An Expert Reviews Our Impact Data

View Dr. Richard Carter's post here.

Real-world WASH programmes and their impacts

Wild, unjustified claims of impact. I imagine that we're all familiar with the claims made by many NGO WASH programmes about their "impact" on the health of their beneficiaries[1]. It seems to be all too easy for some programme managers – and presumably their funders too - to assume that if water supply, sanitation or hygiene facilities are improved in some way, then self-evidently the health of those who have received those services will have improved.

Randomised controlled trials. On the other hand, the results of numerous rigorously implemented randomised controlled trials (let's not mention the poorly conducted RCTs of which there are arguably too many) show that the true impacts, especially on infant diarrhoea, may be quite modest or non-existent. What then can we really expect when WASH projects are implemented routinely and competently? I have written about this in relation to sanitation and hygiene elsewhere (Carter, 2017), but this blog addresses the way in which a very professional WASH INGO can routinely generate evidence of its impacts – without the cost and logistical burden of a randomised controlled trial - producing highly credible findings in the process.

Water for Good/Lifewater. Recently one particular WASH INGO[2] posted in a LinkedIn message[3] a summary of its claims regarding the health and wellbeing of the communities it serves. This prompted me to send the following message, accompanied by 15 questions interrogating the methodology by which the claims were evidenced: "*I applaud the work that Lifewater/Water for Good is doing, and I am sure that it has some impact on the knowledge, practices and health of people in the communities where Lifewater works. However I am always cautious about claims of health impacts, especially in the absence of controls, and especially where respondent bias is likely*."

Water for Good's response. Lifewater (now Water for Good) responded very graciously and fully to my questions, and I reproduce their answers in full below. My reason for doing so is that I see Water for Good's methods as an exemplary, pragmatic, rigorous observational approach that does not carry the attributional claims of full randomised controlled trials (the only way of unambiguously establishing causality). Water for Good is at pains **not** to claim causality, but it also has intentions in future of undertaking evaluations that are capable of doing so. They are not the subject of this blog.

Monitoring impact pragmatically. What follows, the answers to my questions, sets out an approach which is measured and sufficient enough to provide evidence (but not absolutely rigorous attribution) of the impacts of WASH interventions. The remainder of this blog is reproduced with permission from Water for Good's replies to my questions. It is edited only in the interests of clarity.

Water for Good's Monitoring and Evaluation Approach

Baseline Surveys

Q1. Over what duration are your baseline surveys typically undertaken?

Our household knowledge, attitudes, and practices survey is the main data source for our evaluations. Our goal is to collect all necessary surveys within a two-week time window. We also collect institutional data from schools and health facilities. Some of this data is collected by Water for Good staff (including the institutional assessments, in which we document the WASH access at the institutions at a point in time) and some is provided by school/health facility records. These institutional data are collected during the project planning phase up until the baseline survey collection, and the data collected is for a fixed period (the previous school year, calendar year, etc.) to ensure we are not collecting partial data.

---

[1] I use the word "beneficiary" in a purely descriptive way to mean those who gain some benefit from a programme. The term carries no patronising or derogatory connotation.

[2] Lifewater – recently merged with Water for Good and adopting the latter name across its country programmes.

[3] Some more of the findings from Water for Good's projects are available online: Ethiopia, Kokosa Project 4, Tanzania Shinyanga P1 Endline Evaluation Report (2024), Cambodia Svay Leu P3 Endline Evaluation Report (2024).

Q2. At what time of year are the baseline surveys carried out (in relation to weather and farming seasons)?

Baseline surveys are collected +/- one month of starting the Vision of a Healthy Village (VHV) program in that area. The staff for that specific program (based in the area) provide information on the best time to start a new project, taking into consideration seasons and other potential influences. Budget and previous project activities are also taken into consideration. Therefore, there is not a set period for all baseline studies, but rather each project runs on an independent timeline regarding evaluations. Endline surveys are collected three years after the baseline, +/- a maximum of one month to reduce seasonal variation between baseline and endline datasets.

Q3. How are households selected for inclusion in baseline household surveys? What number of hh are typically surveyed?

Our goal is to sample at a 95% confidence level and a 5% margin of error, with an additional 25% buffer. We calculate this using a sample size calculator and the total number of households counted in the participating area during the household census that takes place during the project planning phase. In larger areas, sometimes we reduce the confidence level or increase the margin of error if too many surveys are required (it is difficult to collect more than 450 surveys). We do not go below a 90% CL or above a 7.5% margin of error. We typically collect 400-450 surveys and sample every government village in the project area usually using the probability proportional to size method, weighing the number of surveys per village by the population. Sometimes a Cluster Sampling Method is used to reduce the number of required surveys, if the villages are internally heterogenous (households in the cluster/village are as different as possible) and externally homogenous (the clusters/villages are as similar as possible).

Households are randomly sampled using the Spin a Pen Method, in which:

a. Enumerators determine the center of the village, with the help of community leaders (including elders, chiefs, government officials, etc.),
b. Enumerators spin a pen on the ground and walk in the direction it points,
c. Enumerators survey every third household they encounter (or cast a die to determine the number), and
d. If they reach the edge of the village before completing the required number of surveys, they return to the center of the village and start the process over.
e. Households are excluded if they do not consent to being surveyed, if no one is home at the time of surveying, or if the respondent is under 18.

Q4. In reporting on diarrhea, does the duration of survey only cover one week? How does diarrhea incidence vary over the year?

We apply a one week recall period when surveying diarrhea related questions to minimize recall bias, because:

a. A longer recall period may lead to underreporting, as people forget past illness episodes and a short recall period (7–14 days) helps improve accuracy in reporting, as respondents are more likely to remember recent episodes.
b. A shorter recall period ensures that data reflects current trends (point-in-time evaluation) rather than an averaged-out, potentially inaccurate picture over longer periods.
c. Diarrhea is a relatively frequent and short-duration illness, making it difficult for respondents to remember accurately if asked about a longer period.
d. To align with the National Demographic and Health Survey approach of many countries where we work and with approaches used by other organizations.

Diarrhea prevalence likely varies over the year, as you noted. The result we find for this question only represents a point in time. This is one of the reasons why we are strict about when we collect our endline data – we keep the variance to +/- one month to reduce potential "noise" in our data due to seasonal variances in diarrhea prevalence. We also collect data from the health facilities in the project area regarding diarrhea cases in children under five before and during the project (collected by completed calendar years, so if the project ran June 2021 – May 2024, health facility data was collected for calendar years 2021, 2022, and 2023) to triangulate our results.

We have observed overall reductions in child diarrhea between the baseline and endline studies, from both the household survey data we collect and data from the institutions in our project areas. The percentage of households with children under 5 reporting at least one case of diarrhea in the week prior to the survey in this

age group has decreased by a median of 90% from baseline to endline across all our evaluation studies to date. Our projects are often the only WASH promotion programs to our knowledge implemented in these areas between baseline and endline. This suggests that our WASH program has contributed to the observed reduction of diarrhea in our project areas, although, as noted above, we do not believe that our research methods are sufficient to suggest direct causality to attribute this change to WFG.

Q5. How do you ascertain handwashing behavior?

In the evaluation studies, handwashing behavior is assessed by asking respondents: 1) Handwashing behavior in the previous 24 hours (materials, when), 2) Observing household handwashing facilities (materials, location, type), and 3) asking the following questions:

a. [READ Script] "Imagine that you have just finished feeding animals. Now you are hungry. Please describe exactly what you do from the moment you leave the animals until you eat." [FOR ENUMERATOR ONLY] Does the respondent mention washing his/her hands? *[Enumerator note] Do not probe respondent about handwashing.
b. [READ Script] "Imagine that you have just used the toilet. Now you are about to go to the market. Please describe exactly what you do from the moment you leave the toilet until you walk to the market." [FOR ENUMERATOR ONLY] Does the respondent mention washing his/her hands? *[Enumerator note] Do not probe respondent about handwashing.

Results from questions (a) and (b) are compared to the self-report handwashing questions and the reported facilities available on the compound. As expected, we often find that the number of people who report having washed their hands in the past 24 hours is often much higher than the percentage who have handwashing facilities at baseline, suggesting social desirability bias.

Q6. Generally, what survey instrument or questionnaire do you use?

Water for Good uses an internally developed household knowledge, attitudes, and practices (KAP) survey which includes questions on demographics, health and wellbeing, community assets and institutions, water access, water management and use, sanitation, hygiene, menstrual hygiene, and waste management. This survey was developed at the beginning of the VHV model to understand household WASH access, beliefs and practices and most of these questions are based on existing tools from leaders in the sector.

We use the mobile data collection platform of mWater which enumerators can use on their phones or tablets to conduct the survey. Every survey is reviewed daily to 1) ensure high quality data and 2) resolve any errors/discrepancies in survey data as close to the time of data collection as possible.

Project Intervention

Q7. What exactly does the intervention consist of, and over what duration does it run?

The Vision of a Healthy Village is implemented over three years, and results are monitored for five years. More recent evaluation reports contain this description of the program:

To meet the goal of "reducing WASH-related diseases and improving the health and wellbeing of children and families in the [] district through safe WASH", the project will focus on changing WASH behaviors as well as constructing WASH infrastructure (water points, ventilated improved pit latrines, and handwashing stations). Every household in the project area will be targeted with behavior change activities and set clear, attainable goals such as a Healthy Home certification[4]. Schools will be mobilized to teach their students about WASH (through teacher trainings about issues such as WASH and menstrual hygiene management), cultivate a WASH conducive environment, and implement an operation and maintenance plan for WASH infrastructure. Once communities and schools provide initial investments and establish a WASH committee, Lifewater will begin construction of improved water sources, school ventilated improved pit latrines (VIPLs), and school handwashing facilities. In addition, health workers at local health care facilities will be equipped with training on preventing WASH-related diseases.

[4] Healthy Home Certification: Healthy Home is a Water for Good certification meaning that a household is drinking safe water (by treatment or using an improved source), storing water safely, has an improved latrine with dignity (including a pit cover, door, walls, and a roof), has handwashing devices with soap/ash and water near the latrine and near the kitchen, has a drying rack, has a bathing shelter that allows water to drain, and has a compound clean of feces and rubbish.

Post-Intervention Surveys

Q8. What is the duration of the post-project survey?

Similarly to baseline and endline data collection, our goal is typically to collect the data within a two-week time window. Sampling is done similarly to baseline and endline, although as the household census data is five years old, we also apply the national population growth to the calculations.

Q9. What time of year does the post-project survey take place (in relation to baseline timing)?

The post-project evaluation (PPE) survey data is collected two years after the endline data is collected, +/- one month max. So, from the time of baseline, the post-project survey is collected 5 years later, +/- one month of the baseline data collection. We try to conduct the baseline, endline and post-project evaluations in the same month to avoid any seasonal variance in the data.

Q10. Is the post-project survey planned with the community or done unannounced?

Communities are aware that there will be future evaluations after the endline study, and the program teams maintain contact with these communities during the time between studies through regular monitoring activities. We complete preparation and planning of surveys with countries and program teams and, at the time of the surveys, enumerators/data collectors get participants' consent before conducting the surveys.

Q11. What form of statistical analysis is conducted?

Descriptive statistics – the percentage of respondents answering each answer choice for categorical data, average/median/max/min for continuous data.

The percentage of respondents meeting the criteria for each key indicator are analyzed in Excel or directly in mWater. Then:

a. For key indicators expressed as a percentage, we conduct a two-sample proportion hypothesis z-test or a Fisher's exact test.

b. Key indicators that are not proportions/percentages (e.g., water fetching trip time in the dry season and medical expenditures in four weeks prior to survey), conduct either a Mann-Whitney U test or a two-sample t-test.

Analysis is done using several different software platforms. Some calculations are set up and completed directly in mWater, our data collection tool. The bulk of the analysis is completed in Excel or using Excel add-ons such as XLSTAT.

Qualitative data from focus groups is analyzed using the Atlas.ti program.

Q12. Generally, what survey instrument or questionnaire do you use?

We collect this data in the same manner as baseline and endline, although the household survey is often shortened to only include key data and indicators.

a. Household KAP Survey
b. School Enrollment and Dropout data
c. Healthcare Facility Case Data
d. Focus Group Discussions
e. Key Informant Interviews

Community Characteristics

Q13. What does it mean that "82% of respondents felt that their wealth had increased"? How is wealth defined and measured?

For this figure, we asked the respondents directly, "Since this time last year (12 months ago), how has your wealth changed?" [Increased since last year/Stayed the Same/Decreased since last year] to understand the respondent's perceptions of changes to their overall wealth. We also collect other information regarding the respondents' socio-economic status that are not included in this figure, including months with shortages of income and food, education level, electricity access, and cell phone ownership.

Q14. Where do community members get water in the rainy seasons?

If improved safe water sources are found in their villages, they prefer getting water from there, and at Endline and Post-Project evaluations most respondents report using the same water source in both seasons. Otherwise in some cases, such as in Ethiopia, many households travel to nearby unprotected springs or collect rainwater.

Post-project follow-up

Q15. What if any follow-up is done with communities to encourage further uptake of hygienic practices and to sustain changes that have already taken place?

We conduct follow-up after the endline, during the sustainability phase which continues for five years following the completion of the project cycle. We do this by establishing and training community-based organizations (CBOs) and support professional operations and maintenance services to both continue promoting WASH and ensure the reliability of water supply. Additionally, water committee members, religious leaders, local government leaders and other stakeholders continue to encourage the community to sustain changes. WFG staff continue to maintain contact with the communities and collect sustainability monitoring data within communities, at institutions, and at WASH infrastructure sites. Internally, this monitoring data is compiled and analyzed annually, and then a reflection meeting is held with the program team to discuss findings, identify trends, and develop recommendations and improvements."

Further information direct from Water for Good can be obtained from Seda Arzumanyan, Chief Development Officer sarzumanyan@waterforgood.org

Reference

Carter, R.C. (2017) Can and should sanitation and hygiene programmes be expected to achieve health impacts? Waterlines 36(1):92-103, doi:10.3362/1756-3488.2017.005