from the Horvitz-Thompson estimator.

Parametric tests (e.g. t-tests, chi-square tests), non-parametric tests (e.g. Wilcoxon tests), and other common statistical models (e.g. generalized linear models) are easily extended to more complex survey designs, but we leave the details of this to others (e.g. Lohr, 1999, Lumley, 2010), as it is not

the intent of this article to provide a full exposition on such methods. Most major statistical software programs (including SAS, R, Stata, and SPSS) include functionality for such analyses for complex survey designs.

Data analysis: Discussion

Incorporating population totals

Where available, census data or clinic records can be used to remove or reduce bias in estimates that are based upon survey data, assuming the catchment areas these records refer to are well defined. It was originally intended that the survey data collected in the NAMWASH study would be supplemented by health data from local clinics. Unfortunately, due to the way data is collected and stored by hospitals and district health authorities in Nampula Province, alongside the relatively short study timeline, such data could not be obtained. In future studies it would be highly beneficial to source this information, particularly for health data that could be biased by self-reporting.

If census data are available and they coincide with variables reflected in a survey, they can be used for post-stratification. For instance, if the total number of males of type k in the population is known to be N_k and we post-stratify on the types for males, then the sum of the sample weights w_1, w_2, \dots, w_{n_k} (where $w_i = 1 / \pi_i$) of the n_k sampled males of type k should satisfy

$$\sum_{i=1}^{n_k} w_i = N_k.$$

Post-stratification consists of renormalizing the weights to ensure that this relationship holds. In the case of the schools survey, we post-stratified on town population. Post-stratification can also be used to adjust for non-response, provided that data are missing completely at random (Lohr, 1999).

If census data or clinic data are available for a bivariate (or multivariate) relationship consisting of an outcome of interest and a variable (or variables) for which that outcome is measured, these can be used to reduce bias in the estimates and variances of regression parameters for generalised linear models (GLMs). For example, estimates of the relationship among mortality due to cholera, age, and the water source and latrine type used by the household from a logistic regression model can be improved by incorporating mortality rates due to cholera by age as derived from clinic data. Handcock et al. (2005) show how population totals (e.g. mortality rate due to cholera by age) can be incorporated as a set of constraints on the parameters corresponding to variables for which these population totals exist, and the likelihood for the GLM is then maximised subject to these constraints. Naturally, the parameter estimates subject to constraints (e.g. age of the person) are improved, but the resulting parameter estimates for those variables not subject to constraints (e.g. water source type, latrine type) are also improved. If data are not missing for any of the variables in the model, the maximum likelihood estimators for the parameters of this model are asymptotically unbiased, efficient, and Gaussian (so standard Wald tests can be used). This method can also be used in the presence of non-response and does not require the highly restrictive assumption that data be missing completely at random. In fact, Handcock et al. (2005) show that in cases where data are not missing completely at random, the approach still reduces bias in parameter estimates.

Geospatial Data

Increasingly, Global Positioning System (GPS) data are recorded with a variety of surveys. For instance, in the NAMWASH study GPS coordinates were recorded for both household locations and water points. These open up possibilities for both descriptive and inferential statistics. Plots of household locations and water points can reveal clustering of certain outcomes according to geographic area, allowing for more concentrated exploration of why these outcomes are occurring in

specific locales. For instance, high levels of diarrhea, cholera, or other water-borne diseases in households close to particular water sources (akin to John Snow's plot of cholera deaths (Snow, 1855)) may reveal issues with specific water sources. Additionally, plots of household locations can reveal sampling anomalies that may otherwise go unnoticed, such as households sampled within a geographic cluster tending to lie in a concentrated area of the cluster.

GPS data can also be useful in terms of distance measurements. This can be used to produce new variables related to distance or to measure reliability of distance measurements reported by respondents. For instance, respondents in the NAMWASH household study reported the distance to their primary water source. For those for whom the location of this water source is known, accurate distance measurements based on GPS coordinates can be used in place of such reported distances, and it can be used to detect under- or over-reporting of distances. If locations for all water sources are not known, systematic biases in reporting provide a sense of the reliability of analyses based on reported distances. Additionally, distance measurements can be used to model correlation among observations when evidence exists for such dependence (Cressie, 1993).

The collection of GPS data can also play a role when examining relationships among individual households. In a multi-arm WASH program, the overall intervention will consist of many smaller interventions (e.g. educational campaigns, infrastructural improvements and microfinance programs), and it is highly unlikely that every target household will receive all of these. This has been identified as a potential issue in the NAMWASH study, particularly because it is expected that urban target households will receive more WASH interventions than rural target households. Knowledge of which households received each intervention, alongside their geospatial location, will allow for the examination of WASH variables at individual households compared to their distance from town centres, coupled to which WASH interventions they received. Similarly, GPS data will assist in assessing whether WASH interventions in some clusters influence the WASH variables of neighbouring clusters who did not receive interventions.

Conclusion

The results of WASH interventions are often reported, but it is uncommon for WASH practitioners to either share the methodologies used to determine the efficacy of said interventions or discuss ways in which the study might reasonably have been improved. Even comprehensive WASH baseline studies such as that developed for the NAMWASH study can be improved with hindsight. Analysis of the study post-data collection has suggested multiple additions which could be incorporated into future studies so as to enhance our understanding of relationships between WASH parameters. For example:

- a) Self-reporting data can be supplemented with observational or health data to remove bias;
- b) Surveys can be designed to prevent respondent fatigue;
- c) Population totals can contribute further information on specific parameters and remove some sources of bias from survey results;
- d) Geospatial data can be used for distance measures, pinpointing sampling anomalies and accounting for spatial correlation.

In order to improve the efficacy of WASH interventions across the globe, thus contributing to the achievement of the MDGs, it is crucial that the WASH community of practice share their experiences with both pre- and post-implementation studies, admitting where weaknesses lie and suggesting future improvements for both their own projects and those of others.

Acknowledgements

The authors wish to thank UNICEF and UNICEF Mozambique (particularly Dr Samuel Godfrey, Matteus Van der Velden and Alfonso Alvestegui), the DNA/AIAS Government of Mozambique (particularly Dr Olinda de Sousa) and AusAID Pretoria (particularly Dr Laila Smith) for their assistance with this study and the content of this paper.

References

- Alexander, K. A., Carzolio, M., Goodin, D. & Vance, E. 2013. Climate Change is Likely to Worsen the Public Health Threat of Diarrheal Disease in Botswana. *International Journal of Environmental Research and Public Health*, 10, 1202-30. doi:10.3390/ijerph10041202
- Arnold, B., Arana, B., Mausezahl, D., Hubbard, A. & Colford, J. M. 2009. Evaluation of a pre-existing, 3-year household water treatment and handwashing intervention in rural Guatemala. *International Journal of Epidemiology*, 38, 1651-1661. doi:10.1093/ije/dyp241
- Aziz, K. M. A., Hoque, B. A., Hasan, K. Z., Patwary, M. Y., Huttly, S. R. A., Rahaman, M. M. & Feachem, R. G. 1990. Reduction in diarrheal diseases in children in rural Bangladesh by environmental and behavioural-modifications. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 84, 433-438. doi:10.1016/0035-9203(90)90353-g
- Bennett, S., Woods, T., Liyanage, W. M. & Smith, D. L. 1991. A simplified general method for cluster-sample surveys of health in developing countries. *World Health Statistics Quarterly*, 44, 98-106.
- Blum, D. & Feachem, R. G. 1983. Measuring the impact of water-supply and sanitation investments on diarrheal diseases- Problems of methodology. *International Journal of Epidemiology*, 12, 357-365. doi:10.1093/ije/12.3.357
- Cairncross, S. & Valdmanis, V. 2006. Water Supply, Sanitation, and Hygiene Promotion. In: Jamison, D. T., Breman, J. G., Measham, A. R., Alleyne, G., Claeson, M., Evans, D. B., Jha, P., Mills, A. & Musgrove, P. (eds.) *Disease Control Priorities in Developing Countries*. Washington, D.C.: World Bank.
- Cressie, N. 1993. *Statistics for Spatial Data*, New York, Wiley.
- Gakidou, E., Oza, S., Fuertes, C. V., Li, A. Y., Lee, D. K., Sousa, A., Hogan, M. C., Hoorn, S. V. & Ezzati, M. 2007. Improving child survival through environmental and nutritional interventions - The importance of targeting interventions toward the poor. *Journal of the American Medical Association*, 298, 1876-1887. doi:10.1001/jama.298.16.1876
- Handcock, M. S., Rendall, M. S. & Cheadle, J. E. 2005. Improved regression estimation of a multivariate relationship with population data on the bivariate relationship. *Sociological Methodology*, 303-346. doi:10.1111/j.0081-1750.2006.00169.x
- Horvitz, D. G. & Thompson, D. J. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 663-685. doi:10.1080/01621459.1952.10483446
- Huber, A. C. & Mosler, H. J. 2013. Determining behavioral factors for interventions to increase safe water consumption: a cross-sectional field study in rural Ethiopia. *International Journal of Environmental Health Research*, 23, 96-107. doi:10.1080/09603123.2012.699032
- Lee, E., Hu, M. Y. & Toh, R. S. 2004. Respondent non-cooperation in surveys and diaries: an analysis of item non-response and panel attrition. *International Journal of Market Research*, 46, 311-326.
- Lohr, S. L. 1999. *Sampling: Design and Analysis*, Pacific Grove, United States of America, Duxbury Press.

- Lumley, T. 2010. *Complex Surveys: A Guide to Analysis Using R*, Hoboken, United States of America, Wiley.
- National Statistics Institute of Mozambique 2009. *Final Report of the Multiple Indicator Cluster Survey*. Maputo, Mozambique: National Statistics Institute.
- Perez, E., Pyne, H. H., Amelink, C., Briceno, B., Coombes, Y., Devine, J., Peschiera, R. F., Frischmuth, C., Grossman, A., Koita, S. N., Kullmann, C., Nguyen, M. T., Nguyen, N. K., Orsola-Vidal, A., Paynter, N. & Watkins, K. 2011. *Global Scaling Up Handwashing Project: Progress Report. Water and Sanitation Program: Progress Report*. Washington D. C.: Water and Sanitation Program.
- Snow, J. 1855. *On the Mode of Communication of Cholera*, London.
- United Nations. 2013. *Goal 7: Ensure Environmental Sustainability* [Online]. UN Publications. Available: <http://www.un.org/millenniumgoals/environ.shtml> [Accessed 28th July 2013].
- WHO & UNICEF 2012. *Rapid Assessment of Drinking-water Quality: A Handbook for Implementation*, Geneva, World Health Organization.