Kibera: Urban slum filter implementation study

Final Report

December 22, 2022

Prepared by: Aquora Research and Consulting

Introduction to the report

In fall 2021, Sawyer Products, Inc. [Sawyer] contracted with Aquora Research and Consulting [Aquora] to partner with The Bucket Ministry [TBM; NGO partner] to provide comprehensive study design, survey design, implementation, analysis, water quality testing and dissemination to assess two distinct research questions:

Research question #1. Does the Sawyer PointOne filter [filter] get used and have beneficial health impacts in an urban environment over a 60 to 90 day implementation window?

Research question #2. Does the filter demonstrate comparative benefits over anti-parasitic medications in an urban environment over a 60 to 90 day time period?

Report structure:

The structure of this report is as follows:

Section 1. Executive summary

Project goal: The Bucket Ministry[1] is leading a campaign to install Sawyer International Bucket System®[2] water filters throughout an approximately 400,000-person slum (Kibera), located within Nairobi, Kenya. Kibera is one of the largest urban slums in Africa[3] and one of the most deprived slums globally having greater magnitude and more severity of deprivation than other regions (e.g., Latin America)[3]. Aquora Research and Consulting[4] is leading efforts to assess the efficacy of the filters on population health and water quality in a high population density environment with limited sanitation services.

Summary of results on diarrhea reduction: Aquora Research and Consulting examined a set of 10,000 households that received Sawyer Point One water filters and WASH training from November 16, 2021- April 8, 2022, in 7 neighborhoods in Kibera, Nairobi, Kenya. Households that received filters were monitored over a 90-day follow-up period. Significant data cleaning was necessary to remove data that was deemed unusable due to survey errors, interviewer heterogeneity, incomplete/missing/unusable data and other data inconsistencies. The resulting sample of 6795 households was then analyzed using advanced statistical models to account for geospatial and interviewer clustering, temporal effects, household size and other covariates. These models estimated self-reported, household diarrhea prevalence in the two weeks prior to receiving a filter to be 52.7%. After seventy days of filter use, the self-reported, household diarrhea prevalence was estimated to be 2.2%. (see figure below).

Summary of water quality testing of drinking water sources: Aquora tested 25 drinking water sources representing each of the 7 neighborhoods in Kibera surveyed and the diversity of water source types. Each drinking water source was assessed for in-field *E. coli* and Total Coliform concentrations, potential bacterial pathogens based on community sequencing, and concentrations of 8 metals with known long-term health effects. Results of in-field testing showed that 6 water sources were in the WHO risk category of Unsafe for *E. coli*: 13 water sources were in the Low Risk/Safe category for *E. coli*. In contrast, 18 water sources were categorized as Unsafe for Total Coliforms with only 2 water sources being characterized as Low Risk/Safe. Potential bacterial pathogens were identified in all water sources tested by community sequencing. These results lend support to the high levels of baseline diarrhea rates observed. Metals testing of drinking water sources showed that all but one source had either undetectable levels or detectable levels below WHO health guidelines. One drinking water source (Source 374) showed arsenic levels above WHO health guidelines. These results support the use of Sawyer filters in the Kibera slum.

Summary on study findings from anti-parasitic medications: A low-dose, single administration of Albendazole was provided to all households at distribution as part of the standard protocol used by TBM. As part of the overall study design to isolate potential health effects of the Sawyer filter and the administration of Albendazole, a subset of 2642 households did not receive Albendazole at distribution, instead receiving it at the second follow-up (approximately 5 weeks after filter distribution). After data cleaning and statistical modeling, estimated differences in self-reported diarrhea prevalence between households receiving Albendazole at distribution vs. those that didn't was negligible and not statistically significant ($p>0.05$).

---

[1] https://thebucketministry.org/kibera/
[2] https://www.sawyer.com/products/international-bucket-system
[3] https://en.wikipedia.org/wiki/Kibera; State of the World's cities (2006/2007), United Nations Human Settlements/UN Habitat https://issuu.com/unhabitat/docs/2101_alt/24.
[4] http://www.aquorarc.com

Primary findings

1. The introduction of Sawyer filters to households in this urban slum with high population density is associated with reduced diarrhea and other health related problems significantly over the course of the study.
2. The introduction of Sawyer filters to households in this urban slum is not associated with a statistically significant reduction in monthly costs for obtaining water; however, a significant reduction in monthly costs for health care expenses was observed.
3. Representative water quality testing indicates a high frequency of drinking water source contamination with *E. coli* and Total Coliforms. Potential waterborne pathogens were detected in all water sources tested. Filtration of drinking water with Sawyer filters can reduce the risk from this contamination.
4. Representative water quality testing indicates a very low frequency of dissolved metals present above WHO guidelines for drinking water. Use of Sawyer filters may be sufficient to remove known contaminants of concern.
5. Anti-parasitic medication distribution had little to no impact on the results.

Key limitations

1. This study is observational. Thus, the resulting associations are not necessarily causal. However, given (a) the prospective nature of the study and (b) the control of numerous covariates via the statistical modelling process, the resulting associations are strongly suggestive.
2. This study is based on self-reports of health, economic and other outcomes. Caution should be used if attempting to utilize the data as true measures of absolute levels of disease or economic benefit.
3. The need to eliminate approximately 30% of the data as part of data cleaning is substantial and we are assuming that missing data is approximately distributed at random.
4. Water quality testing of chemical contaminants was limited to dissolved metals. Other types of chemical contaminants found in urban environments were not tested (e.g., volatile organic chemicals (VOCs) associated with petroleum and other organics).

Section 2. Additional background and significance

The Bucket Ministry[5], is leading a campaign to install Sawyer International Bucket System® [6] water filters throughout an approximately 400,000-person slum (Kibera), located within Nairobi, Kenya. Kibera is one of the largest urban slums in Africa[7] and one of the most deprived slums globally having greater magnitude and more severity of deprivation than other regions (e.g., Latin America)[8]. Aquora Research and Consulting[9] is leading efforts to assess the short-term efficacy of the filters on population health and water quality in a high population density environment with limited sanitation services. To date published, peer-reviewed papers investigating the efficacy of the Sawyer International Bucket System have focused on rural contexts[10]. This project is the first of its kind to investigate the efficacy of the Sawyer International Bucket system in an urban slum. Proof-of-principle in this unique context is critical to establish use of the system in urban slums globally.

[5] https://thebucketministry.org/kibera/
[6] https://www.sawyer.com/products/international-bucket-system
[7] https://en.wikipedia.org/wiki/Kibera; State of the World's cities (2006/2007), United Nations Human Settlements/UN Habitat https://issuu.com/unhabitat/docs/2101_alt/24.
[8] https://en.wikipedia.org/wiki/Kibera; State of the World's cities (2006/2007), United Nations Human Settlements/UN Habitat https://issuu.com/unhabitat/docs/2101_alt/24.
[9] http://www.aquorarc.com
[10] https://pubmed.ncbi.nlm.nih.gov/33397511/ and https://pubmed.ncbi.nlm.nih.gov/31410085/

Section 3. Methods

Data collection: Data was collected as part of a larger campaign to install Sawyer International Bucket System®[11] water filters throughout an approximately 400,000-person slum (Kibera), located within Nairobi, Kenya. Interviewers were trained by TBM in conjunction with Aquora Research and Consulting on data collection procedures. Aquora team members participated in onsite training in Kibera in November 2021, with a follow-up visit in March 2022 to ensure compliance with field protocols. Visits included embedding with randomly selected interviewers (as well as with interviewers identified with aberrant survey data) to improve compliance. TBM began by canvasing neighborhoods and inviting community members to a distribution event, typically 470 households (range: 60-1050). At distribution events, all attendees are provided a filter for their home (one per household), baseline survey data is collected, group training on the use of filters is conducted, individual demonstration of backflushing is required, and a basic lesson on the importance of hand washing is provided. Data collectors then visited each home with scheduled visits at 7-10 days, 30-45 days and 60-90 days after distribution. To build relational capital, the same data collector was scheduled to conduct all three follow-up visits with each household. Over 100 data collectors were active in the field throughout the course of the study. Data collection procedures were approved by the Institutional Review Board at Dordt University (a federally compliant IRB: FWA00028884), by contract with Aquora.

Survey All household residents 18 years of age and older were eligible to participate and informed consent was received prior to participation. Receiving a filter was not contingent upon participating in the survey. A single household resident answered questions on behalf of the entire household. The surveys (baseline and follow-up) were developed by Aquora Research and Consulting to maximize scientific rigor (valid and reliable measures), simplicity and minimal time in the field. Multiple rounds of input from various stakeholders yielded preliminary surveys which were then forward and back-translated into Swahili/English using two separate leaders of the field data collection teams. Back-translation yielded minor translation inconsistencies which were then addressed with the translators yielding two reliably translated surveys. Initial training on the surveys and pilot testing in mid-November 2021 identified additional minor survey implementation questions which were addressed in the field before roll-out of the final surveys in late November 2021.

Survey contents and measures Surveys were developed in order to address the following areas: basic household demographics (e.g., gender, age, number of household members, job status, school status), where water for use in the household is typically obtained and how much money is typically spent on water, self-reported diarrhea (two-weeks), self-reported impact of diarrhea on school/work participation (two weeks) and medical costs (previous month), self-reported other health issues (last two weeks) and self-reported hand washing. Follow-up surveys also included measures of filter utilization (how often refill bucket), interviewer perceptions of filter utilization/functionality/user ability to clean filter and self-reported use of hand washing before/after different activities. At follow-up time points, in case the interviewer perceived a lack of knowledge or understanding of respondent filter utilization and/or hand washing, the interviewer retrained/re-educated the respondent. In addition to questions on water/filter use, demographics and health/household economics, a separate set of questions related to evangelism/discipleship were also asked of respondents after all other questions, but before the interviewer left the home.

Sample The target sample was 10,000 households with data collection beginning in late November, 2021 after in-field training and interviewer calibration completed. Data collection ended July 3, 2022 after the last of the initial 10,000 households completed their third follow-up. Data was collected in 7 neighborhoods in

[11] https://www.sawyer.com/products/international-bucket-system

Kibera. By design this report only covers the initial 10,000 households, even though data collection is continuing across the remaining ~90,000 households as part of the Kibera-wide implementation project being run by TBM.

Water testing Water sources were collected in the seven neighborhoods of Kibera that were included in distribution and follow-up surveys as part of this study. Water sources were selected to represent typical drinking water sources that are used in Kibera on a daily basis. In addition, ad hoc testing of the river that runs through Kibera and other sources that could be used as drinking water sources was conducted. To our knowledge, these are not currently used as drinking water sources. For each source tested, three, ~20L samples were collected in jerry cans. Each jerry can was rinsed with the source water prior to filling the can. Jerry cans were transported to centralized testing locations near collection points. Tandem filter kits were used to test each water source. The kits included 1) one Aquagenx CBT EC + TC (Compartment Bag Test) Most Probable Number (MPN) kit[12] for in-field *E. coli* and Total Coliform testing, 2) three tandem filter assemblies for collection of full bacterial communities for 16S rRNA sequencing, and 3) three metal testing kits for collection of water for ICP-MS analysis at a certified testing laboratory in the United States. The tandem filter kits are designed to test three, independent replicates of each water source (with the exception of the in-field EC + TC test). Tests were conducted by Aquora personnel or local TBM team members trained in the field by Aquora personnel according to protocols developed specifically for this testing approach. Information about each water source and testing procedure, including EC + TC testing results, were recorded using FastField Forms or Survey123 mobile device apps in real time. All testing forms were developed by Aquora. Tandem filter kits used for water sources were transported back to the United States either in airline luggage or shipped via FedEx. Upon receipt in the laboratory, tandem filter assemblies were stored at -80°C until processing; metals testing kits were acidified and submitted to ALS Environmental for testing of Arsenic, Barium, Cadmium, Chromium, Copper, Lead, Nickel and Selenium. Metals results were compared to WHO guidelines for metal concentration limits in drinking water. EC + TC test results were analyzed to verify accuracy in-field test interpretation by comparison of recorded results to photographs of the test kits. These validated results were translated to MPN concentrations (cells/100 mL) and assigned to WHO Health Risk Categories based on MPN. Tandem filter assemblies were processed to recover bacteria filtered from source water by individually backflushing A filters (capturing source water bacteria) and B filters (control for each replicate). Total genomic DNA was extracted from backflushed material for downstream sequencing of 16S rRNA bacterial genes. Library preparation and QA/QC were performed according to the Schloss Lab MiSeq Wet Lab SOP[13] and sequenced using Illumina MiSeq 2x250 bp v2 paired end sequencing according to manufacturer instructions. Data processing of 16S rRNA sequenced libraries was conducted on a high-performance computing cluster using the Schloss Lab MiSeq SOP mothur pipeline[14]. Data analyses were conducted using custom R scripts in R Studio to assess bacterial community profiles and determine presence/absence of 18 bacterial genera known to contain waterborne pathogen species. In addition, custom R scripts were implemented in R Studio to summarize EC + TC data, metals data, and water source metadata. Individual water quality dashboards were produced for each water source, summarizing the three data types and general water source descriptions captured through in-field surveys.

[12] https://www.aquagenx.com/wp-content/uploads/2022/06/MPN-CBT-ECTC-Instructions-DrinkingWater-20220513.pdf
[13] https://github.com/SchlossLab/MiSeq_WetLab_SOP/blob/master/MiSeq_WetLab_SOP.md
[14] https://mothur.org/wiki/miseq_sop/

Data cleaning Despite numerous internal validation checks built into the survey instrument, and automatic scheduling of follow-ups, substantial data cleaning was necessary to yield a high-quality dataset ready for analysis. First, early on in the project, substantial data heterogeneity was observed by interviewer. Notably, response patterns by interviewer suggested that interviewers were not consistently implementing the field protocols. Aquora coordinated with TBM and field managers in Kibera to generate lists of interviewers for which the data suggested were not consistently implementing field protocols. These interviewers were then observed in the field and received retraining. This process continued for the remainder of the project discussed here. Additional details on surveyor heterogeneity are provided in the Appendix. Second, early on in the project, surveyor's data suggested that data collection tablets and/or login IDs were often shared between interviewers, which made tracking and identifying uncalibrated surveyors challenging. Again, retraining occurred to attempt to end this practice to improve data quality and consistent implementation. To address both of these data limitations, data collectors who generated health outcomes in the top 5% or bottom 5% prevalence (e.g., very high or very low levels of reported health outcomes) were eliminated from the dataset for the purpose of this analysis.

Third, despite a scheduling system, there were substantial differences in the time at which surveys took place as compared to the ideal timing. Ideally, follow-up surveys took place 7-10 days (followup #1), 30-45 days (followup #2) or 60-90 days (followup #3), however, there were substantial deviations from this ideal distribution (se Appendix for details). To address this issue and ensure meaningful interpretation of results, we emphasize *days since baseline* instead of follow-up #, for the majority of the graphs/statistical models utilized in this report.

Fourth, at the beginning of the data collection process we observed that questions on the survey follow-up forms were being pre-propagated with answers from the distribution form. After identifying the issue, Aquora immediately initiated efforts to correct the error. Efforts were subsequently taken to rectify data propagation issues.

Finally, there were numerous other observed data inconsistencies, including (a) widely varying household sizes reported over time, (b) very large household sizes reported (e.g., 30+), (c) unclear survey responses, (d) surveys with substantial amounts of missing data, (e) locations mapped more than 5 miles from Kibera, (f) duplicate filter IDs, (g) extreme values (more than 4SD) for work or school missed, (h) follow-up dates recorded before distribution dates, (i) surveys with interviewer notes suggested poor data quality and (j) other issues. Surveys with poor quality and/or high levels of missing data were eliminated from subsequent analyses.

After data cleaning, the final number of distribution surveys available for consideration was n=6,795. Households were only available for analysis in the final cleaned dataset if distribution data was available and at least 1 follow up (e.g., households for which distribution data was deemed inaccurate were dropped not only from the distribution dataset, but also all follow-up data was ignored), to ensure that households being tracked were consistent over the course of this longitudinal study. The total number of surveys available for analysis was 25,208, including 18,413 follow-up surveys covering a range of 1 to 203 days post-distribution.

Table 1. Data available for analysis after substantial data cleaning to improve data quality and validity

Time point	Distribution	Follow-up #1	Follow-up #2	Follow-up #3
Surveys available	6,795	6,255	6,153	6,005

Statistical methods

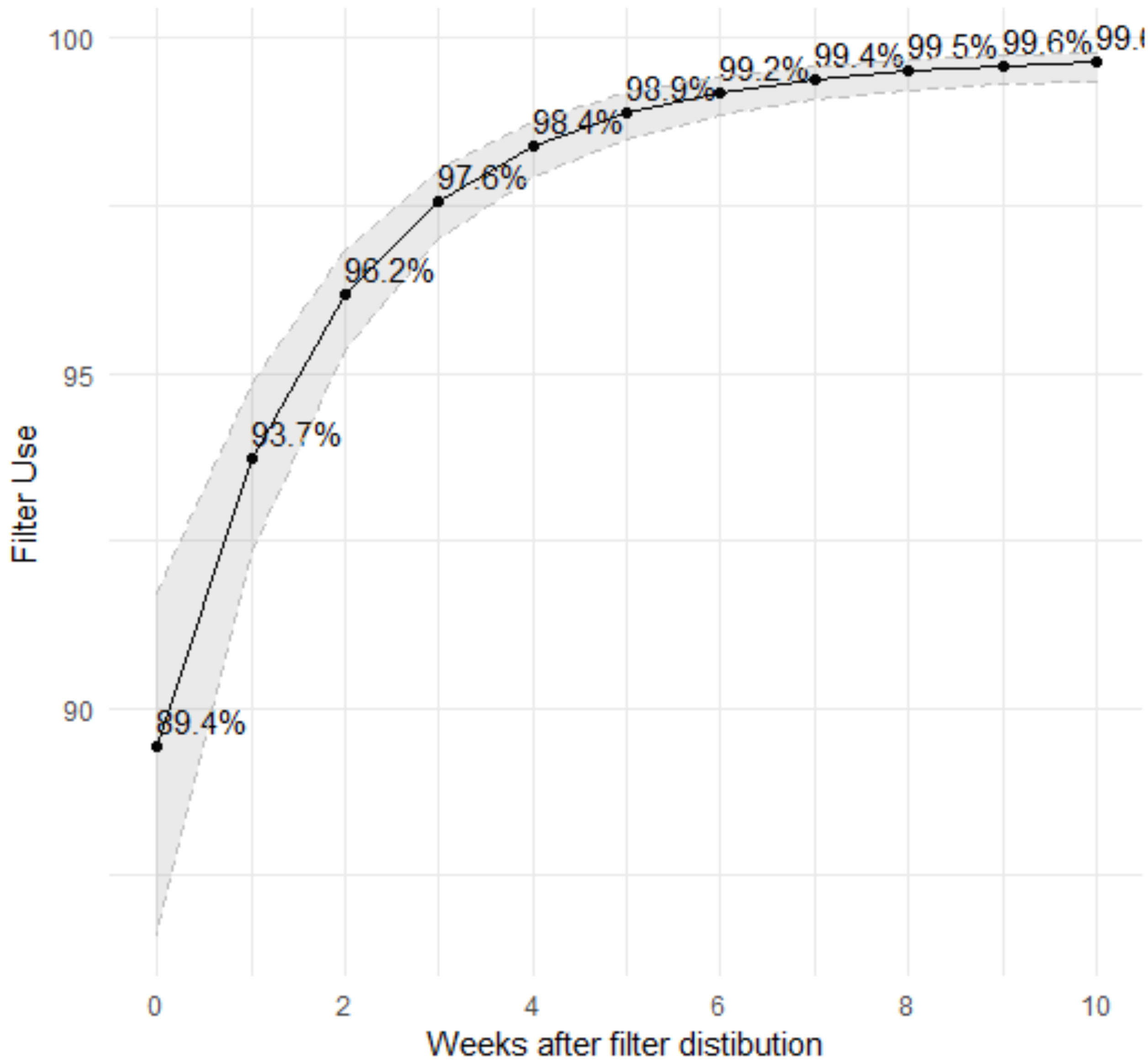
To correctly model the longitudinal nature of the data, and account for numerous sources of confounding, generalized, linear (or logistic) mixed effects models were used. Random effects were included in the model to account for the repeated measures nature of the data (random effect for household) and for interviewer (random effect term). Fixed effects (covariates) were included in the model to account for numerous sources of potential confounding: (a) household size, (b) seasonality (fixed effect for each month), (c) filter installed (yes/no), (d) neighborhood, (e) knowledge of WASH practices and (f) filter use. Separate models were for fit for each different health and economic outcome. Similar models were fit for filter utilization removing filter installation and filter use from the covariates list.

Results provided in this report focus on graphs of adjusted health, economic and utilization rates – accounting for covariates, mixed effects and after data cleaning. 95% confidence bands are included on all graphs. Results shown are statistically significant ($p<0.001$), due to the large sample size available for this study, with non-significant differences noted in the text. Covariate adjusted prevalence estimates on the graphs are obtained by selecting median levels of covariates. Alternative figures are provided which use two week-categories, to obtain (unsmoothed) estimates. Additional model details, covariate effect estimates, are not provided here but are available upon request.

Section 4. Filter Utilization

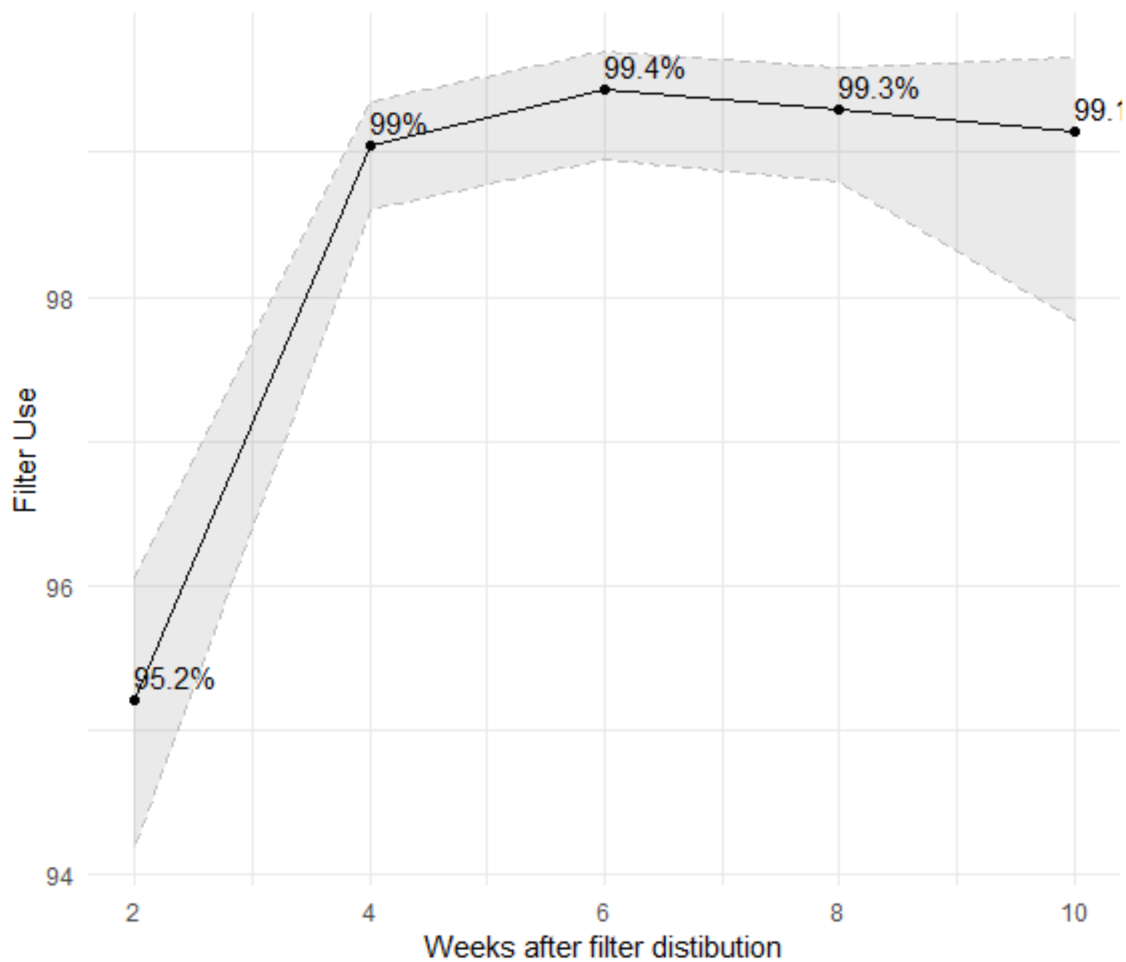
The following graphs illustrate results from statistical models (see Section 3 for details) exploring filter utilization patterns over time. Initial interpretations are provided below each figure.

Figure 4A. Model adjusted estimates of filter utilization



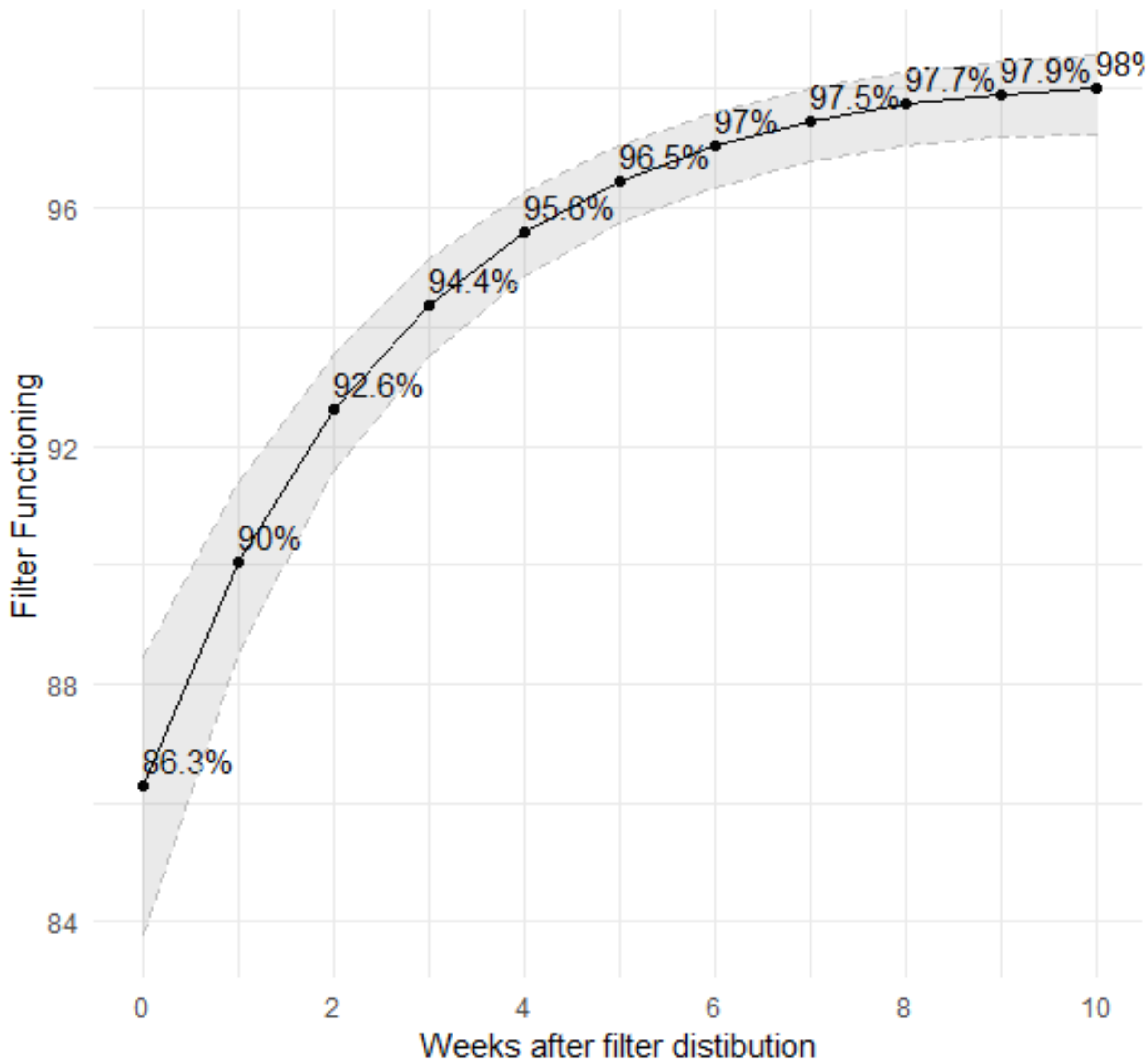
Example of interpretation: After accounting for covariates (e.g., seasonality; interviewer heterogeneity), approximately 99.2% of participants report using their filter within 6 weeks of distribution. This increasing trend likely indicates participants increasing confidence in the filter's utility during the first (and second) follow-up visits.

Figure 4B. Model adjusted estimates of filter utilization – without smoothing



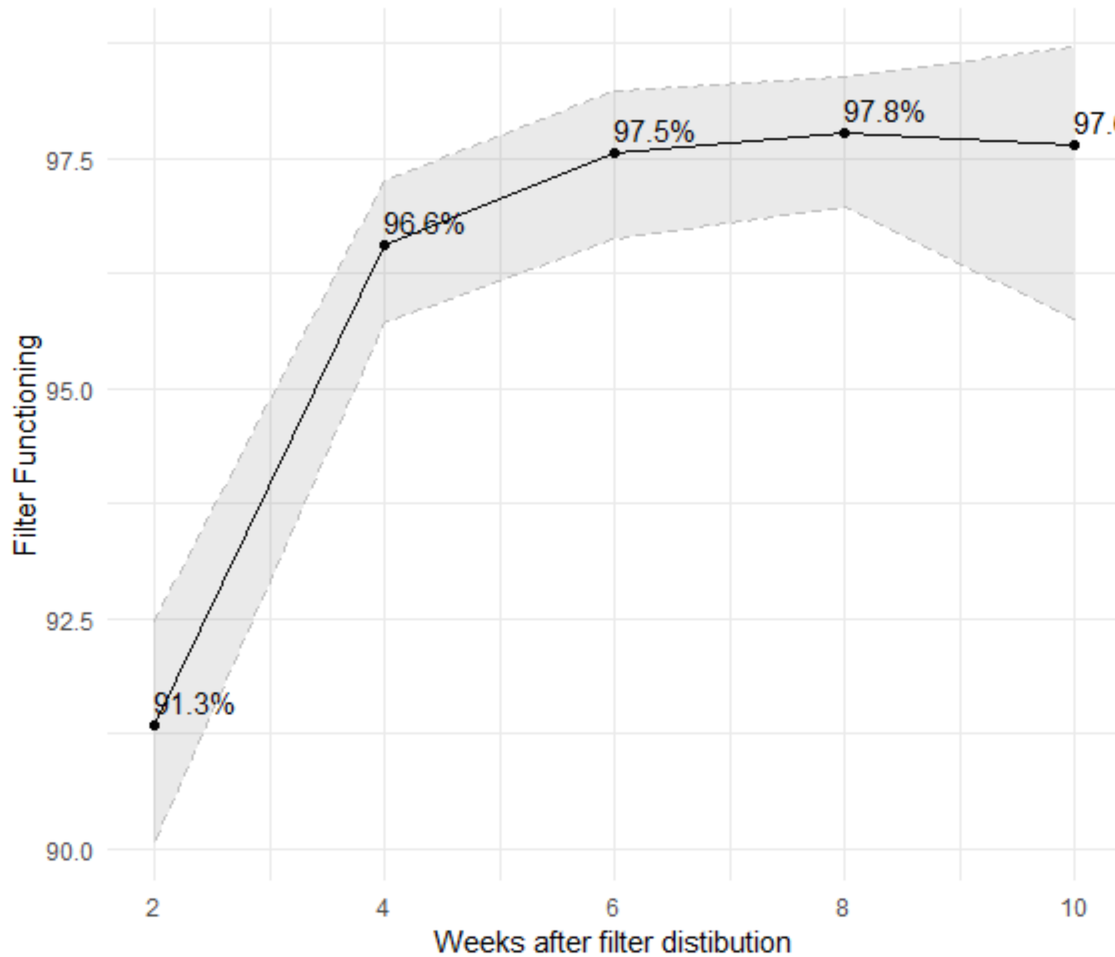
Example of interpretation: After accounting for covariates (e.g., seasonality; interviewer heterogeneity), approximately 99.4% of participants report using their filter within 6 weeks of distribution. This increasing trend likely indicates participants increasing confidence in the filter's utility during the first (and second) follow-up visits.

Figure 4C. Model-adjusted estimates of reports of filter functioning over time



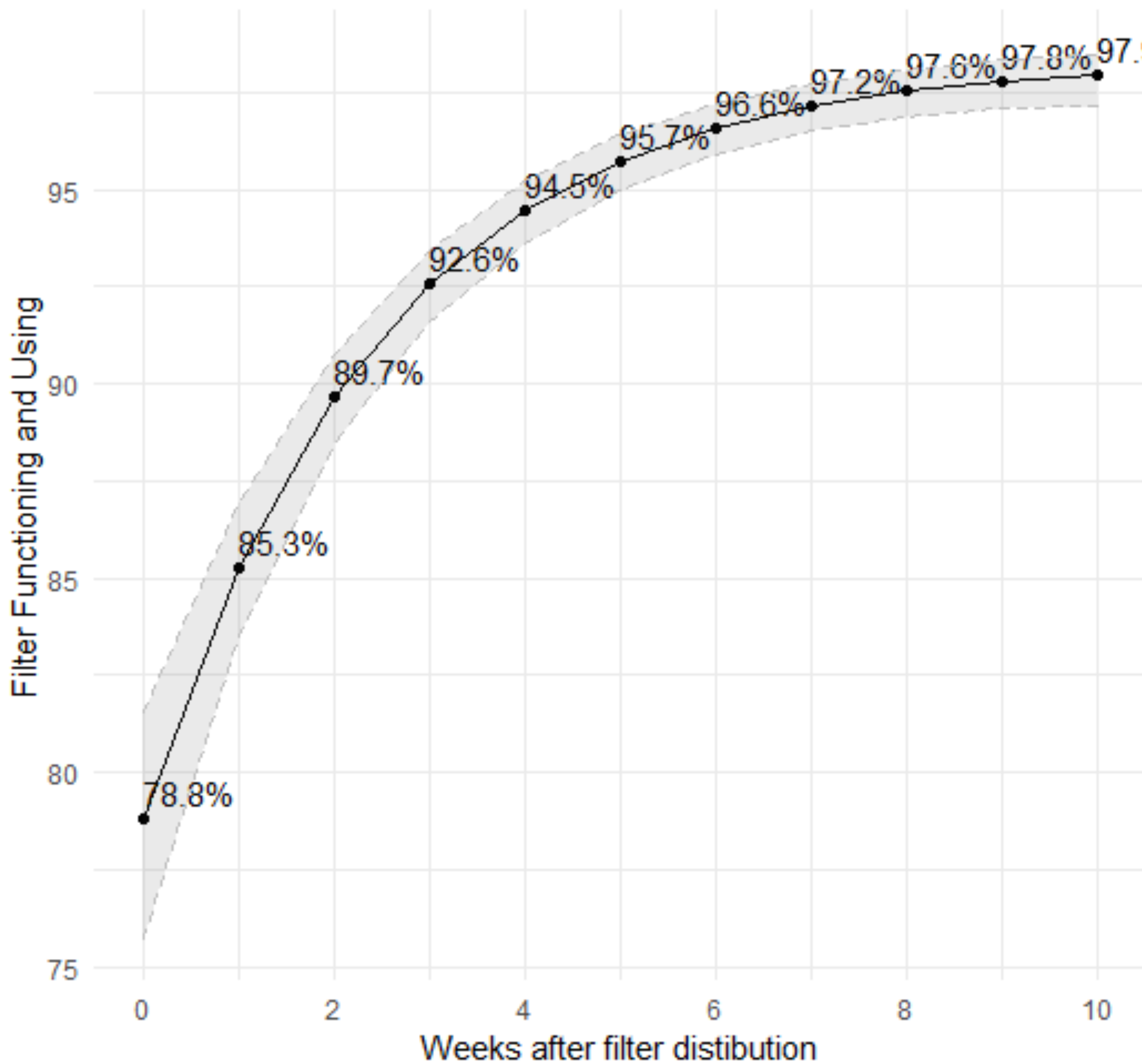
Example of interpretation: After accounting for covariates (e.g., seasonality; interviewer heterogeneity), approximately 97% of participants report having a functioning filter within 6-weeks of distribution. This increasing trend likely indicates retraining, and ensuring parts, etc. are available for participants during the first (and second) follow-up visits.

Figure 4D. Model-adjusted estimates of reports of filter functioning over time – without smoothing



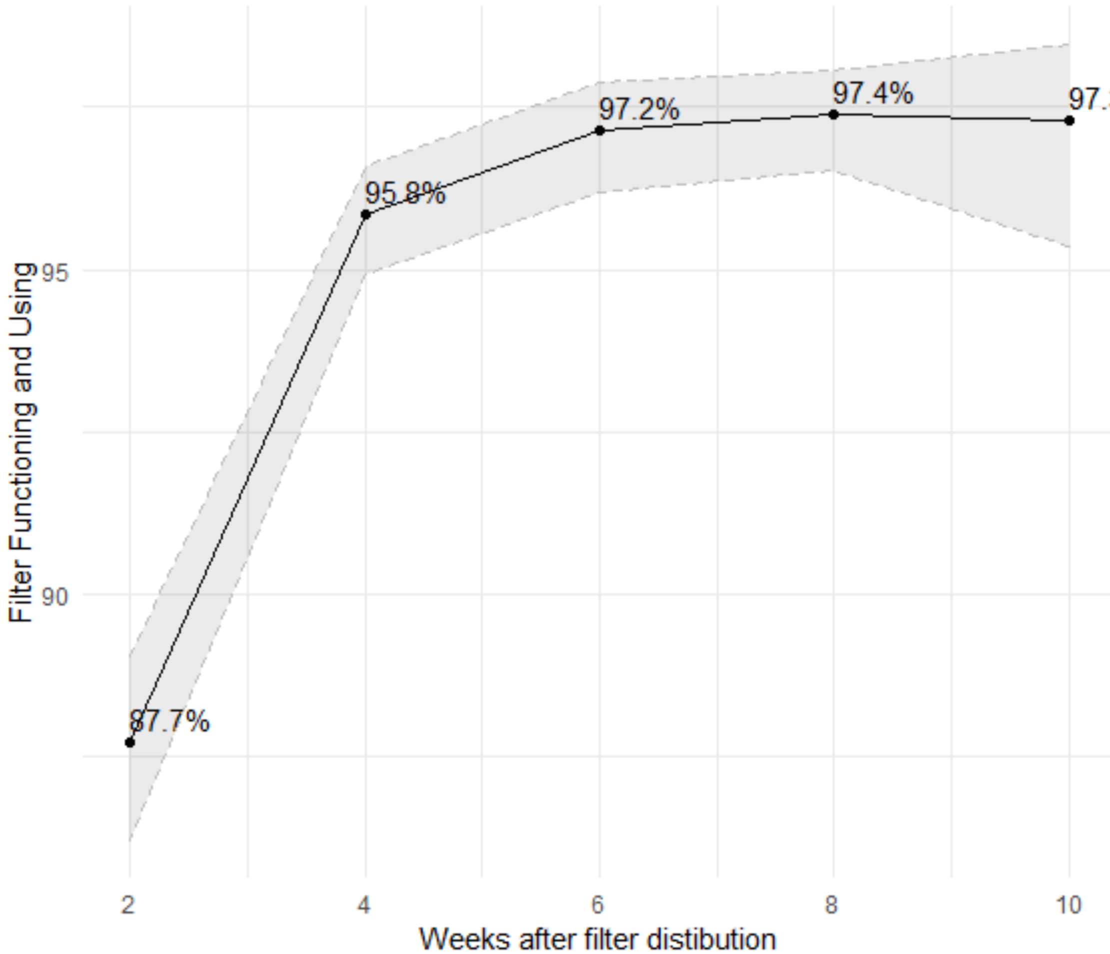
Example of interpretation: After accounting for covariates (e.g., seasonality; interviewer heterogeneity), approximately 97.5% of participants report having a functioning filter within 6-weeks of distribution. This increasing trend likely indicates retraining, and ensuring parts, etc. are available for participants during the first (and second) follow-up visits.

Figure 4E. Model adjusted estimates of simultaneous filter function and utilization



Example of interpretation: After accounting for covariates (e.g., seasonality; interviewer heterogeneity), approximately 96.6% of participants report having a functioning filter and using it within 6-weeks of distribution. This increasing trend likely indicates retraining, ensuring parts are available and building confidence in the utility of the filter for participants during the first (and second) follow-up visits.

Figure 4F. Model adjusted estimates of simultaneous filter function and utilization – without smoothing

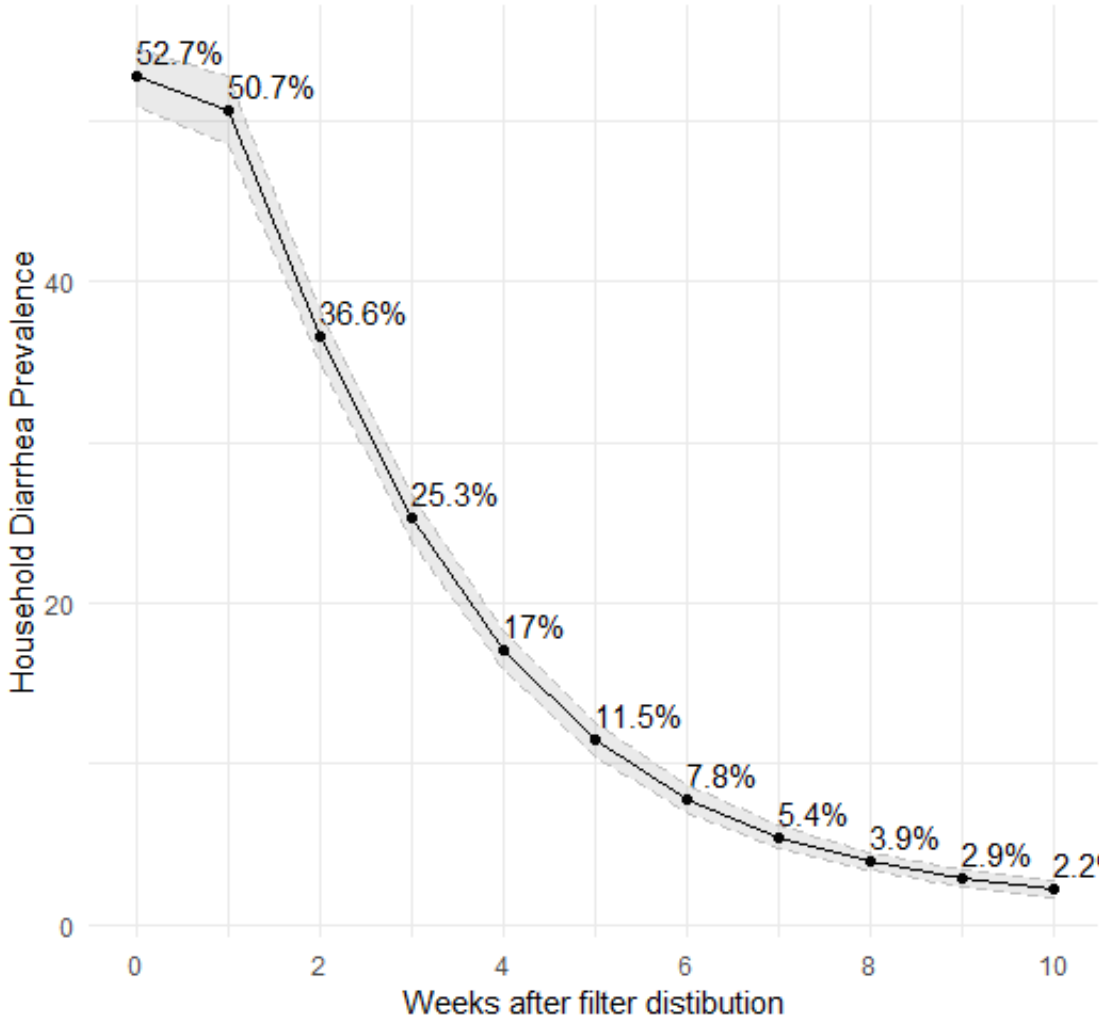


Example of interpretation: After accounting for covariates (e.g., seasonality; interviewer heterogeneity), approximately 97.2% of participants report having a functioning filter and using it within 6-weeks of distribution. This increasing trend likely indicates retraining, ensuring parts are available and building confidence in the utility of the filter for participants during the first (and second) follow-up visits.

Section 5. Health Impacts

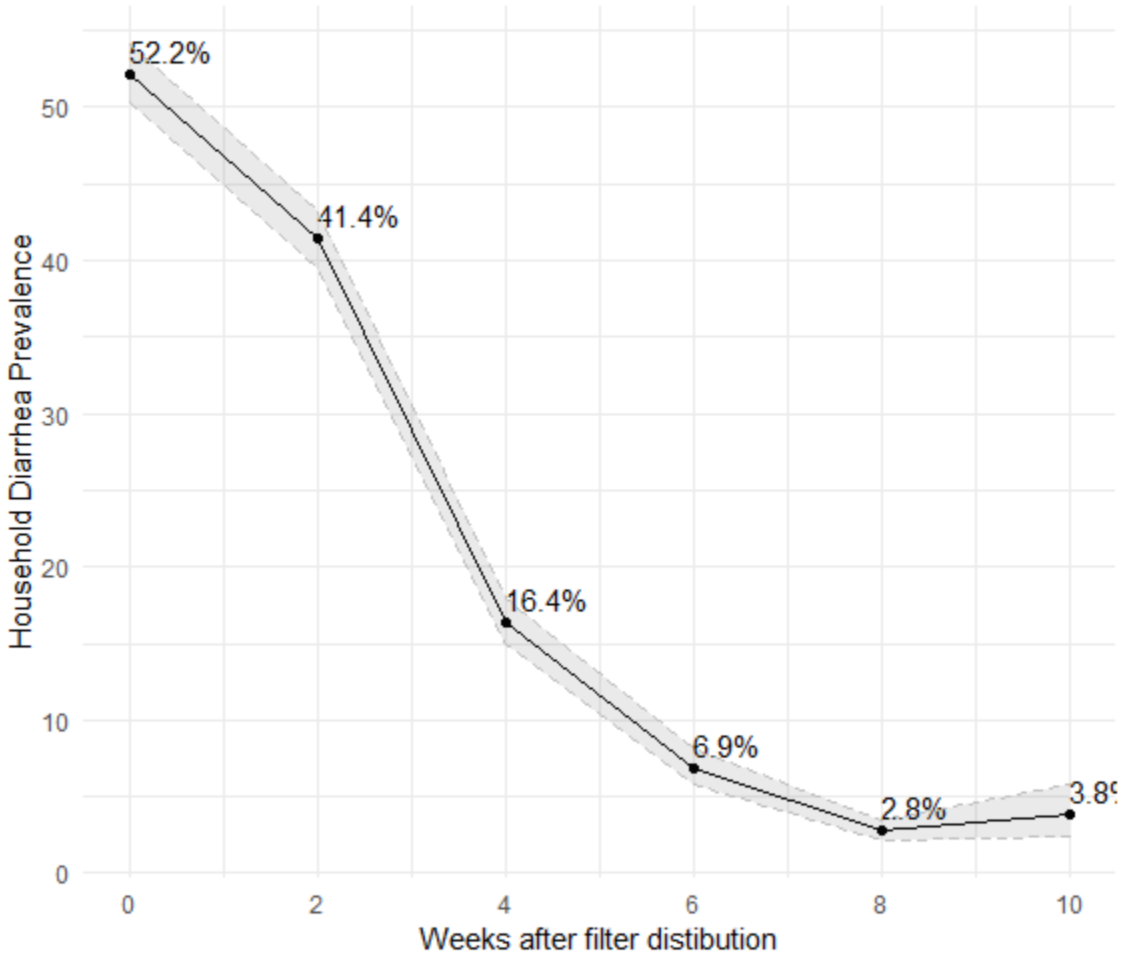
The following graphs illustrate results from statistical models (see Section 3 for details) exploring health patterns over time after filter distribution, WASH training and anti-parasitic medication distribution. Initial interpretations are provided below each figure.

Figure 5A. Model adjusted estimates of Diarrhea Prevalence



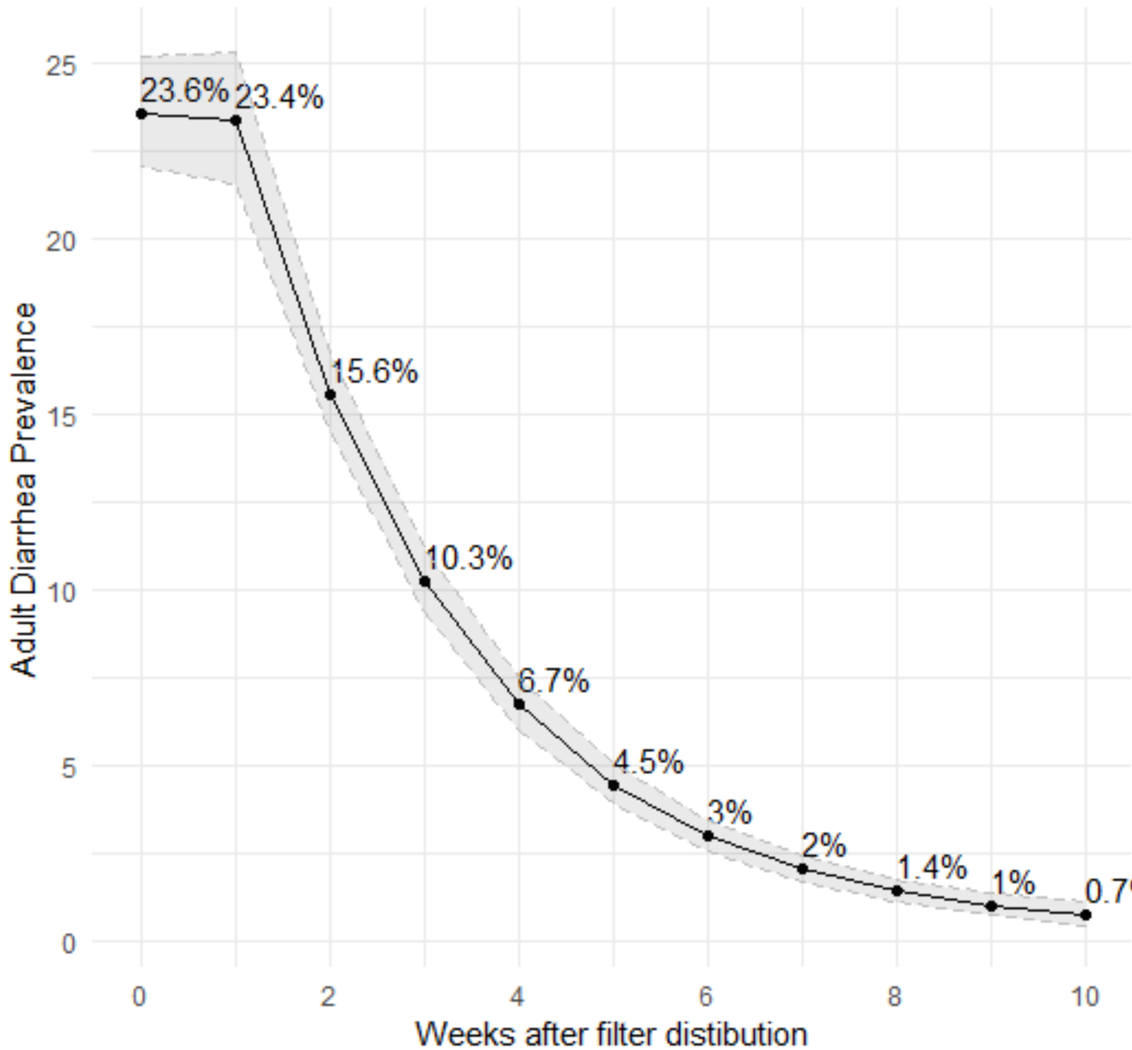
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), approximately 7.8% of participants report having had someone in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5B. Model adjusted estimates of Diarrhea Prevalence – without smoothing



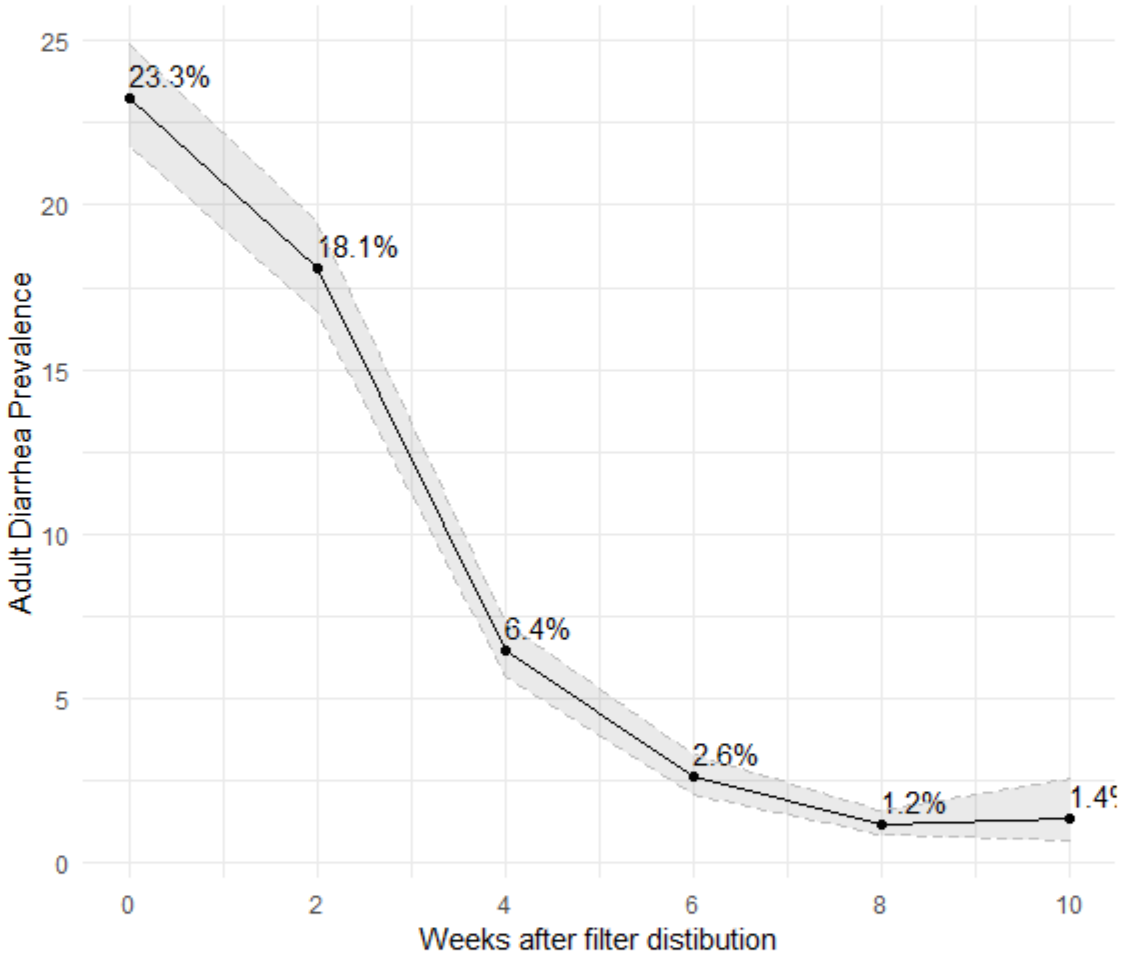
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), approximately 6.9% of participants report having had someone in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5C. Model adjusted estimates of Diarrhea Prevalence among adults (18+)



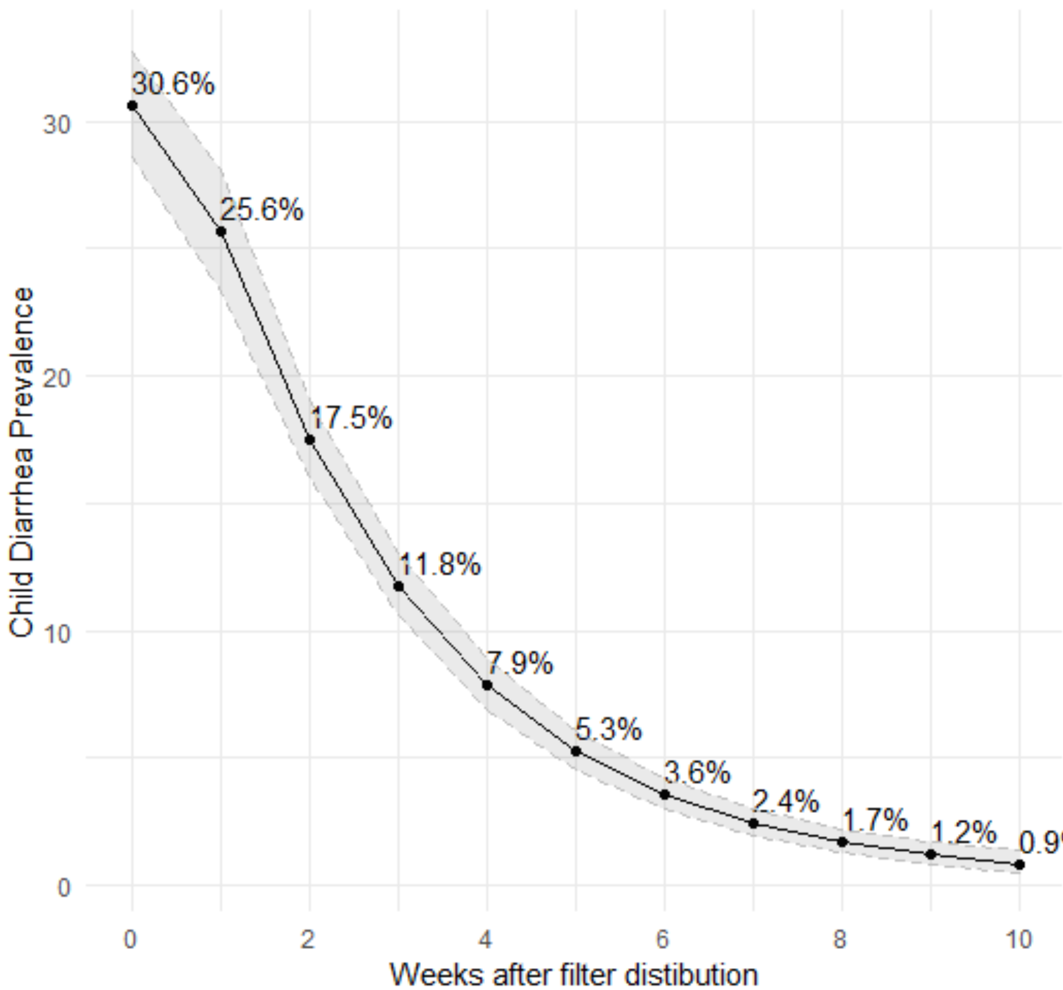
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), approximately 3% of participants report having had an adult in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5D. Model adjusted estimates of Diarrhea Prevalence among adults (18+) – without smoothing



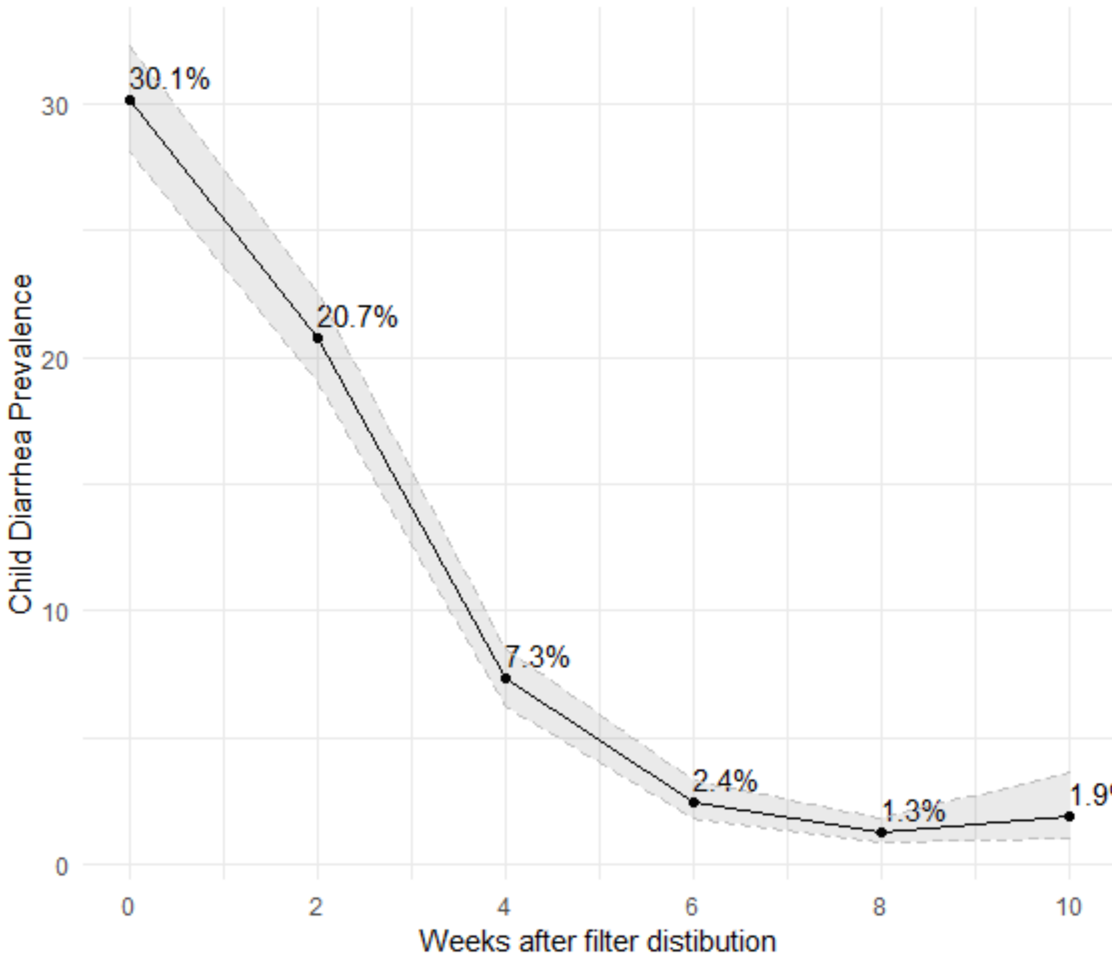
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), approximately 2.6% of participants report having had an adult in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5E. Model adjusted estimates of Diarrhea Prevalence among School aged children (ages 5-17)



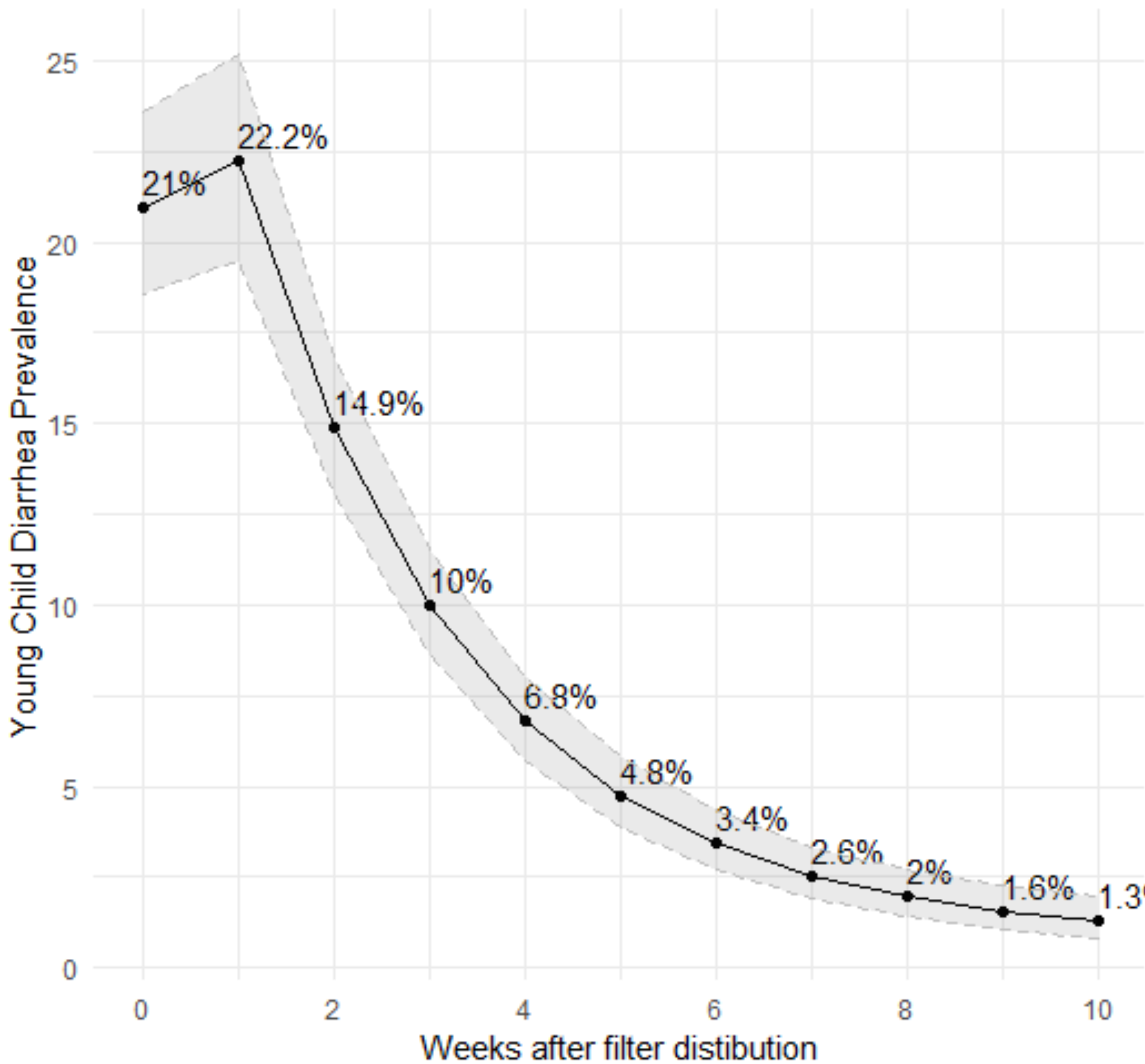
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with school-aged children, approximately 3.6% of participants report having had a school-aged child in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5F. Model adjusted estimates of Diarrhea Prevalence among School aged children (ages 5-17) – without smoothing



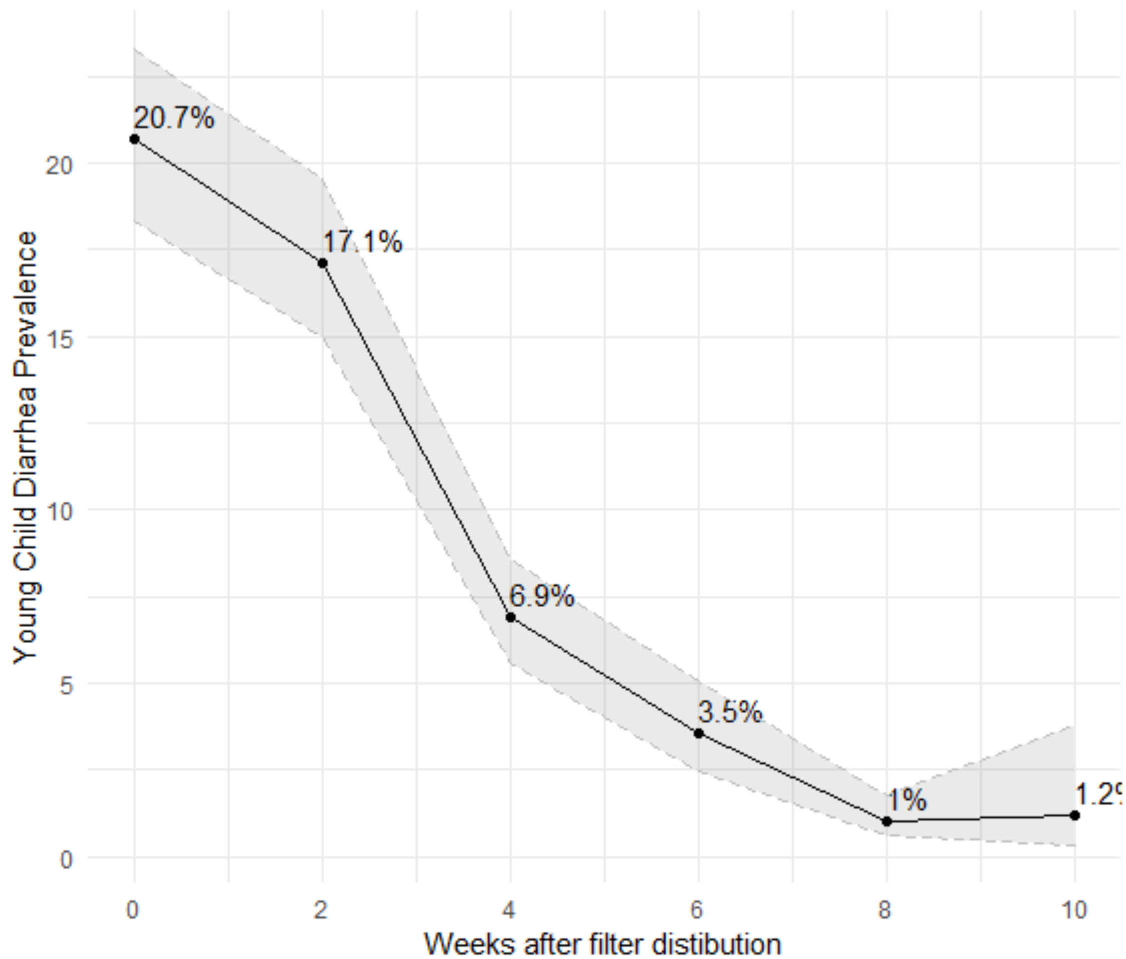
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with school-aged children, approximately 2.4% of participants report having had a school-aged child in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5G. Model adjusted estimates of Diarrhea Prevalence among young children (aged 0-4 years)



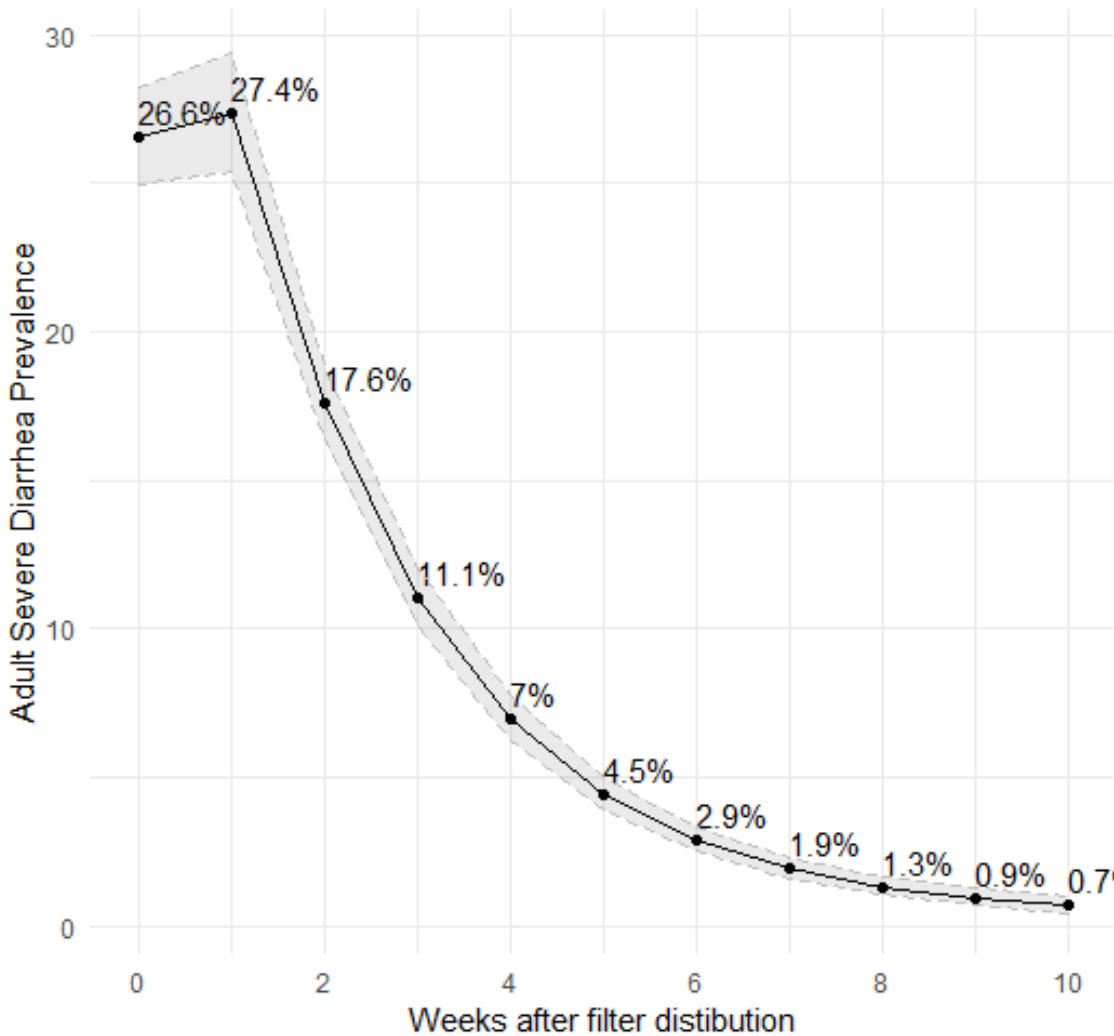
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with young children (aged 0-4 years), approximately 3.4% of participants report having had a young child in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5H. Model adjusted estimates of Diarrhea Prevalence among young children (aged 0-4 years) – without smoothing



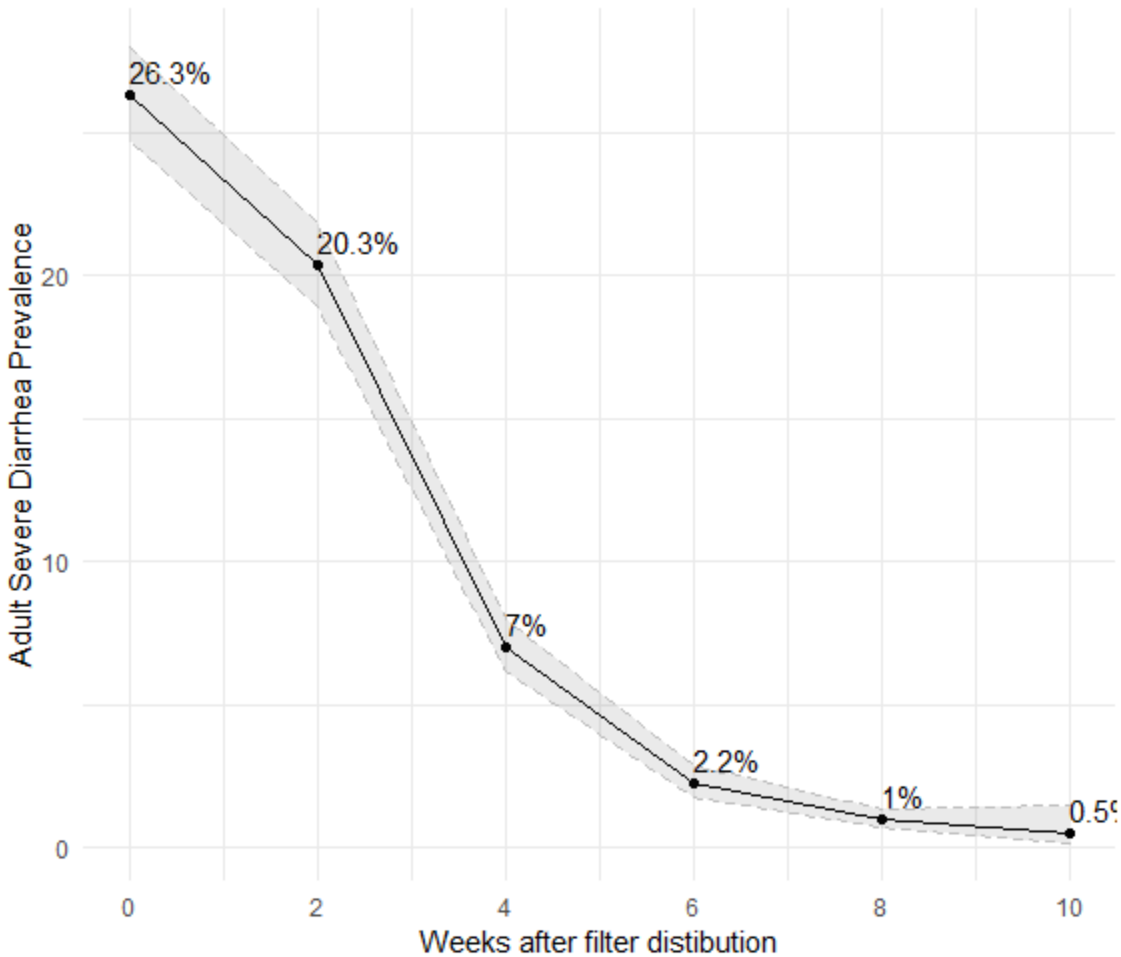
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with young children (aged 0-4 years), approximately 3.5% of participants report having had a young child in the household have diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5I. Model adjusted estimates of Severe diarrhea Prevalence among adults (aged 18+ years)



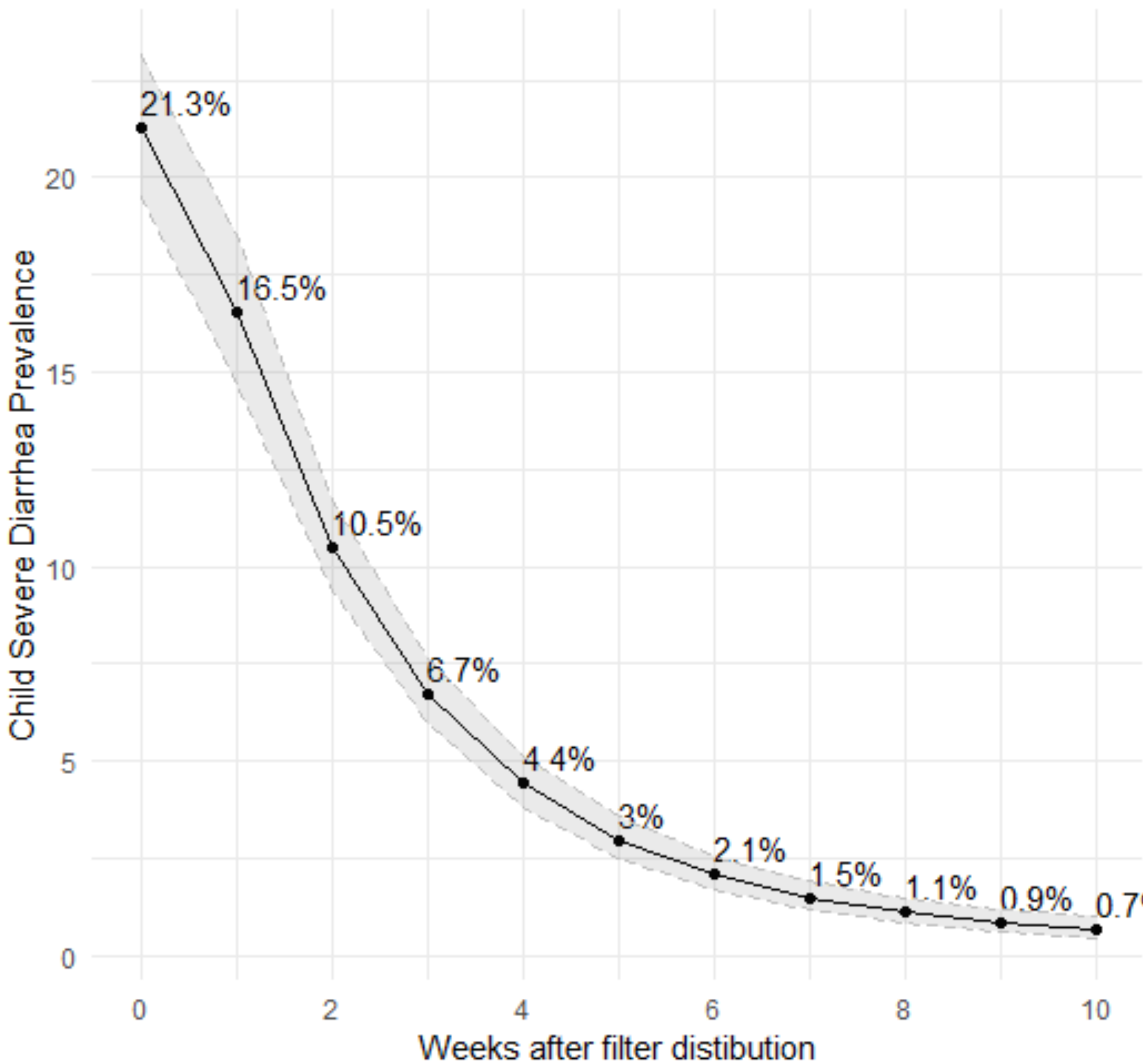
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with working adults, approximately 2.9% of participants report having had an adult in the household miss work due to having diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5J. Model adjusted estimates of Severe diarrhea Prevalence among adults (aged 18+ years) – without smoothing



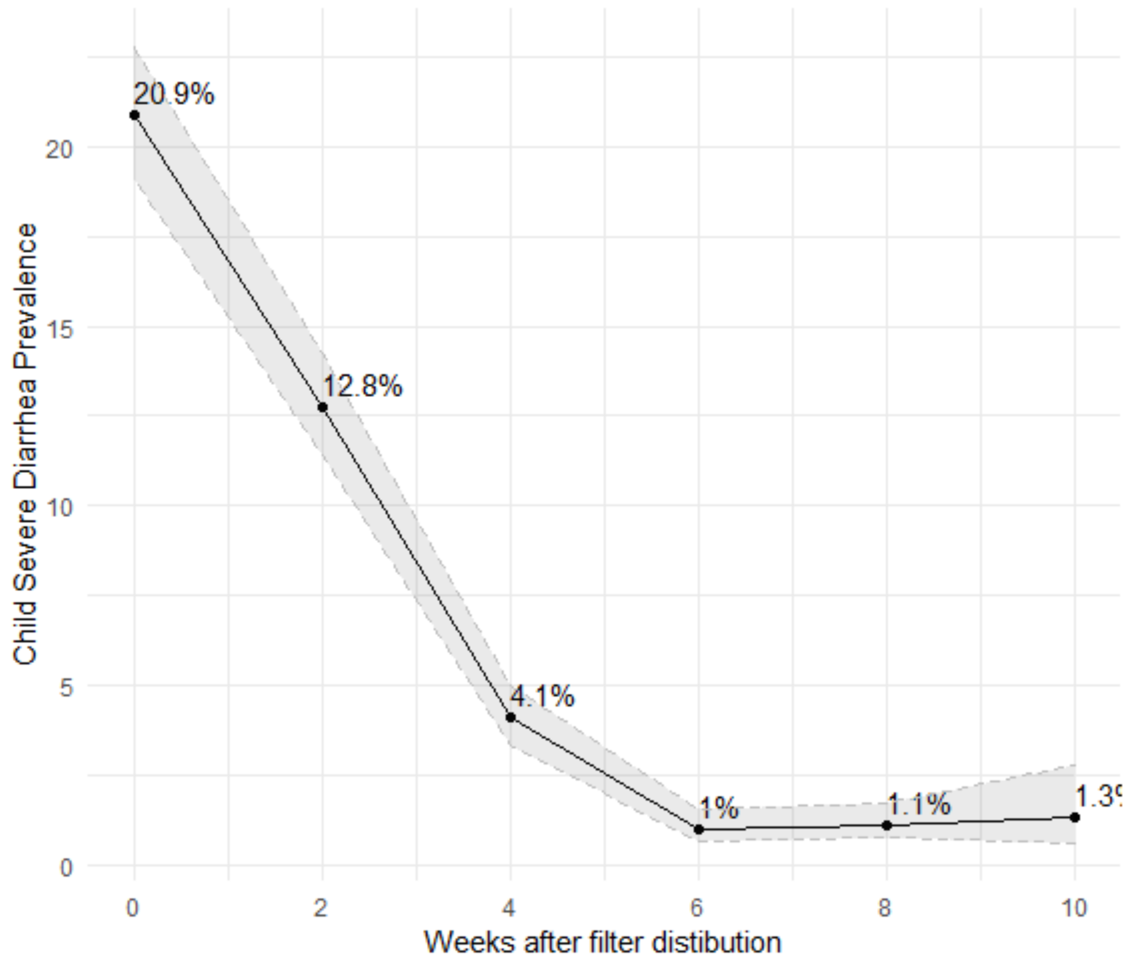
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with working adults, approximately 2.2% of participants report having had an adult in the household miss work due to having diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5K. Model adjusted estimates of Severe diarrhea prevalence among school aged children (aged 5-17 years)



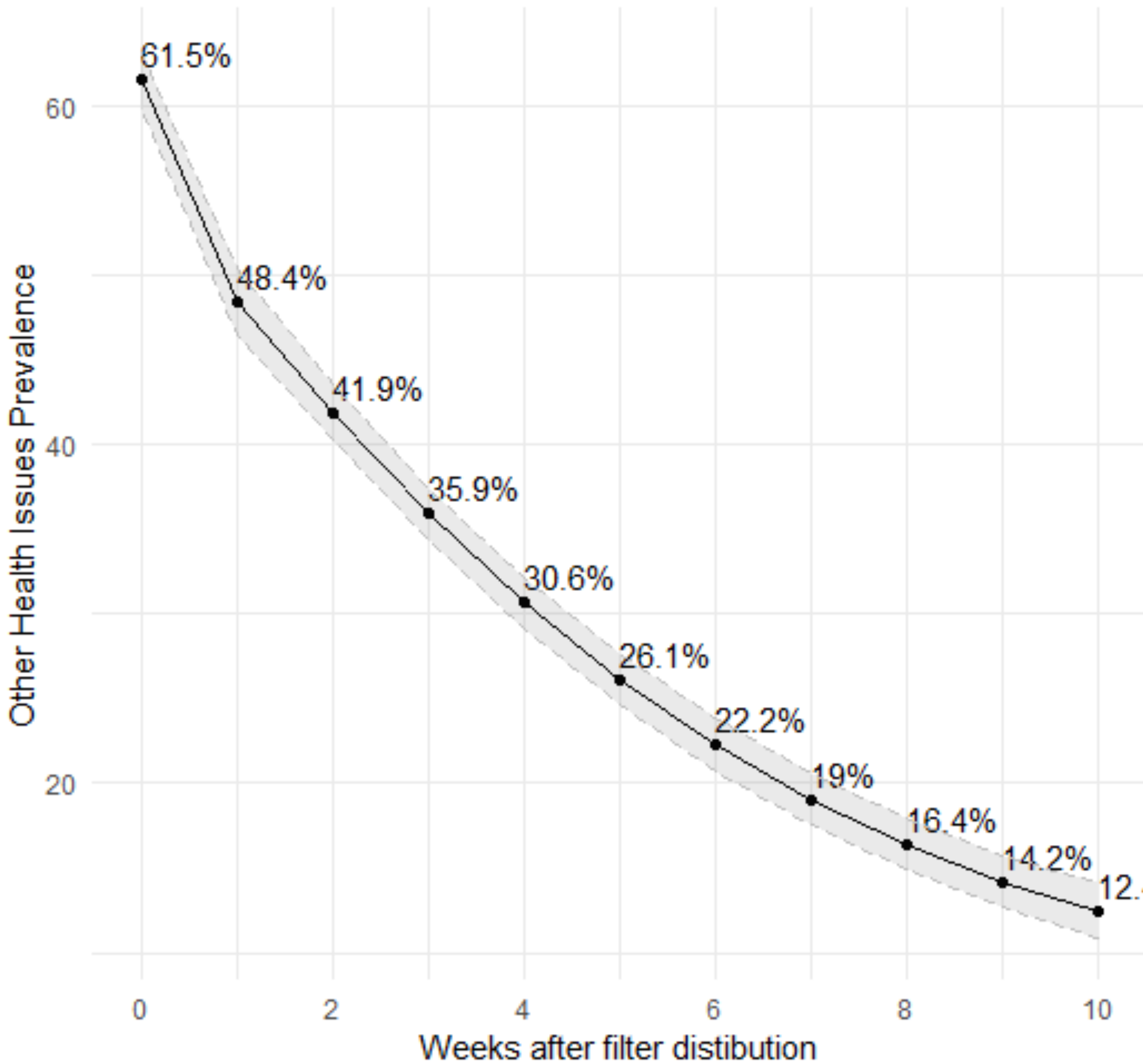
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with school-aged children, approximately 2.1% of participants report having had a child in the household miss school due to having diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5L. Model adjusted estimates of Severe diarrhea prevalence among school aged children (aged 5-17 years) – without smoothing



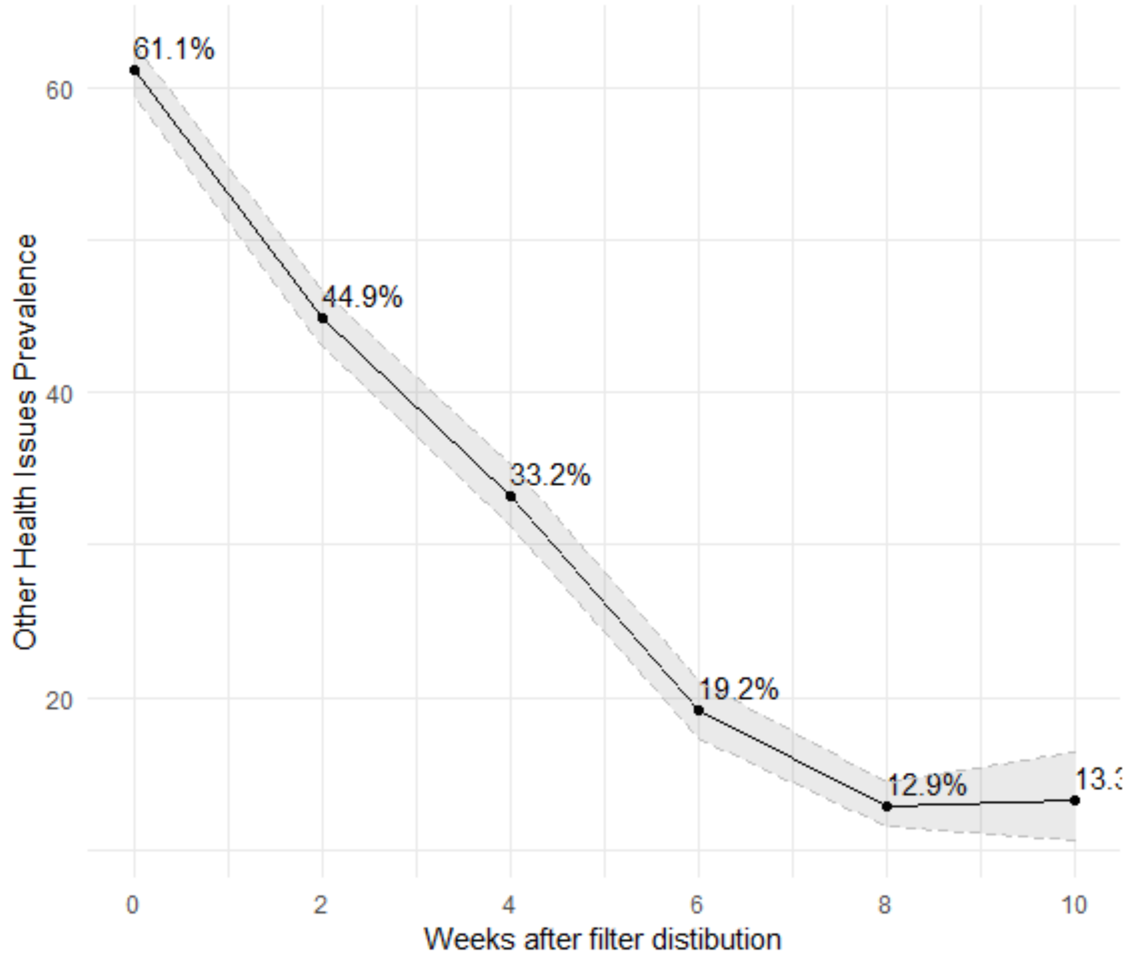
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), and among households with school-aged children, approximately 1.0% of participants report having had a child in the household miss school due to having diarrhea within the previous two weeks within 6-weeks of filter distribution.

Figure 5M. Model adjusted estimates of any other household health problem (last two weeks)



Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), approximately 22.2% of participants report having had any other household health issues within the previous two weeks within 6-weeks of filter distribution.

Figure 5N. Model adjusted estimates of any other household health problem (last two weeks) – without smoothing

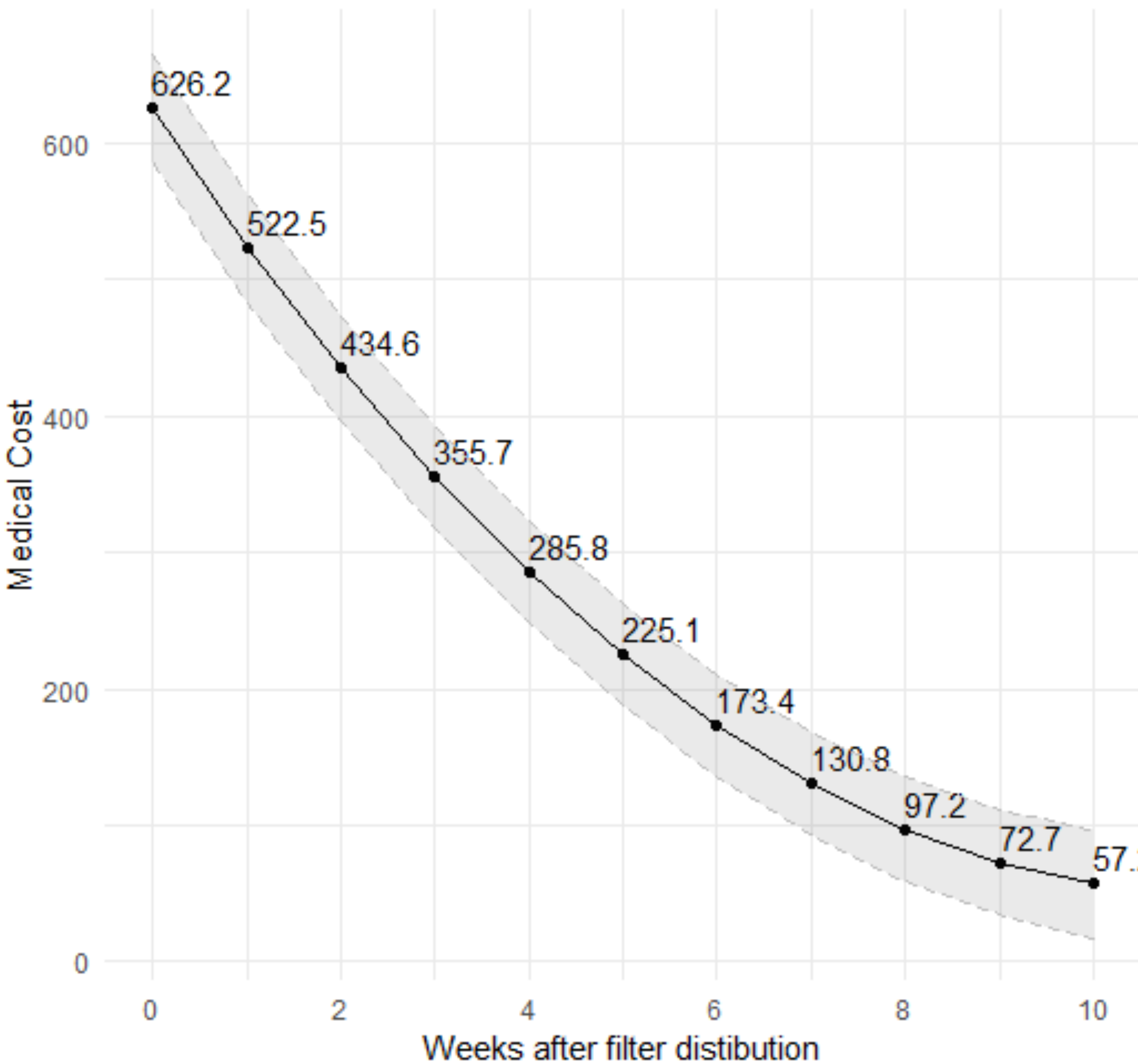


Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), approximately 22.2% of participants report having had any other household health issues within the previous two weeks within 6-weeks of filter distribution.

Section 6. Economic Impacts

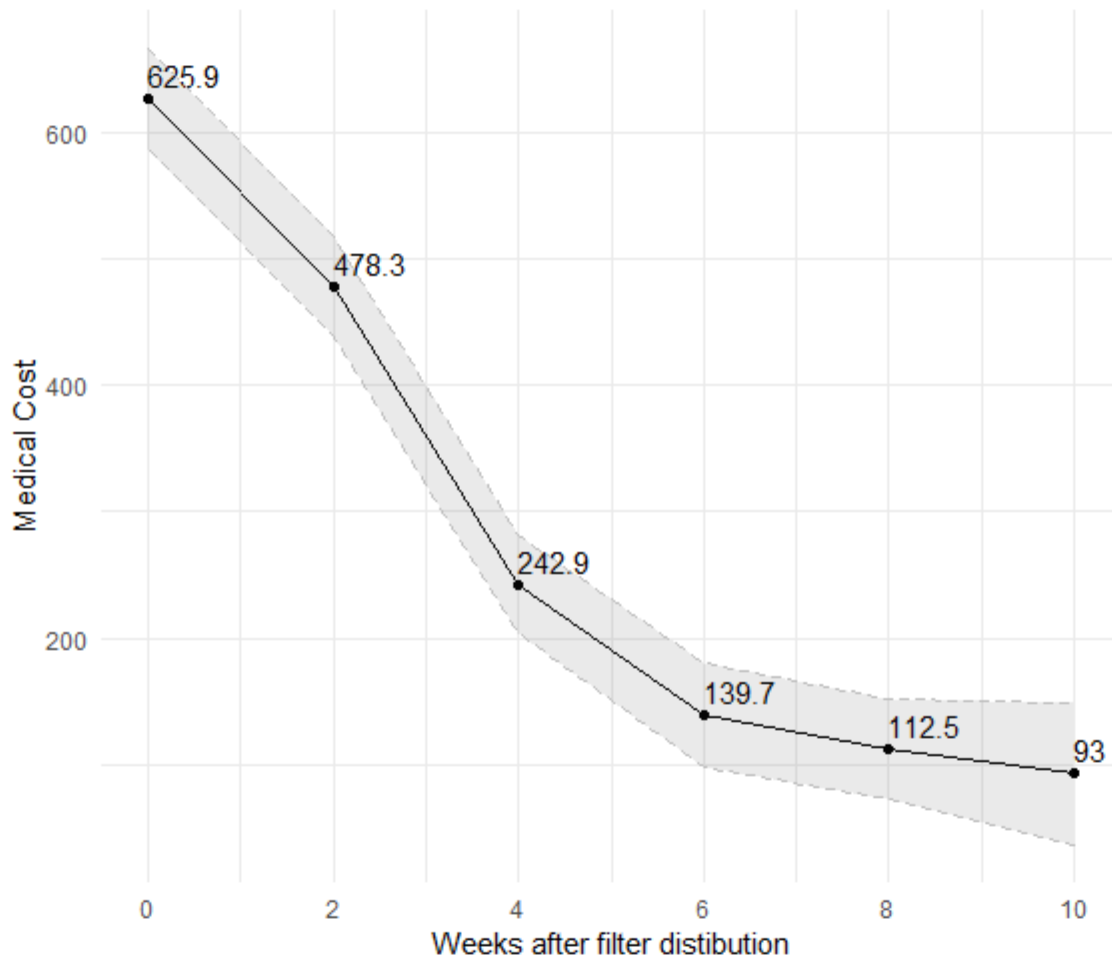
The following graphs illustrate results from statistical models (see Section 3 for details) exploring health patterns over time after filter distribution, WASH training and anti-parasitic medication distribution. Initial interpretations are provided below each figure.

Figure 6A. Model adjusted estimates of reported medical costs (last 30 days) per household (Ksh)



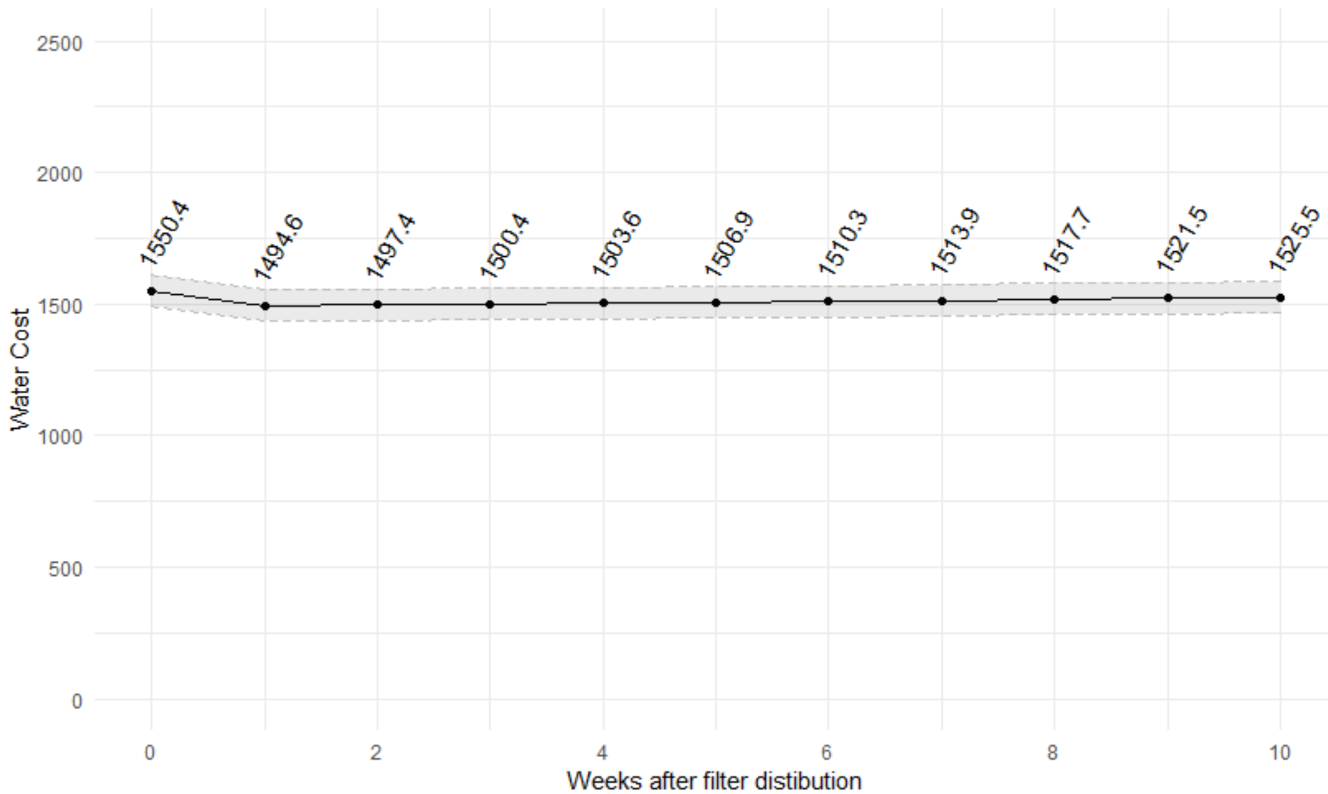
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), monthly household medical costs are estimated at 173.4 Ksh for a typical household within 6-weeks of filter distribution.

Figure 6B. Model adjusted estimates of reported medical costs (last 30 days) per household (Ksh) – without smoothing



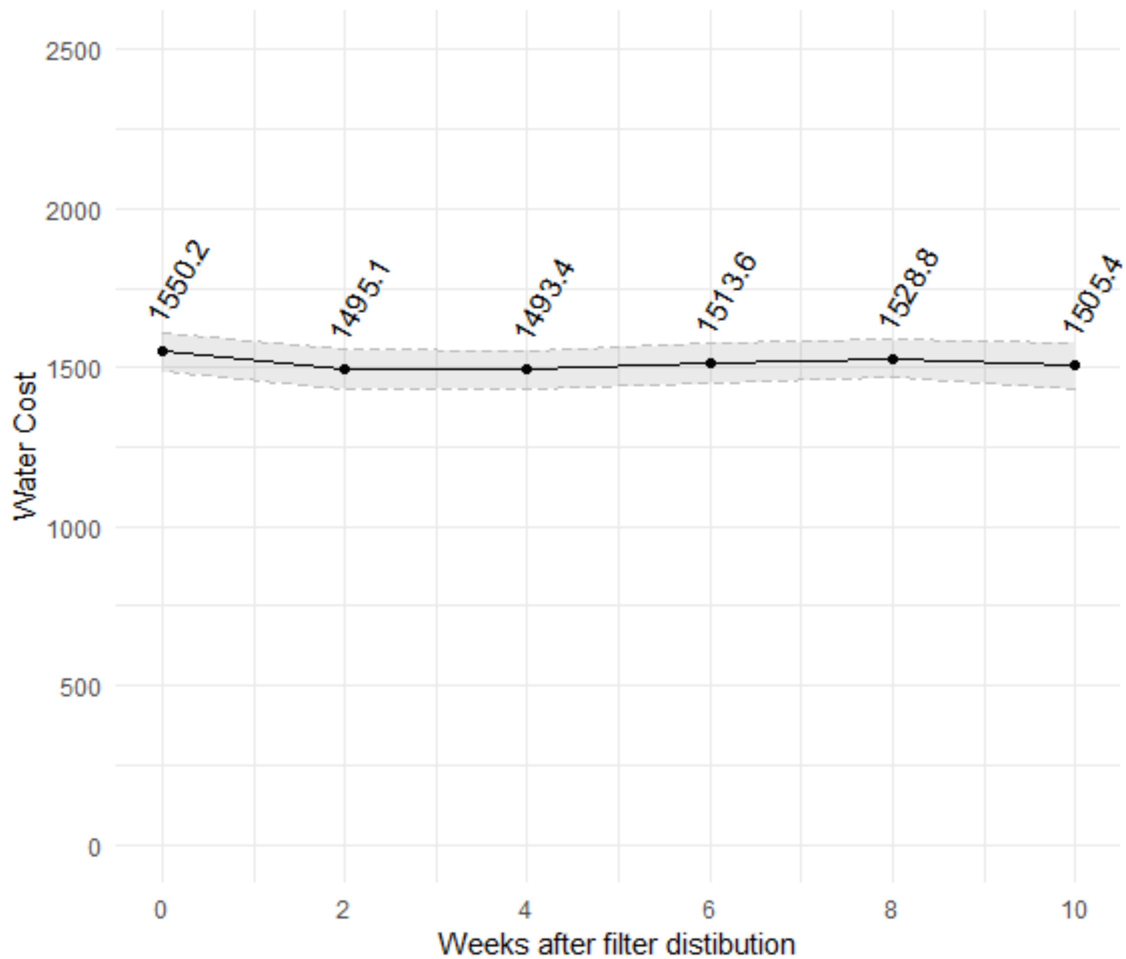
Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), monthly household medical costs are estimated at 139.7 Ksh for a typical household within 6-weeks of filter distribution.

Figure 6C. Model adjusted estimates of reported water costs (last 30 days) per household (Ksh)



Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), monthly household water costs are estimated at 1510.3 Ksh for a typical household within 6-weeks of filter distribution. This cost is a non-significant (p>0.05) from baseline, reflecting the culture in Kibera in which (virtually) all water is purchased.

Figure 6D. Model adjusted estimates of reported water costs (last 30 days) per household (Ksh) – without smoothing



Example of interpretation: After accounting for covariates (e.g., household size; seasonality; interviewer heterogeneity), monthly household water costs are estimated at 1513.6 Ksh for a typical household within 6-weeks of filter distribution. This cost is a non-significant ($p>0.05$) from baseline, reflecting the culture in Kibera in which (virtually) all water is purchased.

Section 7. Water testing results

In addition to household surveys at distribution and follow up, Aquora has conducted representative water quality testing of drinking water sources in the seven neighborhoods of Kibera where distribution of filters took place during the study period. In total, 25 water sources were tested (21 drinking water sources, 3 sources that could be used for drinking water, and 1 surface source that is never used for drinking according to residents of Kibera). The water sources represent 11 different types, which captures the range of where residents obtain their water (Fig. 7A). Samples were collected in November 2021, March 2022, and May 2022, correlating to the timing of distribution of filters in different neighborhoods.

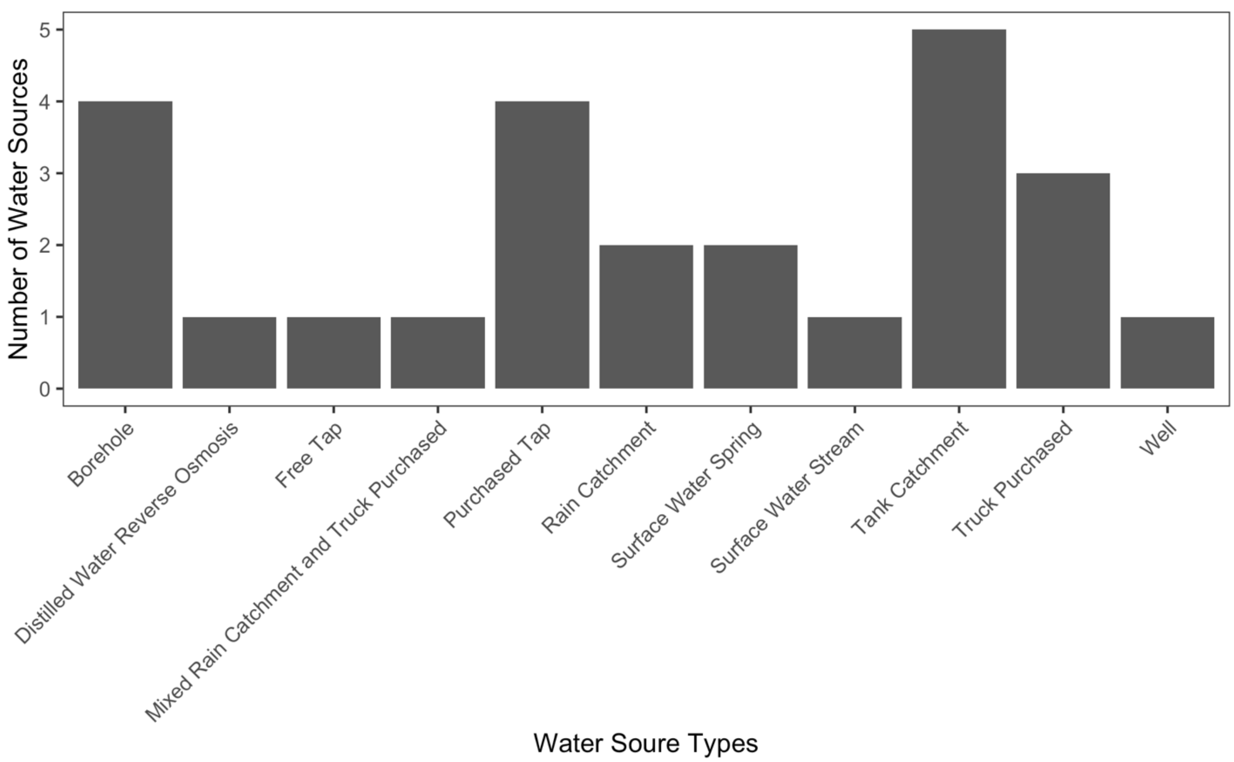


Figure 7A. Distribution of water source types sampled in Kibera.

Each water source was tested to determine 1) WHO Risk Category for *E. coli* and Total Coliforms, 2) potential presence of other bacterial pathogens through bacterial community 16S rRNA sequencing, and 3) presence of toxic metals above WHO limits for safe consumption of water over time. The purpose of the first two testing methods is to characterize the potential risk to residents of drinking untreated water from sources typically available to them. The Sawyer filters can remove these potential bacterial pathogens, rendering the water safe to drink. The third testing method addresses potential long term health risks associated with drinking untreated or Sawyer filtered water.

Results of Bacterial Testing: In field testing of *E. coli* and Total Coliform concentrations leads to the conclusion that most water sources sampled represent a high risk of water-borne illness to residents. Results of in-field testing for *E. coli* showed that 6 water sources were in the WHO risk category of "Unsafe"; 13 water sources were in the "Low Risk/Safe" category for *E. coli* (Fig. 7B). In contrast, 18 water sources were categorized as

"Unsafe" for Total Coliforms with only 2 water sources being characterized as "Low Risk/Safe" (Fig. 7C). It is clear from these data that most water sources contain bacterial contamination at levels that are unsafe for consumption.

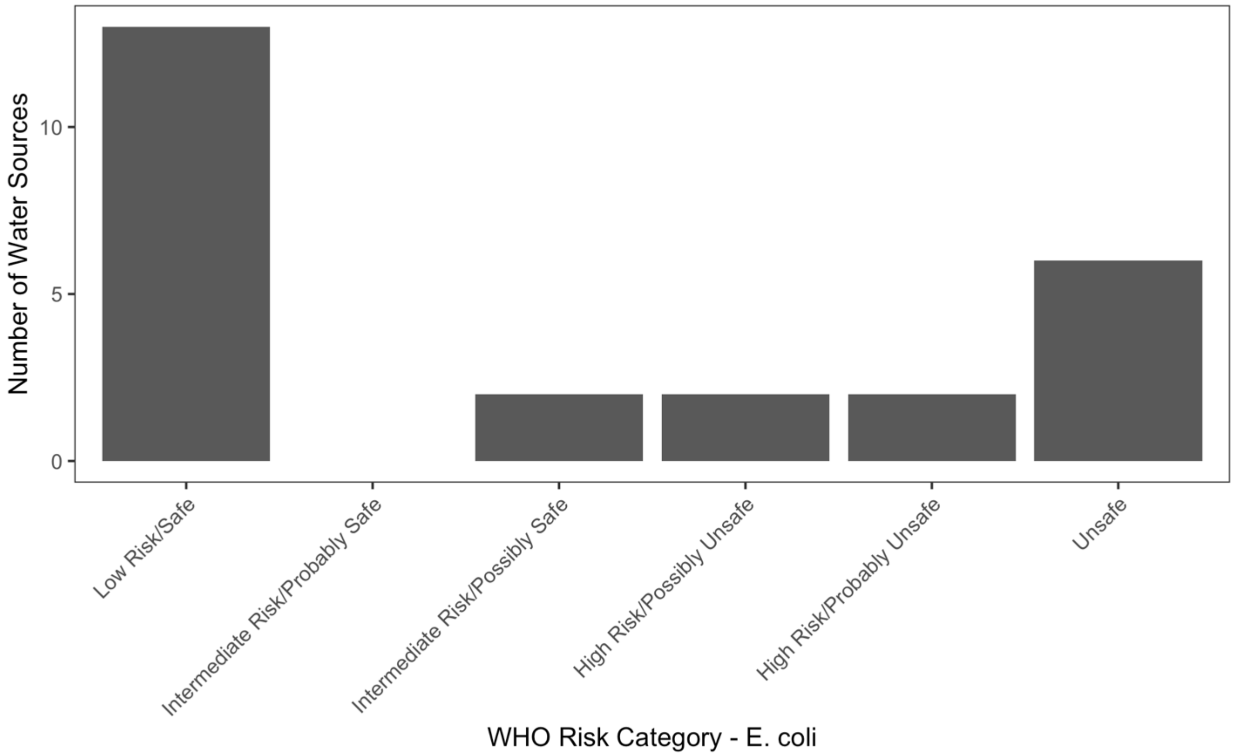


Figure 7B. Field testing of water sources for *E. coli* concentration. Aquagenx Compartment Bag Test EC+TC Most Probable Number tests were conducted for all water sources sampled. Based on the determined MPN (cells/100 mL), each water source was categorized into a WHO Health Risk category. The number of water sources in each category is depicted by the height of the bar for each category.

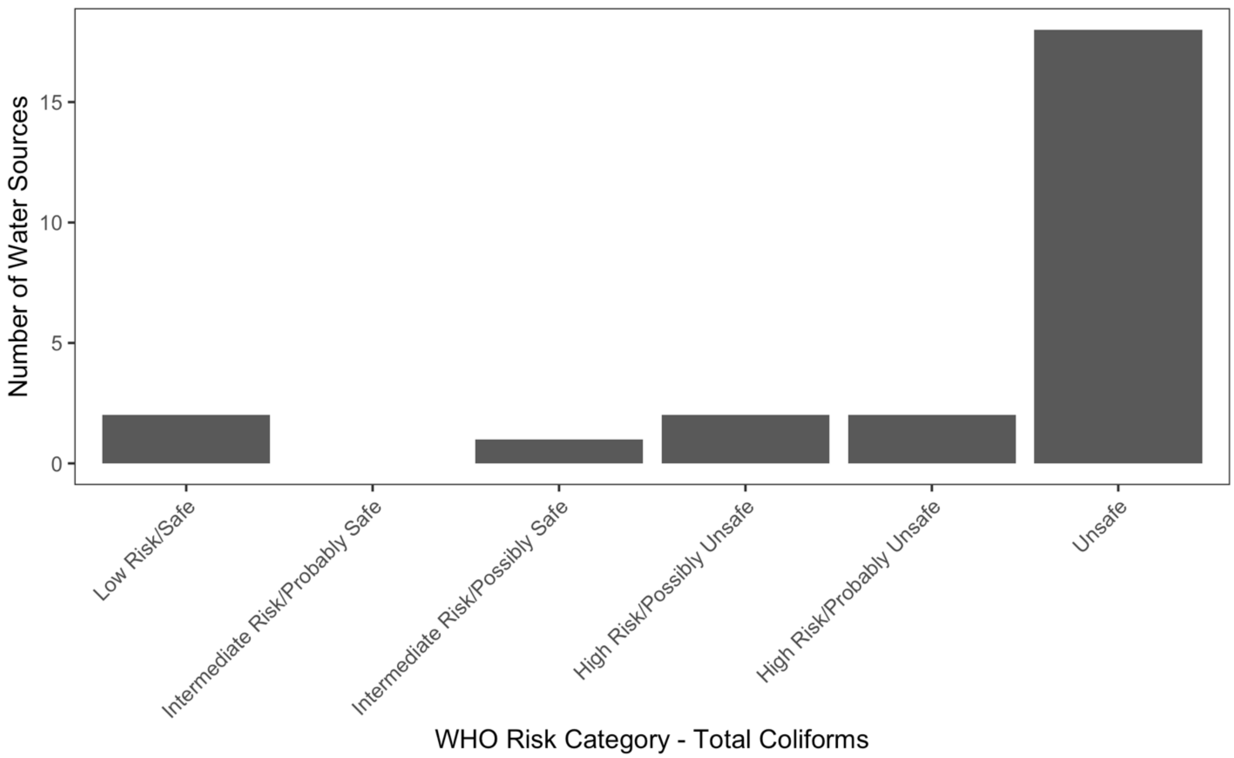


Figure 7C. Field testing of water sources for Total Coliform concentration. Aquagenx Compartment Bag Test EC+TC Most Probable Number tests were conducted for all water sources sampled. Based on the determined MPN (cells/100 mL), each water source was categorized into a WHO Health Risk category. The number of water sources in each category is depicted by the height of the bar for each category.

Other potential pathogens known to be associated with water were identified in all water sources tested to date by bacterial community sequencing (Table 7A). This methodology allows for a more specific description of potential health risks than the EC+TC MPN testing methods. Each of the bacterial genera listed are known to cause either respiratory, gastrointestinal, or skin, eye and wound diseases transmitted through contaminated water. All 24 water sources with sufficient sequencing read depth were observed to contain at least three of the genera listed in the table. Water source 370, representing the river that runs through Kibera and that is not used as a drinking water source, contained seven of the genera.

Water Source	*Acinetobacter*	*Aeromonas*	*Campylobacter*	*Escherichia-Shigella*	*Legionella*	*Leptospira*	*Mycobacterium*	*Pseudomonas*	*Staphylococcus*	*Yersinia*
312	Yes	Yes	No	No	Yes	No	No	Yes	No	No
313	No	No	No	No	Yes	No	No	No	Yes	No
316	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	No
352	Yes	No	No	No	Yes	No	No	No	Yes	No
364	No	Yes	No	Yes	Yes	No	Yes	No	No	No
365	No	Yes	No	No	Yes	No	No	Yes	No	No
367	No	No	No	No	Yes	No	Yes	Yes	Yes	No
368	No	Yes	No	No	Yes	No	Yes	No	No	No
369	Yes	Yes	No	No	Yes	No	No	Yes	Yes	Yes
370	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes
371	Yes	No	No	No	Yes	No	No	Yes	No	No
373	No	Yes	Yes	Yes	No	No	No	No	No	No
374	No	No	No	No	Yes	No	No	No	Yes	No
375	No	No	No	No	Yes	No	Yes	Yes	Yes	No
377	No	No	No	Yes	No	No	No	Yes	Yes	No
389	Yes	Yes	No	No	No	No	Yes	No	Yes	No
391	Yes	Yes	No	No	Yes	No	No	Yes	Yes	No
392	No	Yes	No	No	Yes	No	Yes	No	No	No
394	Yes	No	No	Yes	Yes	No	Yes	No	No	No
397	Yes	Yes	No	No	Yes	No	No	No	No	No
401	Yes	Yes	No	No	No	No	No	Yes	No	No
402	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
403	Yes	No	No	No	Yes	No	Yes	Yes	No	No
404	Yes	Yes	No	No	No	No	No	Yes	No	No

Table 7A. Presence or Absence of Potential Bacterial Pathogens. 16S rRNA community sequencing of 24 water sources was used to identify the presence or absence of 18 bacterial genera with known waterborne pathogenic species. Ten genera were identified amongst the samples tested. All samples contained at least three genera. Presence is indicated as "Yes" and shading. Water source 372 produced sequencing read depth below the filtering cutoff and is not reported in this table. Full details about each water source are available in the Water Quality Dashboards.

The results from both *E.coli* / Total Coliform and bacterial community testing support the high baseline rates of diahrrea and potentially other waterborne health issues observed from the self-reported household surveys.

Results of Dissolved Metals Testing: To assess water quality for potential long term health effects, we tested all the sources sampled for the presence of eight metals known to cause disease and for which the WHO sets out allowable limits in drinking water. Metals testing of drinking water sources showed that all but one source had either undetectable levels or detectable levels below WHO health guidelines (Table 7B). One drinking water source (Source 374) showed arsenic levels above WHO health guidelines. Water Source 374 is a borehole drilled near a local hospital; the water is intended for use by the hospital, but it is also available as a free source of water to the local community. It is estimated that this source serves around 70 people per day. It is recommended that this source not be used for drinking water in the future and that those using the source switch to one of the other sources readily available that do not have arsenic. Overall, the results of metals testing support the use of Sawyer filters in the Kibera slum and do not indicate the need for secondary level filtration to remove metals.

Water Source	Arsenic	Barium	Cadmium	Chromium	Copper	Lead	Nickel	Selenium
312	No	No	No	No	No	No	No	No
313	No	No	No	No	No	No	No	No
316	No	No	No	No	No	No	No	No
352	No	No	No	No	No	No	No	No
364	No	No	No	No	No	No	No	No
365	No	No	No	No	No	No	No	No
367	No	No	No	No	No	No	No	No
368	No	No	No	No	No	No	No	No
369	No	No	No	No	No	No	No	No
370	No	No	No	Yes	No	No	No	No
371	No	No	No	No	No	No	No	No
372	No	No	No	No	No	No	No	No
373	No	No	No	No	No	No	No	No
374	Yes	No	No	No	No	No	No	No
375	No	No	No	No	No	No	No	No
377	No	No	No	No	No	No	No	No
389	No	No	No	No	No	No	No	No
391	No	No	No	No	No	No	No	No
392	No	No	No	No	No	No	No	No
394	No	No	No	No	No	No	No	No
397	No	No	No	No	No	No	No	No
401	No	No	No	No	No	No	No	No
402	No	No	No	No	No	No	No	No
403	No	No	No	No	No	No	No	No
404	No	No	No	No	No	No	No	No

Table 7B. WHO Above/Below health guidelines for metals in drinking water. Water sources were tested for eight metals with regulatory guidelines from the WHO. Metals detected at concentrations above the WHO guidelines are indicated with “Yes” and shading. A “No” indicates that metals were not detected at concentrations above the guidelines, however, it is possible that a metal was present but below the limits (not shown). One drinking water source (374) shows arsenic levels above the guidelines. Chromium is present in one water source (370), which is not used as drinking water source.

Individual dashboard summaries of each water source collected are included as an appendix to this report. The dashboards include pictures and general descriptions of the source along with results for the three water quality tests conducted.

Section 8. Anti-parasitic medication, WASH training and POU filter effects on health

Goal: The Bucket Ministry teaches WASH principles (e.g., hand-washing before meals, after being in the bathroom and distributes an anti-parasitic medication, as part of the group filter distribution process currently occurring in Kibera. Within the larger investigation of the impact of filter distributions on health in Kibera, we designed and implemented an ethical substudy to attempt to isolate filter effects on health.

Protocol:

Numerous discussions took place with TBM during January/February 2022 to design a substudy to isolate filter effects on self-reported health. To ensure sufficient statistical power, ensure ethical treatment of participants, be feasible for TBM/teams on the ground and still accomplish the study goals, anti-parasitic medication was not distributed at baseline or the first follow-up for distributions occurring within the month of March. Instead, anti-parasitic medication was distributed at the second follow-up visit for these households.

In early March 2022, a member of the Aquora team (Best) traveled to Kibera with TBM to train data collectors/field teams on the new protocol, as well as embed with field teams to ensure consistent implementation of surveys for the overall study.

During March 2022, distributions to N=2642 households did not receive anti-parasitic medications until the second follow-up time point (approximately 5 weeks after filter distribution).

We followed a similar statistical modelling process to that described elsewhere in this report (see Methods) to adjust for numerous effects (seasonality, surveyor, changing household demographics), in order to yield adjusted estimates of self-reported diarrhea prevalence for households not receiving anti-parasitic medications.

Substudy Results

Figure 8A below provides adjusted estimates of self-reported household diarrhea prevalence.

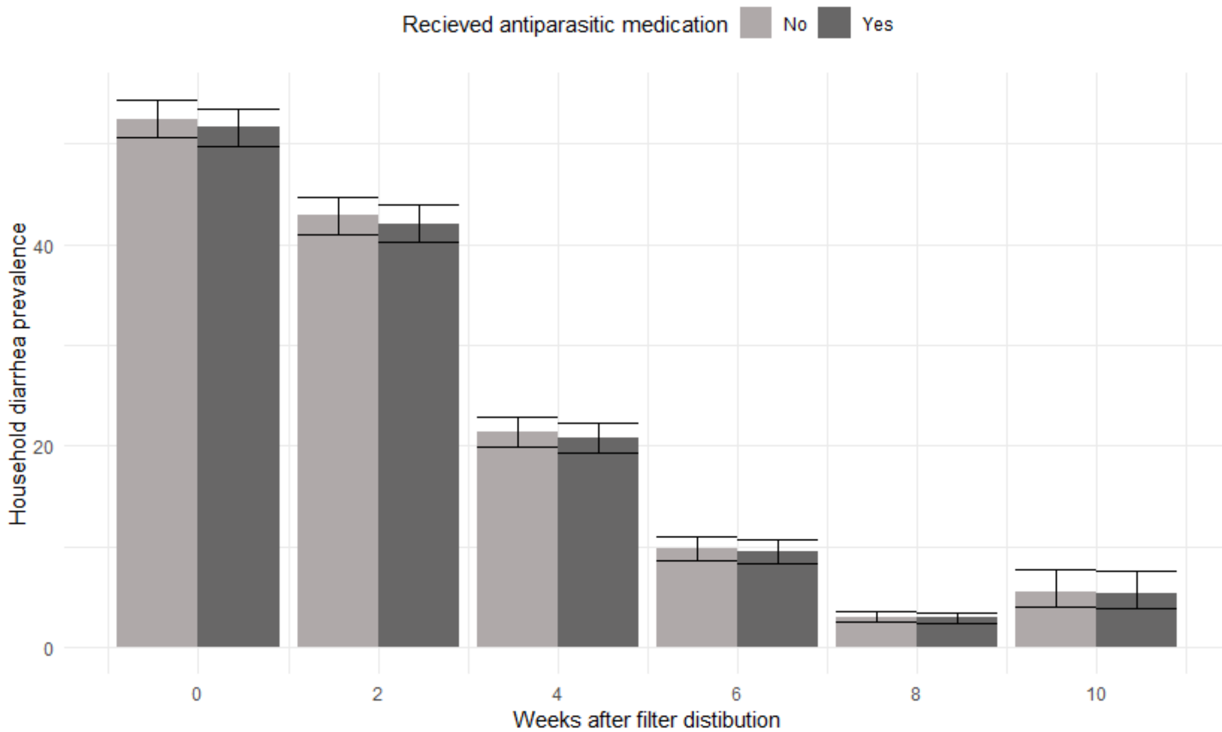


Figure 8A illustrates little to no evidence of an effect of anti-parasitic medication distribution on self-reported diarrhea prevalence at any time during the study (P>0.05 in all cases).

Attempts to isolate the effects of WASH training separate from filter implementation were unsuccessful, because the vast majority of individuals claimed to wash their hands after defecation (97.4% distribution; 98.3% follow-up), before eating (98.6% distribution; 98.5% follow-up) and before food prep (86.7% distribution; 92.6% follow-up).

Substudy discussion/Interpretation

Analyses provide fairly conclusive evidence of a lack of impact from the use of anti-parasitic medications on the observed reduction in self-reported household diarrhea prevalence. Attempts to further isolate filter effects from WASH training were unsuccessful due to the large percentage of individuals self-reporting regular hand washing.

The reasons for this the lack of impact of Albendazole in this study may be due to the limited nature of the antiparasitic medication intervention. Due to Kenyan regulatory constraints, the distribution of antiparasitic medications consisted of only a single, low dose of Albendazole at baseline for the majority of participants in the study (the substudy pushed the timing of this dose to the second follow-up).

Substudy Limitations

Randomized controlled trials (RCTs) remain the gold standard to attribute causality to interventions. However, practical concerns (feasibility; ethical) regarding implementing an RCT for the present study led to the implementation of the study protocol outlined here. Statistical adjustment of confounding variables, a straightforward field protocol, and oversight from the Aquora team mitigate many of the potential limitations of this non-randomized interventional trial.

While the anti-parasitic substudy took place entirely in March to meet feasibility concerns of TBM, potentially confounding seasonal health effects with the treatment arm (withholding anti-parasitic meds), no noticeable change in self-reported prevalence of diarrhea at baseline was observed before (February) or after (April) the substudy, making it unlikely to materially contribute to the findings observed here.

The study was analyzed with intent-to-treat analysis, which ignores whether or not individuals chose to take the Anti-parasitic drug and at what time.

Other limitations of the entire study still apply here (e.g., inability to know to what extent self-reported diarrhea prevalence aligns with actual diarrhea prevalence, etc.).

Section 9. Discussion/Conclusions

Key take home messages

1. The Sawyer filter, combined with anti-parasitic medication use and WASH training, leads to reductions in health outcomes, and improved in economic measures in the Kibera slum at rates in line with those observed in non-urban contexts in prior studies[15].
2. Filter utilization rates are high (as has been observed in other contexts) after the first follow-up is completed, suggesting that the group distribution model is effective, but that the first follow-up is critical for retraining and to ensure filter utilization.
3. Representative water quality testing indicates a high frequency of drinking water source contamination with *E. coli* and Total Coliforms. Potential waterborne pathogens were detected in all water sources tested. Filtration of drinking water with Sawyer filters can reduce the risk from this contamination.
4. Representative water quality testing indicates a very low frequency of dissolved metals present above WHO guidelines for drinking water. Use of Sawyer filters may be sufficient to remove known contaminants of concern.
5. There was little to no effect of the anti-parasitic drugs on this study.

[15] https://pubmed.ncbi.nlm.nih.gov/31410085/ and https://pubmed.ncbi.nlm.nih.gov/33397511/

Section 10. Limitations

While the study, design and analysis were carefully planned and implemented, there are a number of limitations worth noting:

1. This study is observational. Thus, the resulting associations are not necessarily causal. However, given (a) the prospective nature of the study and (b) the control of numerous covariates via the statistical modelling process, the resulting associations are strongly suggestive. Notably, however, randomized controlled trials are necessary to confirm the suggestive associations as truly causal.
2. This study is based on self-reports of health, economic and other outcomes. While self-reports of health outcomes are generally valid and reliable, self-report can be substantially impacted by a variety of biases (interviewer, recall, etc.). Thus, the best way to view the associated results is (a) in light of other similar studies using self-reported measures (comparatively) and (b) in terms of relative changes in self-reported outcomes (e.g., 2-fold reduction, etc.). Caution should be used if attempting to utilize the data as true measures of absolute levels of disease or economic benefit.
3. To minimize interviewer time, a single member of the household (typically the female head of the house) was interviewed and reported on the health of the entire household. Thus, self-reported information is further compounded by the ability of the respondent to recall and know health, economic and utilization information for individual household members.
4. Substantial efforts were made to clean the data (see Methods), however, the reality of challenges with field implementation in an urban slum, hundreds of unique data collectors, challenges with survey item propagation and a prospective study with multiple-follow-ups means that data is not as well-controlled and 'clean' as would be obtained in ideal circumstances. The need to eliminate approximately 30% of the data as part of data cleaning is substantial and we are assuming that missing data is approximately distributed at random.
5. The survey itself was implemented in the context of a larger discipleship/evangelism project by TBM. Efforts were made to separate questions related to filter utilization, health impacts and related questions from other survey questions and discussion with surveyors about discipleship related topics. However, there is no guarantee that order effects/carry-over effects from the discipleship process run by the interviewer are impacting study findings.
6. The study was implemented over a 6-month period in Kibera (Nov 2021 – May 2022). While statistical models accounted for temporal effects, the 6-month period study means that seasonal cycles of rainy/dry, which can impact water access, cleanliness and related infectious disease cycles were only observed for 6 months of a calendar year. Notably, however, the study period cover parts of two rainy seasons (Nov/Dec; Apr/May), as well as a dry season (Jan/Feb/Mar), though data was not collected during the winter dry season (Jun-Oct), which sees the coldest temperatures of the year and peak incidence of infectious diseases that typically spike in winter months (e.g., influenza).
7. Filter distributions were done in a group format, which yields substantial efficiencies and consistencies of training and data collection. However, data analysis suggests that many individuals are not using or receiving the filter benefits until after the first follow-up which occurs at the individual home with a single data collector who have provide training.
8. Water quality testing was limited to 25 water sources in seven neighborhoods in Kibera. While this gives a representative sense of the types of water sources used, it was not designed to be a comprehensive study of water quality in the area.
9. Sequencing of bacterial communities targeting the 16S rRNA gene limits identification of bacteria to the genus level. This allows for the identification of genera that are known to contain bacterial

waterborne pathogens as members. However, this method is not a direct detection of pathogens and should not be interpreted as such.

10. Water quality testing of chemical contaminants was limited to dissolved metals. Other types of chemical contaminants found in urban environments were not tested (e.g., volatile organic chemicals (VOCs) associated with petroleum and other organics).

Section 11. Suggestions for next steps

1. Future controlled trials are needed to isolate and attribute causality to the Sawyer filter
2. Studies are needed which further validate the use of self-reported health measures, a single respondent, and the potential impact of a combined health/evangelism approach
3. Increased water quality testing, including dissolved metals and volatile organic chemicals, is needed to ensure that filter choice is targeted to the water issues of the household
4. Long term studies of filter utilization, filter cleaning compliance and long-term health effects are needed – Kibera is a good study environment for this given TBM infrastructure and plans to be in place for the next two to three years
5. Consideration of the group distribution model is needed given the low utilization rates prior to follow-up #1
6. Further studies in other urban environments/slums and with other non-profit/government partners is necessary to ensure that findings are replicable

Section 12. Questions and timeline

While the results presented constitute an initial set of results to be shared internally at Sawyer as well as with other interested parties, the results here have not yet been subject to peer-review and, thus, should be viewed as *preliminary*. Over the next months (anticipated completion by February 2023), Aquora will use the preliminary results above to conduct additional robustness testing, develop an enhanced literature review and scientific discussion, and submit the resulting materials for peer-review and publication. Upon submission for peer-review, the draft manuscript will be submitted to a pre-print server for immediate, open access to anyone in the world. The draft manuscript, and, ultimately, the final, peer-reviewed manuscript should be viewed as the most authoritative, complete and comprehensive analysis of the related data and should be cited, when available, in lieu of this report.

Section 13. Appendices

Section 13.A. Surveyor heterogeneity.

Despite on-going monitoring and intervention, certain surveyors consistently yielded questionable survey results. Surveyors with very high reported diarrhea prevalence (across participants they surveyed), very low diarrhea prevalence, or dramatic changes (from high to low or vice versa after intervention), were flagged and dropped from the dataset as part of the data cleaning processing. See Tables 13.A.1 and 13.A.2 for examples.

Table 13.A.1 – Diarrhea prevalence by Surveyor ID#, overall and by month - distribution.

UserID	Nov	Dec	Jan	Feb	Mar	Apr	n	Total	Kept
Surveyor41	NA	NA	99%(141)	78%(41)	90%(101)	90%(67)	350	92%(350)	No
Surveyor13	21%(81)	NA	67%(133)	54%(52)	65%(17)	0%(1)	284	51%(284)	Yes
Surveyor18	35%(49)	0%(1)	25%(110)	60%(43)	48%(46)	NA	249	37%(249)	Yes
Surveyor3	15%(20)	0%(4)	9%(90)	20%(30)	27%(98)	NA	242	18%(242)	Yes
Surveyor19	24%(21)	NA	28%(68)	47%(70)	45%(80)	NA	239	39%(239)	Yes
Surveyor26	14%(14)	NA	84%(45)	68%(56)	61%(114)	NA	229	64%(229)	Yes
Surveyor50	NA	NA	66%(128)	63%(43)	NA	49%(45)	216	62%(216)	Yes
Surveyor33	NA	0%(2)	35%(81)	50%(54)	41%(59)	67%(12)	208	42%(208)	Yes
Surveyor10	68%(88)	NA	40%(20)	18%(33)	39%(64)	NA	205	48%(205)	Yes
Surveyor8	21%(14)	NA	16%(49)	45%(29)	38%(78)	36%(28)	198	32%(198)	Yes
Surveyor70	NA	NA	NA	47%(60)	35%(55)	32%(80)	195	37%(195)	Yes
Surveyor20	53%(19)	100%(1)	70%(91)	47%(72)	45%(11)	NA	194	59%(194)	Yes
Surveyor59	NA	NA	30%(53)	34%(83)	55%(40)	61%(18)	194	40%(194)	Yes
Surveyor43	NA	NA	12%(52)	64%(28)	75%(87)	74%(19)	186	55%(186)	Yes
Surveyor65	NA	NA	46%(24)	88%(52)	88%(78)	71%(28)	182	80%(182)	Yes
Surveyor11	50%(70)	100%(1)	22%(41)	33%(12)	NA	31%(51)	175	37%(175)	Yes
Surveyor5	38%(52)	NA	94%(33)	100%(50)	100%(39)	NA	174	80%(174)	No
Surveyor7	0%(17)	0%(1)	20%(56)	0%(1)	18%(62)	24%(29)	166	17%(166)	No
Surveyor15	29%(87)	NA	75%(4)	48%(31)	100%(1)	21%(38)	161	32%(161)	Yes
Surveyor21	35%(20)	NA	43%(51)	52%(33)	46%(56)	NA	160	45%(160)	Yes
Surveyor25	69%(13)	NA	37%(38)	39%(36)	48%(31)	48%(40)	158	45%(158)	Yes
Surveyor30	9%(11)	NA	4%(57)	20%(15)	16%(75)	NA	158	11%(158)	No
Surveyor55	NA	NA	6%(17)	67%(39)	46%(102)	NA	158	47%(158)	Yes

Table 13.A.2 – Diarrhea prevalence by Surveyor ID#, overall and by month – follow-up #1.

UserID	Nov	Dec	Jan	Feb	Mar	Apr	n	Total	Kept
Surveyor17	12%(8)	11%(19)	38%(8)	42%(43)	10%(41)	0%(54)	190	15%(190)	No
Surveyor20	12%(8)	58%(38)	29%(14)	11%(35)	2%(84)	NA	179	18%(179)	Yes
Surveyor25	86%(7)	97%(36)	98%(46)	86%(7)	38%(81)	NA	178	69%(178)	No
Surveyor72	NA	NA	14%(37)	83%(6)	60%(35)	33%(70)	155	35%(155)	Yes
Surveyor9	50%(2)	0%(9)	29%(7)	40%(40)	63%(35)	83%(54)	150	59%(150)	Yes
Surveyor50	NA	11%(28)	25%(8)	30%(54)	15%(13)	22%(32)	145	21%(145)	Yes
Surveyor18	0%(2)	93%(14)	81%(32)	84%(37)	25%(16)	31%(29)	143	58%(143)	Yes
Surveyor59	NA	NA	70%(53)	100%(5)	54%(78)	NA	136	62%(136)	Yes
Surveyor3	30%(10)	6%(17)	17%(6)	22%(46)	50%(14)	21%(39)	132	23%(132)	Yes
Surveyor44	NA	43%(35)	36%(14)	36%(28)	67%(3)	37%(51)	132	39%(132)	Yes
Surveyor48	NA	81%(16)	100%(7)	92%(12)	100%(51)	89%(45)	132	93%(132)	No
Surveyor28	20%(10)	11%(18)	83%(6)	41%(41)	36%(11)	20%(41)	130	29%(130)	Yes
Surveyor57	NA	NA	48%(58)	80%(10)	63%(59)	33%(3)	130	57%(130)	Yes
Surveyor45	NA	5%(19)	0%(8)	11%(37)	10%(63)	0%(1)	128	9%(128)	Yes
Surveyor81	NA	NA	58%(12)	29%(7)	45%(66)	44%(41)	128	45%(128)	Yes
Surveyor41	100%(1)	NA	95%(43)	50%(4)	52%(71)	57%(7)	126	67%(126)	Yes
Surveyor6	25%(4)	65%(17)	75%(8)	74%(34)	74%(27)	48%(29)	122	64%(122)	Yes
Surveyor89	NA	NA	0%(25)	25%(20)	32%(76)	NA	121	24%(121)	Yes
Surveyor10	14%(7)	0%(20)	0%(47)	43%(7)	100%(36)	100%(3)	120	36%(120)	No
Surveyor1	12%(16)	20%(20)	14%(14)	47%(32)	65%(34)	0%(2)	119	38%(119)	Yes
Surveyor55	NA	0%(1)	8%(36)	31%(16)	25%(65)	0%(1)	119	20%(119)	Yes
Surveyor70	NA	NA	42%(43)	43%(7)	44%(66)	0%(2)	119	42%(119)	Yes
Surveyor68	NA	NA	39%(31)	33%(3)	38%(26)	46%(54)	114	42%(114)	Yes

Section 13.B. Heterogeneity of days since baseline by follow-up.

Figure 13.B.1. Distribution of follow-up survey timing by reported follow-up

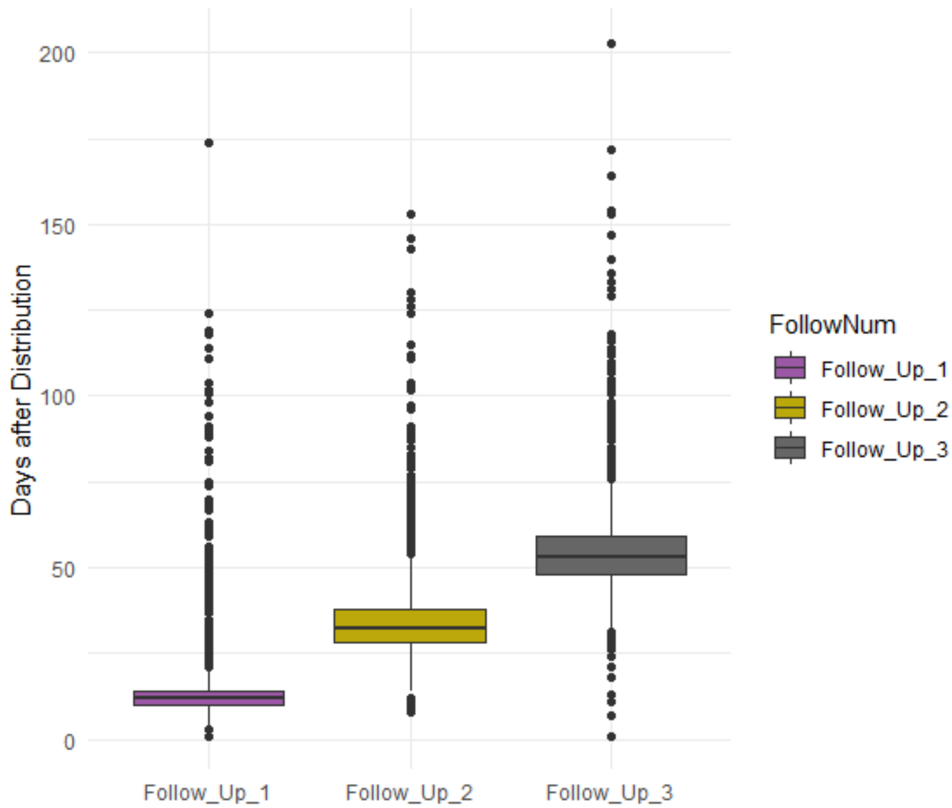


Figure 13.B.1. illustrates that while the majority of follow-up data occurred within a fairly narrow two-week window for each of the three scheduled follow-ups, there was still substantial variation in the timing of follow-up data collection. In order to deal with this observed variation, statistical models focus on days since follow-up instead of follow-up number.

Section 13.C. Water Quality Dashboards.

Individual water quality dashboards for water sources are included in a separate file.