WATER AND SANITATION PROGRAM: **TECHNICAL PAPER**

Global Scaling Up Handwashing Project

Improving Measures of Handwashing Behavior

Pavani K. Ram, Stephen P. Luby, Amal Krishna Halder, M. Sirajul Islam, Stewart Granger

May 2010

wsp
water and sanitation program

The Water and Sanitation Program is a multi-donor partnership administered by the World Bank to support poor people in obtaining affordable, safe, and sustainable access to water and sanitation services.

By Pavani K. Ram, MD

I sincerely thank the Bangladeshi respondents and their households who allowed our field team into their homes for hours in order to complete high-quality data collection. In many cases, participating households had to deal with flood-related consequences and still welcomed the field team. The study could not have been designed without invaluable thoughtful input from Steve Luby, and could not have been conducted without the enthusiasm and rigor introduced by Amal Krishna Halder, the study coordinator. Our field team consisted of a highly trained, dedicated set of individuals who were meticulous about acquiring high-quality data.

The field team included Khaleda Begum, Salim Akter, Iffat Sharmin, Mahfuza Islam, Tahmina Parvin, Salma Sultana, Nasrin Akter, Mansura Khanam, Sayeda Mahsina Akter, Halima Hawa, Ziaonnessa Begum, and Mahmuda Afroz Siddika. The following laboratory workers, led by Dr. Sirajul Islam, contributed zeal and diligence to the microbiology portion of the study: Iqbal Jahid, Debasish Paul, Shahanewaz Khan, Partho Sharathi Gope, and Iffat Tasmin Haque.

Benjamin Nygren, a master's degree student from Emory University, conscientiously developed the microbiology field protocols and provided useful input to other aspects of the methodology. Our colleagues from Unilever, Stewart Granger, Peter Hall, and Therese Jones, provided invaluable input into the deployment of the sensor products, analysis of the logger data, and review of other findings from the study. Members of the Public Private Partnership for Handwashing Steering Committee and the Monitoring and Evaluation Subcommittee provided substantial input into the thinking that led to this work. Mark Sobsey and Jack Colford greatly informed the study design, reviewed initial analysis with care, and provided substantial guidance along the way. Val Curtis and Orlando Hernandez kindly provided instruments to aid in the development of study tools. Jack Molyneaux asked many thought-provoking questions and provided much valuable guidance during the study design and analysis process. The study was graciously funded by the Water and Sanitation Program (WSP).

SANITATION FOR ALL

Global Scaling Up Handwashing is a Water and Sanitation Program (WSP) project focused on learning how to apply innovative promotional approaches to behavior change to generate widespread and sustained improvements in handwashing with soap at scale among women of reproductive age (ages 15–49) and primary school-aged children (ages 5–9). The project is being implemented by local and national governments with technical support from WSP. For more information, please visit www.wsp.org/scalinguphandwashing.

This Technical Paper is one in a series of knowledge products designed to showcase project findings, assessments, and lessons learned in the Global Scaling Up Sanitation Project. This paper is conceived as a work in progress to encourage the exchange of ideas about development issues. For more information please email Pavani Ram at wsp@worldbank.org org visit our website at www.wsp.org.

WSP is a multi-donor partnership created in 1978 and administered by the World Bank to support poor people in obtaining affordable, safe, and sustainable access to water and sanitation services. WSP's donors include Australia, Austria, Canada, Denmark, Finland, France, the Bill & Melinda Gates Foundation, Ireland, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom, United States, and the World Bank.

Executive Summary

KEY POINTS

- **This study suggests the potential utility of sensor soap, a novel technology for measuring soap use at the household level.**
- **Structured observation, while yielding detailed information on handwashing behavior, is subject to reactivity, as demonstrated by the use of motion sensors in bars of soap.**
- **Study findings call into question the validity of hand microbiology, given the great variability of results from the same person's hands over the course of just a few hours.**

Study Questions

As handwashing (HW) promotion is scaled up and adopted by organizations of varying research capacity, evaluation techniques that are more field-friendly, inexpensive, and readily adaptable are required. Based on our understanding of the available literature on measures of HW behavior, we posed the following questions:

1) How reactive is a subject, with respect to HW behavior, when observed during an extended structured observation?
   a. Does the collection of hand rinse samples for quantification of hand contamination result in handwashing reactivity?
2) What is the optimum duration of a structured observation for the purposes of measuring HW behavior?
3) What is the change in soap-use behavior during repeated structured observations within the same household?
4) Does hand contamination measured at random times predict hand contamination at times critical to pathogen transmission?
5) What is the rate of recontamination among the target population at two to three hours following a thorough handwashing?

Methods

We conducted a cross-sectional survey in six randomly selected rural areas in Bangladesh to address these questions. In 100 households, handwashing behavior was assessed using sensor soaps, which consist of motion sensors embedded in Lifebuoy soaps, structured observations of 90 minutes (N=50) or five hours (N=50) duration, spot checks, hand microbiology, and questionnaire administration.

Findings

The sensor soap enabled us to identify a subset of households in which the presence of an observer for structured observation resulted in substantial reactivity and, thus, increase in soap use. Reactivity was associated with socioeconomic and educational factors. Among abbreviated (90-minute) observation households, there was substantial loss of ability to observe handwashing behavior during defecation-related events and before the mother's eating. In a small group of nine households that had serial structured observations on three consecutive days, we did not find significant differences in soap-use behavior, as measured either by structured observation or by sensor soap, on the third day of serial observations, compared to the first day of serial observations. We found no correlations between levels of fecal coliform or *Escherichia coli* contamination of hands at random and hands at critical times. Moreover, we found large absolute differences between levels of contamination at random and at critical times, suggesting that microbiology testing of the same subject's hands can yield very variable results. We did find correlations between results of hand contamination of two critical time samples. Similarly, we found correlations between the results of the random hand rinse taken before the structured observation was begun and the hand rinse taken two hours after a supervised thorough handwashing. Despite these correlations, the absolute differences between the two random samples and the two critical time samples were actually quite large, again pointing to the lack of repeatability of hand microbiology testing. Following the thorough handwashing with soap by 25 respondents, we found fecal coliforms on all mothers' hands but noted that a sizable proportion (20%) of mothers did not have any detectable *E. coli* contamination.

Conclusions

We identified substantial reactivity to structured observation among a subset of households, with socioeconomic and education markers as potential explanatory markers for reactivity. This study suggests the potential utility of sensor soap, a novel technology for measuring soap use at the household level. Among those households that did not display reactivity, structured observation may provide useful information regarding context-specific soap use and other handwashing behaviors. Our findings call into question the validity of hand microbiology, given the great variability of results from the same person's hands over the course of just a few hours. Notably, we found little correlation between self-report, spot check, structured observation, or hand microbiology indicators of handwashing behavior, and sensor soap data on soap use.

Based on our findings, we have the following recommendations:

1) Evaluate utility and feasibility of sensor soap in large-scale research studies and evaluations of handwashing promotion programs.
2) Deploy sensor soaps among households participating in structured observation, in order to assess for reactivity. Among households found not to demonstrate reactivity, structured observation data should be analyzed for context-specific soap-use behaviors.
3) Confirm that knowledge regarding sensor soap construction, deployment, retrieval, and data download can be transferred speedily and accurately to in-country personnel in various countries.
4) Confirm our findings regarding the lack of correlation between sensor soap data on per capita daily soap use and other measures of handwashing behavior, spot-check measures in particular.
5) Examine consistency in soap use during different critical times for handwashing among households found not to be reactive to the presence of the observer.
6) Examine the relationship between hand contamination, following a supervised thorough handwashing with soap, to soap use, as measured by sensor soap, during the time between the supervised handwashing and the collection of the hand rinse samples.
7) Assess the relative impact on behavior, and if possible, health, of promoting handwashing with soap at specific critical times versus general promotion of handwashing with soap in order to increase overall soap use in a day.

Table of Contents

Tables

Figures

1. Background

KEY POINTS

- **Data on HW behavior may be observed, inferred, or reported. Observed data may be considered the most objective type of data available for measuring HW behavior.**
- **Structured observation provides detailed information about handwashing behavior but may be subject to reactivity.**

Handwashing (HW) efficacy studies have taken the form of randomized controlled trials and the outcome(s) of interest have primarily been health related.[1] That is, the intervention group—which received HW promotion—and the control group have been compared according to disease burden. As HW promotion is scaled up and adopted by organizations of varying research capacity, evaluation techniques that are more field-friendly, inexpensive, and readily adaptable are required.

As HW promotion is scaled up and adopted by organizations of varying research capacity, evaluation techniques that are more field-friendly, inexpensive, and readily adaptable are required.

Data on HW behavior may be observed, inferred, or reported. Observed data may be considered the most objective type of data available for measuring HW behavior. Observations may be conducted using spot checks or continuous observation.[2] Continuous structured observation is frequently used to evaluate HW behavior as well as other hygiene behaviors.[3] This technique consists of the placement of a trained observer in a target family's home for several key hours in the day. The observer can conduct continuous observation of opportunities for hygiene behaviors and track whether or not hygiene behaviors are practiced at those opportunities. Observers can also track the quality of observed hygiene behaviors, for example the use of soap or ash for washing hands.

Although the structured observation technique provides an opportunity for intensive and direct measurement of HW behaviors, it is subject to several limitations.

Although the structured observation technique provides an opportunity for intensive and direct measurement of HW behaviors, it is subject to several limitations. Most importantly, the validity of HW behavior data collected under such observation may be compromised, since the subject may alter her usual practices because she is being observed. Structured observation of water, sanitation, and hygiene behaviors variably generates "reactivity," or a change from usual practices, often dissipating after the subject has been observed over several sessions.[4] Moreover, repeatability, or the ability to detect the same behavior over multiple observation sessions, is challenged by the fact that individuals practice behaviors inconsistently.[5] For example, a subject may variably wash hands with soap after defecation; if observed only once, she may be observed to wash with soap and it may be assumed that she consistently washes hands with soap. However, on a second observation day, it may be noted that she does not wash hands with soap after defecation, reflecting the reality that the subject varies HW behavior after defecation. Since structured observation is so time-intensive, each observer may

[1] Luby, et al. 2004; Luby, et al. 2005.
[2] Ruel and Arimond 2002; Stanton, et al. 1987.
[3] Curtis, et al. 1993.
[4] Cousens, et al. 1996.
[5] Cousens, et al. 1996; Gorter, et al. 1998.

only be able to conduct one session with a single household per day of work. This intensity results in a high cost of collecting structured observation data, compared to collecting other types of data for measuring HW behavior. Thus, it is impractical to conduct repeated sessions with the same household(s), for the purposes of minimizing reactivity or ensuring repeatability, during the course of monitoring and evaluation for large-scale HW promotion programs. Therefore, evaluating the degree to which a single structured observation session generates reactivity would be helpful in placing data collected at the session in the context of usual HW behavior.

One approach to assessing the reactivity of structured observation is to compare that data to data from another type of direct observation of HW behavior, or at least soap use.

One approach to assessing the reactivity of structured observation is to compare that data to data from another type of direct observation of HW behavior, or at least soap use. Unilever, a multinational soap manufacturer, has developed a technology that allows for continuous monitoring of the vertical and horizontal movement of a bar of soap.[6a] A motion sensor is embedded in an otherwise normal appearing bar of Lifebuoy soap. The sensor soap can be given to subject households to use over a week's time, and then collected for data download and analysis at the end of the week of "observation." The sensor soap too may be subject to reactivity. However, according to colleagues at Unilever, their data suggest that the reactivity decreases after a couple of days and HW behavior "normalizes."[6b] If the sensor soap could be given to a household several days prior to an extended structured observation, it would be possible to compare HW behavior during the period when the household is being "observed" by the sensor soap alone to the period when the household is under extended structured observation. This correlation would help clarify the degree to which extended structured observation results in reactivity, or increased use of soap for HW. The sensor soap has been paired with a motion sensor-embedded water vessel, called a sensor *bodna*. Such water vessels, known as *bodna* in Bangladesh, are typically carried by an individual to the defecation site for washing the anus and, sometimes, for washing hands. The data from the sensor *bodna* allows for the timing of defecation events. Once a defecation event is timed, the data from the sensor soap can be reviewed to examine whether soap was used for handwashing during the several minutes following the defecation event.

In the published literature on structured observations, the duration of observation ranges from at least three hours up to seven hours and the observations are collected at the same time of day with each respondent.[7] As noted above, these prolonged periods, and the opportunity to conduct only one observation by an observer each day, add to the cost of the structured observation. We are unable to locate published data in the hand hygiene literature on the optimum duration of structured observation. Presumably, a prolonged duration allows for a greater number of observations on specific HW opportunities, such as after defecation or before preparing food, for a single subject. If the unit of analysis were HW with

[6a, 6b] W. Gibson and S. Granger, Unilever, personal communication.
[7] Stanton et al. 1987; Cousens et al. 1996; Gorter et al. 1998; Manun'Ebo, et al. 1997.

soap during any critical time, rather than HW with soap during a specific critical time (for example, after defecation), then shorter observation periods may facilitate data collection with a larger number of subjects during the study period. For example, if an observer can witness HW behavior during at least one critical time within a 90-minute observation period, she can conduct observations on three or four subjects during a single day.

Compared to observed behaviors, such as those documented during structured observations, inferred and reported data have different strengths and weaknesses with respect to HW behavior measurement. Neither of these data types are direct measures of HW behaviors. Still, inferred and reported data on HW behaviors are often used because they are easier to collect as well as less costly and time-intensive than structured observations. Self-reported information on HW practices is subject to information and courtesy biases. Self-reported HW behavior has been shown repeatedly to overestimate true behavior.[8] Spot-check observations are an inefficient way of directly observing HW behavior at specific critical times, since large numbers of spot checks are necessary to capture sufficient numbers of opportunities to observe HW behavior at critical times.[9] Instead, spot checks may be more useful to obtain information, such as the presence of soap and water in the kitchen or the ability of a mother to display correct techniques when asked to demonstrate HW, from which HW behavior may be inferred.[10] Creating links between inferred data and true HW behaviors requires the incorporation of potentially invalid assumptions.

Compared to observed behaviors, such as those documented during structured observations, inferred and reported data have different strengths and weaknesses with respect to HW behavior measurement.

Testing for the presence of and/or quantification of fecal coliforms, including *E. coli*, on a subject's hands using microbiologic techniques may be considered a source of inferred data on HW behavior.[11] Reduced hand contamination has been demonstrated among persons exposed to HW promotion or persons specifically instructed to wash hands with a cleansing agent, compared to persons who were not similarly exposed.[12] Although the data are variable, at least one study has demonstrated the correlation between reduced hand contamination and reduced diarrhea risk.[13] However, hand microbiology may be limited as a source of information about HW behavior because hands become recontaminated quickly after washing or because the subject washed hands with soap immediately preceding the sample collection.[14]

Presumably, the level of hand contamination at critical times, such as preparing food or feeding a young child, impacts the degree to which pathogens are transmitted at those times. In terms of measuring hand contamination that is relevant to pathogen

[8] Stanton, et al. 1987; Manun'Ebo, et al. 1997.
[9] Strina, et al. 2003.
[10] Ruel and Arimond 2002.
[11] Kaltenthaler and Pinfold 1995.
[12] Hoque, et al. 1995; Luby, et al. 2001; Pinfold 1990.
[13] Pinfold and Horan 1996; Luby, et al. 2007.
[14] Sobel, et al. 1998.

If hand contamination, as detected in random hand rinse samples, were indeed predictive of hand contamination at critical times, this information would bolster attempts to use and refine random sample hand microbiology in HW monitoring.

transmission, this would necessitate the presence of an observer at critical times in the household in order to collect hand rinse samples at those times. This would entail some duration of structured observation, particularly if specific critical times for HW are sought, rather than any critical times for HW. If hand contamination measured at random times could predict hand contamination at times critical for pathogen transmission, the required duration of structured observation could be shortened, thereby reducing personnel requirements and costs. Therefore, we examined whether hand contamination, as detected on random collections of hand rinse samples (similar to collections done during spot checks), can predict hand contamination at critical times, such as before food preparation.

If hand contamination, as detected in random hand rinse samples, were indeed predictive of hand contamination at critical times, this information would bolster attempts to use and refine random sample hand microbiology in HW monitoring. If hand contamination, as detected in random hand rinse samples, were *not* predictive of hand contamination at critical times, this may explain the variability described in previous studies of the association between hand contamination and HW behavior. The collection of hand rinse samples may itself result in reactivity and, thus, enhanced handwashing behavior during the course of an extended structured observation. Since hand contamination represents a potentially useful and objective measure of handwashing behavior, we examined whether the collection of the hand rinse sample itself results in increased handwashing with soap during the course of the observation.

Based on our understanding of the available literature on measures of HW behavior, we posed the following questions:

1) How reactive is a subject, with respect to HW behavior, when observed during an extended structured observation?
    a. Does the collection of hand rinse samples for quantification of hand contamination result in handwashing reactivity?
2) What is the optimum duration of a structured observation for the purposes of measuring HW behavior?
3) What is the change in soap-use behavior during repeated structured observations within the same household?
4) Does hand contamination measured at random times predict hand contamination at times critical to pathogen transmission?
5) What is the rate of recontamination among the target population at two to three hours following a thorough handwashing?

We conducted a cross-sectional survey to address these questions.

To address Question 1, we conducted extended structured observations (~five-hour duration) in a set of households that were given the sensor soap four days in advance of the observer's visit. The degree of reactivity to the observer's presence was determined by comparing soap use during the several days preceding the observer's visit to soap use during the extended structured observation. To address Question 1a, we assessed whether there was an association between hand rinse sampling and reactivity to structured observation.

To address Question 2, we compared the opportunities to observe HW behavior during extended structured observations to opportunities to observe HW behavior during abbreviated structured observations.

To address the Question 3, we examined soap-use behavior, as measured by structured observation and by sensor soap, over the course of three consecutive days in a subset of households undergoing extended structured observation.

To address Question 4, we compared the contamination on hands of mothers of young children at random times, such as upon arrival of the observer in the home, to contamination detected at critical times during structured observation in the same household on the same day, such as when the mother was preparing food, handling water for storage, or feeding the child.

To estimate the rate of recontamination (Question 5), we requested respondents in the abbreviated structured observation group to wash their hands thoroughly with soap and then conducted hand rinses approximately two hours following that thorough hand wash.

2. Study Design and Methods

KEY POINTS

- **The study aimed to answer specific questions on measuring handwashing behavior change.**
- **Handwashing behavior was measured using structured observations, sensor soap, hand microbiology, spot checks, and questionnaires.**

We conducted this project in the context of the large-scale hygiene promotion and sanitation program (SHEWA-B) that is being undertaken by the Department of Public Health Engineering of the Government of Bangladesh (DPHE-GoB) and UNICEF-Bangladesh to impact the water, sanitation, and hygiene conditions of 30 million residents across the country. In addition, DPHE and UNICEF have already been conducting similar hygiene and sanitation promotion programs in some rural areas that are adjacent to SHEWA-B areas but that are not directly impacted by the new project. For this work, we collected data in both the *new* and old project areas. We conducted fieldwork in the new project areas since those households that had not yet been exposed to intensive hand hygiene promotion represent baseline populations. Accordingly, those households in the old areas may have already been exposed and, thus, represent follow-up populations. ICDDR,B is the lead agency responsible for M&E of the SHEWA-B project over the next three years.

Whereas the larger M&E activities for the SHEWA-B are taking place in 50 randomly selected communities across the country, this study was conducted in six randomly selected rural areas in Bangladesh. All six communities were located in Sirajganj and Brahmanbaria districts, which contain both new and old project areas. We listed out all of the unions in the new project areas located within a two to three hour driving distance from Dhaka. From among these, three clusters in Sirajganj district were selected based on population-proportionate-to-size techniques. The three clusters in Brahmanbaria district were selected at random without concern for population size, since Brahmanbaria was affected by flooding during the course of data collection. All Brahmanbaria unions not affected by flooding were alphabetized and then numbered sequentially. The fifth union was selected as the starting point for the cluster. After selecting the six unions, we alphabetized all of the primary schools and the fifth primary school in the alphabetical order was selected from each union. These six primary schools represented the starting points from which respondent selection began (detailed below under Recruitment).

Participants

In the eligible household, the child under two years old was considered the index child. If there was more than one child under two years old, the youngest child was considered the index child. The primary target respondent for structured observation, spot checks, and questionnaire administration, was the primary caregiver of the child less than two years old (usually the mother). The primary caregiver of the child less than two years old was selected as the primary target respondent since she is the closest contact of the child. Previous studies have demonstrated that improving the hand hygiene of the mother can reduce illness in the child.[15]

Eligibility Criteria: Households were considered eligible if the following criteria were met:

- Child less than two years old living in the household, and
- Primary caregiver of the child available gave informed consent.

Study Groups

- *Extended observation group:* We intended to assess a total of 50 households using a five-hour structured observation, spot checks, hand microbiology, and questionnaire administration. Sensor soap, sensor *bodnas*. Field workers collected microbiological hand rinse samples upon arrival for the five-hour observation and at critical times during the course of observation in 25 households.
- *Abbreviated observation group:* A total of 50 households were to be assessed using a 90-minute structured observation, spot checks, and questionnaire

[15] Luby, et al. 2004; Luby, et al. 2005.

administration. Field workers collected microbiological hand rinse samples upon arrival and at critical times during the course of observation in 25 households. They concluded by having respondents perform a thorough handwashing with soap and by collecting hand rinse samples again two hours after the handwashing in these 25 households.

Recruitment

The field research officer visited the village in which the randomly selected primary school was located. S/he worked with the village head to enumerate all the *baris* (multi-household compounds) in the village. All odd-numbered *baris* were approached for potential inclusion in the abbreviated structured observation group, based on the eligibility criteria. All even-numbered *baris* were approached for potential inclusion of in the extended structured observation, based on the eligibility criteria. There was a maximum of one respondent per *bari*.

Abbreviated Structured Observation

Interviewers approached each odd-numbered *bari* in the village. In each odd-numbered *bari*, the interviewer identified all children younger than two years old in the *bari*. If only one household in the *bari* had a child <2 years old, the mother of that child was approached for enrollment.

If more than one household in the *bari* had a child younger than years old, the team member identified all mothers of children younger than two years old in the *bari*. The interviewer requested all of the mothers of the children younger than two to stand in front of their respective homes. Starting from her right, she counted the mothers from one to five and requested the mother who was counted as "five" to participate in the study. For example, let us assume that there were three children younger than two years old in a given *bari*. The mothers of those three children will stand in front of their respective homes. Mother A's household occupied the first home from the right, Mother B's household occupied the second home from the right, and Mother C's household occupied the third home from the right. Mother A was counted as "one." Mother B was counted as "two." Mother C was counted as "three." Mother A was then counted as "four." Mother B was then counted as "five." The interviewer requested Mother B to participate in the study.

If Mother B was unavailable for observation or refused to participate, the interviewer returned to counting from 1 to 5 with Mother A and Mother C. In the scenario where Mother B was unable or unwilling to participate, Mother A was counted as "five" and was approached for participation.

One household per *bari* was enrolled in the abbreviated structured observation group. When one participant had been enrolled, or if no eligible participants could be identified, the *bari* was considered "exhausted." Once a *bari* was "exhausted," the interviewer then proceeded to the next *bari*. There, she repeated the process of identifying all households with children younger than two years old in the *bari*, counting those households off from one to five if necessary, and approached the mother of household "five" for enrollment.

Extended Structured Observation

Each even-numbered *bari* was approached to identify potential respondents. The interviewer identified all children younger than two years old in the *bari*. If only one household in the *bari* had a child younger than two years old, the mother of that child was enrolled. If more than one household in the *bari* had a child younger than two years old, the team member selected the appropriate mother for participation based on the same protocol used in *baris* eligible for inclusion in the abbreviated observation group above.

One household per *bari* was enrolled in the extended structured observation group. When one participant had been enrolled or if no eligible participants could be identified, the *bari* was considered "exhausted." Once a *bari* was "exhausted" the interviewer then proceeded to the next *bari*. There, she repeated the process of identifying all households with children younger than two years old in the *bari*, counting those households off from one to five if necessary, and approached the mother of household "five" for enrollment.

Consent Process

Participation in the study was voluntary. This study involved the collection of information related to hygiene, water, and sanitation practices in the household. These questions are generally not perceived to be of a sensitive nature in Bangladeshi culture. All potential participants were informed about the purposes and intent of the study, and the voluntary nature of their participation. Written informed consent was obtained from each participating primary caregiver. A bilingual study staff person with a graduate degree translated the consent form into Bengali. The interviewer read the consent form aloud since the literacy rate among Bangladeshis remains very low. The potential participant's questions and concerns were addressed. If a potential participant agreed to take part, she signed the consent form or provided her thumbprint in lieu of signature.

Research Procedures

Abbreviated Structured Observation

A total of 50 respondents were to be enrolled in the abbreviated structured observation group. Once an eligible respondent had been identified to take part in the abbreviated structured observation, the trained interviewer explained the particulars of the study and obtained written informed consent. She requested permission to stay in the *bari* or household for the subsequent two hours. She proceeded with the 90-minute structured observation and then administered the questionnaire to the primary respondent. Upon conclusion of the questionnaire administration, the interviewer conducted spot checks, ending with a request for the mother to demonstrate usual handwashing behavior.

A total of 25 from the abbreviated structured observation group were included in the hand microbiology component of the study. In these 25 households, the interviewer collected a random hand rinse sample immediately after receiving written consent and then proceeded with the 90-minute structured observation During the course of the 90-minute observation, when the interviewer witnessed a critical time for pathogen transmission (specifically, preparing food, eating food, feeding a child, or handling water for storage), she collected a hand rinse sample immediately preceding such activity. Only one critical time hand rinse sample was collected in the 25 abbreviated observation households included in the hand microbiology component of the study.

Upon completion of the 90-minute observation, in these same 25 households, the interviewer collected another hand rinse sample and asked the respondent to wash her hands thoroughly with soap. The interviewer returned to the household hours later in order to collect a final hand rinse sample.

Extended Structured Observation

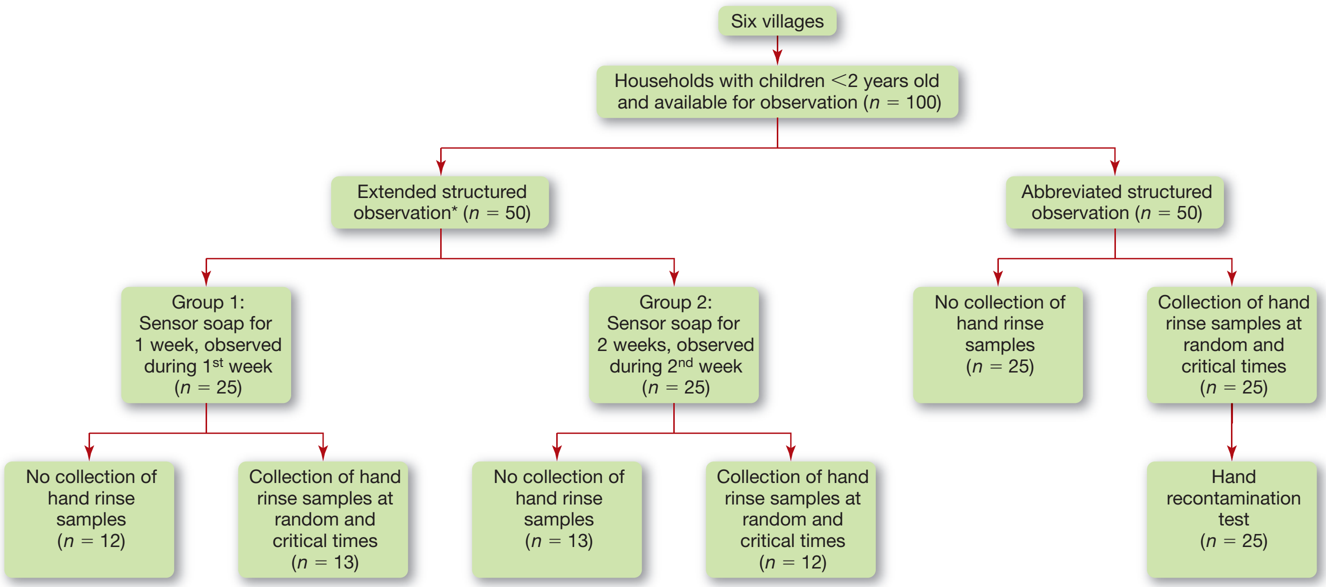
Once an eligible respondent had been identified to take part in the extended structured observation, the interviewer explained the particulars of the study, including information about the sensor soap and the sensor *bodna*, and obtained written informed consent. Once written informed consent was obtained, the interviewer administered the questionnaire to the primary respondent and conducted the spot checks. She then provided a bar of sensor soap to the subject in exchange for the bar of soap used in the home or at the toilet. If two soaps were used anywhere in the household, two sensor soaps were given in exchange for both of the household's soaps. She also provided a sensor *bodna* in exchange for the household's own *bodna*, the water vessel used by the family for cleansing after defecation.

The five-hour structured observations were conducted in 25 households on the fourth full day after the sensor soap and the sensor *bodna* were delivered (Figure 1). The sensor soap and the sensor *bodna* were collected upon conclusion of the five-hour structured observation. In addition, the interviewer asked to collect whatever cloth or other item was being used regularly to clean surfaces around the home, especially in the kitchen.

In another 25 households, the sensor soap was delivered at the time that consent was obtained and retrieved four full days later (Figure 1). A second sensor soap was provided at that time and the five-hour structured observations were conducted four full days after the second sensor soap was delivered, i.e., eight full days after the initial introduction of the sensor soap. Within this group, we selected nine houses at random for completion of structured observations on each of three consecutive days in order to address the question on change in soap-use behavior during repeated structured observations within the same household.

The sensor soaps and sensor *bodnas* were retrieved at the conclusion of the single five-hour structured observation among households undergoing observation on a single day

FIGURE 1: FLOWCHART OF STUDY GROUPS



* The culture of the youngest child's hands will take place in all 50 households in the extended structured observation group.

or the third structured observation among those households undergoing observation on three consecutive days. Then, the interviewer asked the mother to assist in doing a hand rinse for the youngest child.

A total of 25 respondents from the extended structured observation group were included in the hand microbiology component of the study. From 12 respondents observed on the fourth full day after sensor soap placement and 13 respondents observed on the eighth full day after sensor soap placement, a hand rinse sample was obtained immediately upon the interviewer's arrival for the five-hour structured observation. This hand rinse sample will hereafter be referred to as the "random" sample. Thereafter, during the course of the five-hour structured observation, the interviewer watched for critical times for pathogen transmission (specifically, preparing food, eating food, feeding a child, or handling water for storage) and collected a hand rinse sample immediately preceding any of these activities. The interviewer collected a total of two critical time hand rinses per household. In the event that two critical times could not be observed, the interviewer collected a final hand rinse sample at the conclusion of the five-hour observation period.

Description of Data Sources for Measuring HW Behavior

Structured Observations

Continuous observations of a household's HW behavior conducted over several hours, and often repeatedly, have been utilized frequently as part of HW studies and evaluations.[16] The observer arrived at the household and asked the mother where she could spend much of her time for the subsequent 90 minutes or five hours, as appropriate. The observer identified a location for herself such that she was as unobtrusive as possible. There, she was expected to quietly observe the mother's activities and note down all opportunities for HW and the mother's practice at those opportunities. The observer was instructed to maintain the primary respondent as her focus of observation. As other household members engaged in opportunities for handwashing in the full view of the observer, she was expected to note down their specific handwashing practices in those instances.

[16] Bentlye, et al. 1994.

The following indicators of HW behavior were measured by structured observations:

- Percent of caregivers observed to wash hands with soap at any of the critical times
- Percent of caregivers observed to wash hands with soap after defecating/using the toilet
- Percent of caregivers observed to wash hands with soap after cleaning child who has defecated or used the toilet
- Percent of caregivers observed to wash hands with soap before preparing food
- Percent of caregivers observed to wash hands with soap before feeding child
- Percent of caregivers observed to wash hands with soap before eating
- Percent of caregivers observed to wash hands with soap before handling food for storage, and
- Percent of caregivers observed to wash hands with soap before handling water for storage.

Continuous Recording of Soap Use

Respondents from all of the households in the extended structured observations group were requested to use sensor soaps for one week or two weeks, as noted above. Sensor soap, a technology developed by Unilever, consists of a motion sensor embedded in a bar of Lifebuoy soap. The motion sensor allows detection of vertical and horizontal movements of the soap. By analyzing the motion patterns, it is possible to detect individual episodes of soap use. The soap itself is the exact formulation of commercially available Lifebuoy soap. Each sensor soap lasts approximately five to seven days for an average family of four or five people. After that time, the logger is usually visible and the sensor soap should be retrieved.

The consent form provided details and instructions on the use of the sensor soap. Specifically, the interviewer explained that the "special" Lifebuoy soap being given to the household is similar to Lifebuoy soap available in the market. It should be used as soap is typically used in the home. The "special" Lifebuoy soap collects information on soap use. The soap itself is the same and provides the same benefits as regular Lifebuoy soap. People who are sensitive to Lifebuoy soap or who experience skin reactions to Lifebuoy soap will experience similar reactions to the "special" Lifebuoy soap and should not use it. The "special" Lifebuoy soap was to be collected from the home within five to six days. In some households, a second bar of "special" Lifebuoy soap was given.

Soap use was tracked in all of the households undergoing extended structured observations via the sensor soap. The sensor soap was delivered four days before the extended structured observation was to take place.

The following indicator was measured using continuous recording of soap use:

- Number of soap-use episodes per household member per period time.

By pairing the sensor soap with the sensor *bodna*, the following indicator was also measured:

- Percent of defecation events following which hands are washed with the sensor soap.

Hand Microbiology

A total of 50 respondents, 25 each in the extended and abbreviated structured observation groups, were requested to provide hand rinse samples for microbiologic testing for the detection of fecal coliforms and *E. coli*. In addition, one hand rinse sample was collected from the youngest child in the extended structured observation group households. Hand rinses, including rinse of the palm and fingertips, were collected by immersing the entire hand in Whirl-Pak bags containing 200 ml of recovery medium. Both hands were rinsed, in series, in the same Whirl-Pak bag.

Hand rinse samples were collected from primary respondents in the selected households immediately before the structured observation and at critical times. Random hand rinse collections were defined as collections performed immediately before the structured observation. Among those households that did not have the requisite number of critical times available for critical time hand rinse sampling, an additional hand rinse was collected at the conclusion of the structured observation. Here, critical times were defined as before eating, before feeding a child, before preparing food, and before handling water for storage.

In the 25 households from which hand rinse samples were collected in the abbreviated structured observation group, the respondent was requested to wash her hands with soap following the conclusion of the structured observation. The interviewer asked to return to the household within two to three hours to collect a final hand rinse sample in order to estimate the rate of recontamination.

The Whirl-Pak bag containing the hand rinse was labeled and placed in a cool box to be delivered to the ICDDR,B water microbiology laboratory at the end of each data collection day. The 200 ml of hand rinse water was processed using standard membrane filtration techniques, with the goal of quantifying the number of colony forming units of fecal coliforms and *E. coli* in each hand rinse sample.

The following indicators were measured using hand microbiology:

- Number CFU/100 ml of fecal coliforms and *E. coli* on caregivers' hands upon interviewer arrival (random)
- Number CFU/100 ml of fecal coliforms and *E. coli* on caregivers' hands in samples obtained just before feeding child, preparing food, or handling water
- Number CFU/100 ml of fecal coliforms and *E. coli* on child's hands, and
- Percent of caregivers whose hands were found to be recontaminated within two hours after washing hands with soap.

Spot Checks

The interviewer made observations of structural elements around the home related to HW behavior, including the presence of a dedicated HW station, the distance from the latrine and food preparation area to a place where soap and/or water are present.

The following indicators were measured using spot checks:

- Percent of caregivers who use soap when asked to demonstrate HW
- Percent of caregivers who wash both hands when asked to demonstrate HW
- Percent of caregivers who air dry or use a clean towel to dry hands after HW when asked to demonstrate HW
- Presence of soap and water at the same location for HW purposes (HW station)
- Presence of HW station in/near toilet, and
- Presence of HW station in/near kitchen.

Questionnaire Administration

A standardized questionnaire was designed to capture self-reported measures of HW behavior and use of soap or other cleansing agents, such as ash. The primary respondent for the questionnaire was the mother of the child under five years old. The questionnaire consisted of closed-ended and open-ended questions with pre-coded responses.

The following indicators were measured using questionnaire administration:

- Percent of caregivers who report washing hands with soap
- Percent of caregivers who report washing hands with soap at critical times, and
- Percent of caregivers who report washing hands with soap during a majority of times that hands are washed.

Sample Size Calculations

The power calculations were based on the research questions of greatest interest.

First, in order to determine the necessary sample size to assess reactivity of subjects to extended (five-hour) structured observation, we used one-sided paired t-test calculations. We used one-sided calculations since we assumed that there would be reactivity only in the positive direction, i.e., that we would o*nly noti*ce increases in soap use among observed individuals rather than a mix of increases and decreases in soap use.

We performed power calculations in order to assess the detectable differences based on different degrees of correlation (ρ) between the number of HW events during the structured observation five-hour block and the number of HW events during non-observation blocks. Box 1 shows the detectable difference based on various degrees of correlation, which in turn affected the standard deviation of the difference between groups. This is assuming that we held the sample size of 50 households constant.

Assuming that we could distribute sensor soap to and conduct extended structured observations in 50 households, we estimated that we would be able to detect a difference of 0.32 HW events between the structured observation five-hour block and the non-observation five-hour blocks with 90% power at the .05 significance level.

BOX 1: DETECTABLE DIFFERENCES BASED ON DEGREES OF CORRELATION

Degree of Correlation (ρ)	Standard Deviation of Difference Between Groups (σ_{diff})	Detectable Difference (# of Soap-Use Events)
.25	.92	.39
.5	.75	.32
.75	.53	.22
.9	.34	.14

Second, in order to address the question of the optimum duration of structured observation, we used two-sided two-sample t-test calculations to model the equivalence of opportunities to witness HW at critical times between the extended and abbreviated structured observation groups. We assumed that, in the extended (five-hour) structured observation group, we would be able to witness approximately one opportunity for HW during each hour of observation. We also assumed that we would conduct the five-hour extended structured observation in 50 households (as described in the first power calculation above) and the 90-minute abbreviated structured observation in 50 households. This sample size would allow us to demonstrate equivalence between the extended (five-hour) and abbreviated (90-minute) observation groups at the 0.33 level with 90% power and 95% significance. That is, with 50 households in the extended observation group and 50 households in the abbreviated observation group, we expected to have sufficient power to demonstrate that the two groups were equivalent if the number of HW opportunities witnessed per hour ranged from 0.67 to 1.33 in the abbreviated structured observation group.

Third, in order to address the question of whether the degree of hand contamination, as measured during random hand rinse collections, predicted hand contamination at critical times, we used linear regression. We intended to assess the degree of correlation in the # CFU/100 ml of *E. coli* between the random hand rinse samples and the critical time hand rinse samples. If we collected one random and one critical hand rinse sample each from a total of 50 households, we estimated that we would be able to detect a correlation coefficient (ρ) ≥ 0.44 (R^2 of ≥19%) with 90% power at the .05 significance level.

Fourth, in order to address the question of whether the collection of hand rinse samples results in reactivity, or increased soap-use events among extended structured observation households, we used two-sided paired t-test calculations. Here, we used a two-sided test since we could not know whether mothers would increase handwashing events because of their perception that the entire exercise is focused on handwashing or whether mothers would decrease handwashing events because of the perception that the hand rinse sampling was, indeed, a form of handwashing and, thus, that usual handwashing behavior was not as necessary. We assumed that the degree of correlation (ρ) between the number of HW events in the hand rinse households and the number of HW events in the non-hand rinse households would be approximately 0.5. Assuming that we would collect hand rinses in 25 extended structured observation households, we estimated that we would be able to detect a difference of 0.5 HW events with 90% power at the .05 significance level.

Given these calculations related to the research questions of greatest interest, we proposed to enroll 50 households in the extended structured observations group and 50 households in the abbreviated structured observations group.

We present here power calculations for the additional research questions, although these analyses were intended to be exploratory in nature.

First, to assess the change in reactivity of subjects as a result of repeated structured observations within the same household, we used paired two-sided t-test calculations. We assumed that the mean number of soap-use events on the first day of structured observation would be 2.0 and that the mean number of soap-use events on the third day of structured observation would be 2.5. If we assume that the standard deviation of the difference between the means is 0.75, to demonstrate this difference in the means with 80% power at the .05 significance level, we would need to perform repeated observations in 36 households. Given the various project logistics, we were able to conduct these repeated observations in nine households, which allowed us to demonstrate a difference in mean soap events of 0.5 with a power of 26%.

Second, to assess the rate of recontamination following a handwashing with soap, we estimated that 50% of respondents would have recontaminated hands two hours following a thorough hand wash with soap. To detect this rate of recontamination, ±5% error, at the .05 significance level, among the 25 households providing hand rinses in the abbreviated structured observations group, we would have needed to examine hand recontamination in 23 respondents. Therefore, we conducted this exercise in all 25 respondents from the abbreviated structured observation group that were participating in the hand microbiology portion of the study.

Data Analysis

As presented above, our study questions were the following:

1) How reactive is a subject, with respect to HW behavior, when observed during an extended structured observation?
   a. Does the collection of hand rinse samples for quantification of hand contamination result in handwashing reactivity?
2) What is the optimum duration of a structured observation for the purposes of measuring HW behavior?
3) What is the change in soap-use behavior during repeated structured observations within the same household?
4) Does hand contamination measured at random times predict hand contamination at times critical to pathogen transmission?
5) What is the rate of recontamination among the target population at two to three hours following a thorough handwashing?

To address the first question on reactivity of a subject when observed during an extended structured observation, the primary outcome measure is the number of soap-use events recorded by the sensor soap during a five-hour period. Reactivity of structured observations was assessed by examining the difference in means of the soap-use events recorded during the five-hour structured observation, compared to the same five-hour time block during the preceding days on which sensor soap was in the home but when the observer was not in place. We used two-sided paired t-tests to determine the significance of the difference. We quantified reactivity by calculating the absolute difference between the number of soap-use events during the SO time block on the day of the structured observation to the mean number of soap-use events during the SO time block on pre-observation days. We called those in the highest quartile of the absolute difference in the number of soap-use events "reactors" and those in the lower three quartiles as "non-reactors." We examined predictors of reactivity by comparing reactors to non-reactors with respect to knowledge and attitudes, as well as SES strata and participation in the hand microbiology.

To address the second question on the optimum duration of a structured observation for the purposes of measuring HW behavior, we considered the outcome measures to be whether or not the interviewer was able to observe each of the handwashing opportunities of interest in the abbreviated and extended observation groups and used Mantel-Haenszel chi-squares to identify statistically significant differences. Opportunities of interest were defined as follows. Defecation-related events included the respondent's own defecation or other toileting event, as well as cleaning a child who had defecated, or changing the child's diaper. Water-related events included collecting water for storage and retrieving water from a storage container. Events such as preparing, cutting, or serving food were considered food-related. Eating food, feeding a child, and returning to the compound were each separate categories of events. We also examined the difference between extended and abbreviated observation groups with respect to each category of opportunities to observe HW behavior during the structured observation, using two-sided t-tests or the Wilcoxon rank sums method as appropriate, based on the distribution of the variables' responses.

To address the question on change in soap-use behavior during repeated structured observations within the same household, the primary outcome measure was the number of per capita soap-use events recorded by the sensor soap

[17] Rusin, Orosz-Coughlin, and Gerba 1998.

during the five-hour observation period. We assessed differences in mean soap-use events between the first and second day of the serial observations, and between the first and third days of serial observations. Paired t-tests were used to determine the significance of the difference.

To address the question on whether contamination measured during random hand rinse samples predicts contamination at critical times, we estimated the correlation between the degree of contamination in hand rinse samples collected upon arrival (as measured by # CFU/100 ml of fecal coliforms and *E. coli*), considered random samples, and the degree of contamination in hand rinse samples collected at critical times (as measured by the # CFU/100 ml of fecal coliforms and *E. coli* detected in hand rinse samples collected at critical times during the five-hour observation within a particular household). For the purposes of this analysis, the following HW opportunities were considered critical times: before feeding a child, before preparing food, and before handling water for storage. Additionally, we calculated the mean of the absolute differences between random and critical time results and construct Bland-Altman plots to assess whether differences in contamination measured at random and critical times vary according to the mean level of contamination.[18]

To address the question on the rate of recontamination among the target population, we calculated the frequency of contamination, defined as the presence of fecal coliforms or *E. coli* > 0 CFU/100 ml in the hand rinse sample that is collected two to three hours following a thorough handwashing with soap. In addition, we calculated the arithmetic and geometric means and examined the standard deviation. To understand the degree of recontamination, we present below the cut-off points for quartiles within the subgroup that underwent the recontamination assessment.

Additional Data Analysis

Our colleagues at Unilever have examined sensor soap data by calculating soap-use events per hour in order to examine patterns in soap use throughout the day during the days when an observer is NOT present in the household. This provides us with information on patterns of soap use for handwashing and body washing in Bangladesh, which may facilitate scheduling of abbreviated structured observations in the future. This data has been presented separately by Unilever and is not included in the report below.

Analysis of self-reported data obtained via interview was mainly descriptive. As appropriate, bivariate analysis using chi-squares was performed to identify associations between explanatory variables and HW behavior, as measured by the per capita 24-hour soap-use events detected by the sensor soap during the pre-observation days.

Since socioeconomic status (SES) may impact upon hygiene, water, and sanitation conditions within a household, we described the SES of each participating household and examined its association with soap use, as well as its effect on reactivity. Principal components analysis was used to categorize households into SES strata.

Ethical Assurance for Protection of Human Rights

Risks: The level of risk encountered by subjects in this study was no greater than minimal. Our research procedures included structured observations, hand microbiology, spot checks, and questionnaire administration. These all represented routine data sources for measuring hygiene, water, and sanitation behaviors in both research studies and in program evaluations. The novel data sources implemented in this study were the sensor soap and sensor *bodna*, which also posed no greater than minimal risk to the participant. The sensor soap and the sensor *bodna* each use a 3-volt button battery (Lithium type CR1225), commonly used in watches, cameras, and calculators. The battery is securely encased within the sensor *bodna* and does not pose any risk of electric shock. The risk of electric shock to someone using the sensor soap is considered to be extremely low. In the very unlikely event that an individual touches the 3-volt battery, it is expected that this is below perception threshold and will not cause pain. Also the cap of the sensor is made to be waterproof to a depth of 30 cm. The battery is located within the casing, which also minimizes the risk of a subject's exposure to the battery. The actual content of the soap, which protects against diarrhea and

[18] Bland and Altman 1986.

acute respiratory infection, is the same as that found in commercially available Lifebuoy soap. The motion sensor itself is comprised of components similar to those found in cellular telephones, which are widely used in Bangladesh and indeed worldwide, and thus, we did not anticipate that any risk resulted directly from the "exposure" to the motion sensor. As with any other items that contain batteries or with items such as soap, the sensor soap was to be kept away from children who might put it into their mouths.

Benefits: Participants in the extended structured observation group received the small direct benefit of a toy that was given to their child. Also, they received a new bar of soap to replace the sensor soap that was retrieved from the home. Finally, they received a new cleaning cloth.

Adverse Events: The study involved the interview and observation of caregivers of young children. The observational and interview techniques were non-invasive and we did not anticipate or experience adverse events related to study participation. As indicated in the consent form, persons who typically experience skin reactions to regular Lifebuoy soap would likely experience similar reactions to the sensor soap. They were instructed not to use the sensor soap or any other Lifebuoy soap.

Consent: Participation in the study was voluntary. This study involved the collection of information related to hygiene, water, and sanitation practices in the household. These questions are generally not perceived to be of a sensitive nature in Bangladeshi culture. All potential participants were informed about the purposes and intent of the study, and the voluntary nature of their participation. Written informed consent was obtained from each eligible primary caregiver, identified through the recruitment process. A bilingual study staff person with a graduate degree translated the consent form into Bangla. The interviewer read the consent form aloud since the literacy rate among Bangladeshis remains very low. The potential participant's questions and concerns were addressed. If a potential participant agreed to take part, she signed the consent form or provided her thumbprint in lieu of signature.

Scientific and Ethical Review: The protocol received approval by the Research Review Committee and the Ethical Review Committee of ICDDR,B, and also the IRB of the University at Buffalo before its initiation.

Confidentiality: The consent form contained the participant's full name and Unique ID number. The questionnaire, structured observation form, and spot checks form will have the participant's first name and Unique ID number. All paper documents are kept in a locked cabinet at ICDDR,B, the partner institution. The senior research investigator from ICDDR,B will have sole access to the locked cabinet. The specific address for the repository is the Programme on Infectious Diseases and Vaccine Sciences Office, ICDDR,B: International Centre for Diarrhoeal Disease Research, Mohakhali, Dhaka 1212. The database itself will only contain de-identified data.

3. Results

KEY POINT

- A series of analyses were conducted to answer questions on measuring handwashing behavior.

Demographic Characteristics

We enrolled a total of 100 respondents, 50 into the extended five-hour observation group and 50 into the abbreviated 90-minute observation group. Respondents came from two districts: Brahmanbaria and Sirajganj which lie approximately two hours east and three hours west of the Bangladeshi capital, Dhaka. All respondents enrolled were the mothers of children < 2 years old. The mean age of respondents was 25.7 years (SD 5.7, 16–45) and the mean age of the youngest children was 0.9 years (SD 0.6, 0–2). Mean household size was 5.8 persons (SD 2.1) with a range from 3 to 12; there was a mean of 1.5 children < 5 years old per household (SD 0.6, 1–3). Respondents reported a mean of 5.4 of years of education (SD 3.6, 0–12) for themselves and for 4.8 years of education for their respective heads of household (SD 4.3, 0–14).

We collected information on ownership of household goods, home wall and roof materials, type of fuel used for cooking, and homestead and other land ownership in order to assign households to socioeconomic status (SES) categories (Table 1). Among these, we opted to include in principal components

TABLE 1: DEMOGRAPHIC AND SOCIOECONOMIC CHARACTERISTICS OF ALL RESPONDENTS, AND BY DURATION OF STRUCTURED OBSERVATION, BANGLADESH, 2007

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Status of respondent (mother)	100%	100%	10%		
Respondent age, in years (mean, range, SD)	25.7 (16–45, SD 5.7)	25.5 (16–39, SD 5.5)	25.9 (17–45, SD 5.9)	.75	
Household size (mean, range, SD)	5.8 (3–12, SD 2.1)	5.6 (3–12, SD 2.3)	5.8 (3–11, SD 1.8)	.83	
Number <5 in household (mean, range, SD)	1.5 (1–3, SD .58)	1.5 (1–8, SD 0.6)	1.6 (1–3, SD 0.6)	.45	
Age of youngest child (mean, range, SD)	.91 (0.0–1.99, SD 0.6)	0.8 (0–1.9, SD 0.5)	1.0 (0.0–1.99, SD 0.6)	.10	
Electricity	56%	60%	52%	.42	1.2 (0.8–1.6)
Almirah	21%	24%	18%	.46	1.3(0.6–2.9)
Table	78%	74%	82%	.34	0.9 (0.7–1.1)
Chair	81%	82%	80%	.79	1.0 (0.8–1.2)
Watch	71%	72%	70%	.83	1.0 (0.8–1.3)
Khat	61%	70%	52%	.07	1.3 (1.0–1.9)
Chouki	79%	82%	76%	.46	1.1 (0.9–1.3)
Radio	39%	42%	36%	.54	1.2 (0.7–1.9)
Television, B&W	27%	32%	22%	.26	1.5 (0.8–2.8)
Television, color	8%	8%	8%	1.0	1.0 (0.3–3.8)
Refrigerator	1%	0%	2%	1.0	0.3 (0–8.0)
Bicycle	24%	30%	18%	.16	1.7 (0.8–3.4)
Motorbike	3%	6%	0%	.24	7.0 (0.4–132.1)

(Continued)

TABLE 1: (Continued)

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Sewing machine	9%	12%	6%	.48	2.0 (0.5–7.6)
Landline	0%	0%	0%		
Mobile	46%	56%	36%	.05	1.6 (1.0–2.4)
Roof material (tin)	99%	98%	100%	1.0	1.0 (0.9–1.0)
Wall material					
Kaccha	44%	34%	54%	.05	
Brick	6%	12%	0%		
Tin	50%	54%	46%		
Floor material (kaccha)	92%	88%	96%	.27	0.9 (0.8–1.0)
Fuel					
Wood	25%	34%	16%	.89	
Crop/grass	49%	38%	60%		
Dung	23%	24%	22%		
Gas	1%	0%	2%		
Other	2%	4%	0%		
Respondent's education, in years (mean, range, SD)	5.4 (0–12, SD 3.6)	5.8 (0–12, SD 3.6)	5.0 (0–10, SD 3.6)	.32	
Head of household education (mean, range, SD)	4.8 (0–14, SD 4.3)	5.1 (0–14, SD 4.0)	4.6 (0–14, SD 4.7)	.42	
SES Quartile				.35	
• Poorest	21	21%	22%		
• Quartile 2	28	23%	34%		
• Quartile 3	24	27%	22%		
• Wealthiest	25	29%	22%		

analysis (PCA) those categorical variables that were reported by at least 10% but by no more than 90% of respondents, and all continuous variables. After initial analysis, we excluded those variables that loaded on multiple components and were left with the following variables in the final model: electricity, *almirah* (a wardrobe or dresser), table, chair, bicycle, television, type of fuel used for cooking, and number of years of education of the head of the household. Based on previous uses of PCA for SES stratification,[19] we included only the first component, which was found to explain 33% of the variance in the study population. Using the factor score from this first component, we assigned households into SES quartile. To confirm appropriate allocation, we examined whether possession of household goods and other SES-related measures varied by SES quartile (Table 2).

Self-Report and Spot Checks of Handwashing Behavior

Mothers reported washing their hands a mean of 14.9 times (SD 5.2, 3–28) and washing hands with soap a mean of 4.0 (SD 2.1, 0–15) during the previous 24 hours. The most commonly reported reasons for washing hands with soap were bathing (90%), going to the toilet (56%), washing dishes (50%), and washing clothes (42%). Less frequently reported reasons for washing hands with soap are detailed in Table 3.

[19] Vyas and Kumaranayake 2006.

TABLE 2: SOCIOECONOMIC FACTORS, BY SOCIOECONOMIC QUARTILE AS DETERMINED BY PRINCIPAL COMPONENTS ANALYSIS

	Poorest Quartile (N = 21)	Quartile 2 (N = 28)	Quartile 3 (N = 24)	Wealthiest Quartile (N = 25)	P-Value**
Variables Included in the Final PCA Model					
Electricity	33%	29%	67%	92%	<.0001
Almirah	0%	0%	25%	60%	<.0001
Table	10%	96%	96%	100%	<.0001
Chair	14%	100%	100%	100%	<.0001
Bicycle	5%	7%	25%	56%	<.0001
Television, B&W*	5%	4%	42%	56%	<.0001
Television, color*	0%	0%	0%	32%	<.0001
Fuel (use of grass or dung cakes, as opposed to wood or liquid gas)	81%	86%	75%	52%	.0077
Mean number of years of education for HH head	3.4	1.6	6.9	7.6	<.0001
Variables NOT Included in Final PCA Model					
Mean number of years of education for respondent	1.9	3.8	7.4	8.4	<.0001
Mean number of decimals of homestead land	3.8	10.0	23.3	40.5	.0001
Mean number of decimals of other land	8.9	9.5	72.2	119.8	<.0001
Watch	33%	54%	96%	100%	<.0001
Radio	10%	29%	46%	72%	<.0001
Mobile phone	10%	21%	63%	88%	<.0001
Do not share a defecation place with other households	24%	41%	63%	76%	.0002

* For the PCA analysis, black-and-white television and color television were combined into a single variable that coded for no television, black-and-white television, and color television.
** p-for-trend or p for linear correlation, as appropriate.

TABLE 3: SELF-REPORTED HANDWASHING BEHAVIOR

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Number of times hands washed "yesterday" (mean, range, SD)	14.9 (3–28, SD 5.2)	14.4 (4–25, SD 4.7)	15.4 (3–28, SD 5.7)	.49	
Number of times hands washed <u>with soap</u> "yesterday" (mean, range, SD)	4.0 (0–15, SD 2.1)	3.8 (0–8, SD 1.7)	4.2 (0–15, SD 2.4)	.64	
Why Soap Used "Yesterday"					
I was going to eat	2%	0%	4%	.49	0.2 (0–4.1)
I was going to toilet	56%	50%	62%	.23	0.8 (0.6–1.1)
I was feeding my child	6%	6%	6%	1.0	1.0 (0.2–4.7)
I had cleaned my child's bottom	31%	26%	36%	.28	0.7 (0.4–1.3)
I was cooking	25%	30%	20%	.25	1.5 (0.7–3.0)
I just felt like I needed to wash them	12%	14%	10%	.54	1.4 (0.5–4.1)
I felt like my hands were dirty	9%	8%	10%	1.9	0.8 (0.2–2.8)
I was bathing	90%	88%	92%	.74	1.0 (0.8–1.1)
Washing dishes	50%	52%	48%	.69	1.1 (0.7–1.6)
Cleaning/washing clothes	42%	34%	50%	.11	0.7 (0.4–1.1)
Other	13%	14%	12%	.77	1.2 (0.4–3.2)

*p-value calculated based on Mantel-Haenszel chi-square, t-test, or Wilcoxon rank sum, as appropriate.

TABLE 4: SOAP USE BY CRITICAL TIMES FOR HANDWASHING, AS OBSERVED DURING STRUCTURED OBSERVATION

Critical Times for Handwashing	Number of Households with at Least One Observation of This Critical Time	Number of Households Observed to Use Soap to Wash Hands at Least Once During This Critical Time	%
All Household Members			
All opportunities for handwashing	100	48	48
Defecation-related events	56	24	43
Water-related events	62	3	5
Food preparation or serving	88	12	14
Before eating	95	20	21
Before feeding a child	78	15	19
After returning to the compound	64	8	13
Primary Caregivers			
All opportunities for handwashing	99	41	41
Defecation-related events	44	19	43
Water-related events	47	2	4
Food preparation or serving	83	12	14
Before eating	51	0	0
Before feeding a child	74	14	19
After returning to the compound	0	—	—

A total of 97 respondents reported having a specific place to wash their hands after using the toilet. Among these, 89 (92%) had water present at the handwashing station at the time of the interview. Bar soap was observed in 16 (16%), laundry soap in 12 (12%), and mud in 15 (15%) of households. A handwashing station designated to wash hands before cooking, eating, or feeding a child—distinct from the place to wash hands after using the toilet—was observed in 70 households. Of these, 37 (53%) had water available for handwashing at the time of the interview. Bar soap was seen in 1 (1%) household, laundry soap in 3 (4%), and mud in 7 (10%).

Interviewers were instructed to record the appearance of the mother's and youngest child's nails, palms, and finger pads. The mother's nails were reportedly clean for 47% of respondents. Mother's palms and fingerpads were reportedly clean for 74% and 73%, respectively. The child's nails were reportedly clean for 15%, with palms and finger pads noted to be clean for 42% and 39%, respectively.

When asked to demonstrate handwashing, 59% of the 100 respondents used soap and water, 28% used mud and water, 10% used ash and water, and 9% used water alone. Both hands were washed by 58%, with a mean duration of rubbing hands with soap of 10 seconds (SD 6, 2–26). The majority (84%) dried their hands on the clothes they were wearing.

During the structured observations, at least one opportunity for handwashing by any household member was observed in all 100 study households (Table 4). Defecation-related opportunities for observing handwashing behavior were witnessed in 56; these included the respondent defecating, the respondent going to the toilet, and the respondent cleaning a child who had defecated. In 62 households, there was at least one opportunity to observe water-related events, including collecting or storing water. Food preparation or serving events were observed in 88 households and it was possible to observe at least one household member eating in 95. In 78 households, there was at least one opportunity to observe a household member feeding the child. A household member returning to the compound from outside was observed in 64 households. During these opportunities for handwashing, any use of soap for handwashing was observed in 48% of households. Use ranged from 5% in the context of water-related events to 43% in defecation-related events (Table 4).

Interviewers were able to observe at least one opportunity for handwashing by 99 primary caregivers. The proportion of primary caregivers observed to be using soap during specific critical times were as follows: defecation-related events (43%), water-related events (47%), food preparation or serving events (83%), before eating (51%), and before feeding (74%). No primary caregivers were observed returning to the compound. A total of 41% of primary caregivers with at least one opportunity for handwashing were observed to use soap at least once. Soap use by the primary caregivers during specific critical times ranged from 4% during water-related events to 43% during defecation-related events (Table 4).

Sensor products, specifically the sensor soap and sensor *bodna* were assigned to the 50 extended observation households. Of these 50 households, seven had no bar of soap in the home at the time of the interview and were given one sensor soap. One sensor soap was provided to 40 households that had one bar of soap currently in use in the home. Two sensor soaps each were provided to three households. A total of 82 sensor soaps were distributed. For those households provided two sensor soaps, the number of soap-use events below reflects the total number of events recorded by both sensor soaps. At the time of collection, 37 (45%) of 82 sensor soaps weighed less than 65 grams, the weight at which the motion sensor embedded in the soap is expected to become visible. Indeed, 35 (43%) of 82 motion sensors were reportedly visible to the field worker at the time of collection. None of the sensor soaps or sensor *bodnas* were lost. Detailed information, including the raw data, rates of unsuccessful data recovery from loggers, and the 24-hour soap-use patterns are available in a report from colleagues at Unilever. We are awaiting data from sensor *bodnas* at this time.

Reactivity to Sensor Soap

Extended structured observations were conducted in 25 households on the fourth full day of sensor soap placement and in 25 households on the eighth full day of sensor soap placement. The presence of sensor soap in the home may have led to a high degree of soap use on Day 1, the first full day of sensor soap observation, and that soap use might wane over the subsequent days as the novelty of the soap declined. Using paired t-tests to detect such reactivity, we compared the number of soap-use events in each household on Day 1, the first full 24 hours of sensor soap observation, to mean number of soap-use events on Days 2–3 or Days 2–7, for households observed on the fourth full day and the eighth full day of sensor soap placement, respectively. Data for this analysis was available for 21 households observed on the fourth full day and 17 households observed on the eighth full day of sensor soap placement. Among households observed on the fourth full day of sensor soap placement, the mean number of sensor soap uses on Day 1 was 9.0 (SD 3.1, range 5–16) and on Days 2–3 was 7.6 (SD 3.6, range 2–18) ($p = .16$). When we performed a retrospective power calculation, we determined that we had 90% power to detect this difference at the 95% significance level. Among households observed on the eighth full day of sensor soap placement, the mean number of sensor soap uses on Day 1 was 10.6 (range 4.7–23.0) and on Days 2–7 was 10.6 (range 4.7–23.0) ($p = .30$). We had 67% power to detect this difference at the 95% significance level.

Association Between Self-Reported Measures of HW Behavior and Sensor Soap Use

We looked for associations between self-reported measures of HW behavior and sensor soap use, using the mean 24-hour per capita sensor soap use on pre-observation days. For this analysis, Groups 1 and 2 were combined. We found no significant associations between any of the self-reported measures of handwashing behavior and 24-hour per capita sensor soap use (Table 5). We examined the relationship between the self-reported number of times that hands were washed, or washed with soap, on the day before the interview and sensor soap use and found no significant correlations.

Association Between Spot-Check Measures of HW Behavior and Sensor Soap Use

We looked for associations between spot-check measures of HW behavior and sensor soap use, using the mean 24-hour

TABLE 5: ASSOCIATION BETWEEN SPOT CHECK, SELF-REPORTED, AND STRUCTURED OBSERVATION MEASURES OF HANDWASHING BEHAVIOR, AND PER CAPITA SENSOR SOAP USE ON PRE-OBSERVATION DAYS

	Those with the Measure					Those without the Measure					P-Value*
Measure	N	Mean Per Capita Sensor Soap-Use Events	SD	Max	Min	N	Mean Per Capita Sensor Soap-Use Events	SD	Max	Min	
Spot-Check Measures											
Bar soap at the toilet handwash station	9	1.52	0.53	2.5	1.66	37	1.51	0.71	3.33	0.48	0.97
Laundry soap at the toilet handwashing station	6	1.35	0.48	2.2	0.97	40	1.53	0.69	3.33	0.48	0.44
Soap available inside the toilet	3	1.49	0.62	2.2	1.02	41	1.5	0.68	3.33	0.48	0.97
Mother washed hands with soap before giving the interviewer a glass of water to drink	5	1.02	0.42	1.5	0.48	40	1.63	0.67	3.33	0.58	0.04
Private toilet (i.e., not sharing a toilet with any other household)	24	1.38	0.55	2.48	0.54	20	1.67	0.79	3.33	0.48	0.17
Mother's nails are dirty	28	1.47	0.63	2.67	0.48	18	1.57	0.75	3.33	0.6	0.63
Mother's fingerpads are dirty	13	1.59	0.88	3.33	0.48	33	1.47	0.58	2.67	0.58	0.62
Mother's palm is dirty	12	1.38	0.78	2.67	0.48	34	1.56	0.63	3.33	0.6	0.43
Child's nails are dirty	33	1.51	0.67	3.33	0.48	7	1.54	0.73	2.67	0.67	0.89
Child's fingerpads are dirty	27	1.59	0.7	3.33	0.48	19	1.39	0.63	2.67	0.67	0.31
Child's palm is dirty	25	1.65	0.69	3.33	0.48	21	1.34	0.63	2.67	0.54	0.12
Mother used soap to wash hands when asked to demonstrate usual handwashing behavior	26	1.46	0.58	2.5	0.48	18	1.61	0.79	3.33	0.54	0.47
SES Quartile 1	9	1.55	0.9								
2	10	1.85	0.71	**							
3	12	1.68	0.48								
4	13	1.1	0.42								

(*Continued*)

* p-value calculated based on Mantel-Haenszel chi-squares, t-test, or Wilcoxon rank sum, as appropriate.

** There was a significant difference in the mean per capita soap-use events between Quartile 2 and Quartile 4 using analysis of variance ($p < .05$).

TABLE 5: (Continued)

Measure	Those with the Measure N	Mean Per Capita Sensor Soap-Use Events	SD	Max	Min	Those without the Measure N	Mean Per Capita Sensor Soap-Use Events	SD	Max	Min	P-Value*
Self-reported Measures											
Number of times hands were washed during the previous day (continuous)	R = –.090	p = .56									
Number of times hands were washed with soap during the previous day (continuous)	R = .078	p = .61									
Hands were washed after going to the toilet	23	1.54	0.69	3.33	0.48	21	1.49	0.66	2.67	0.54	0.79
Hands were washed before feeding the child	2	2.19	0.39	2.48	1.92	42	1.49	0.67	3.33	0.48	0.15
Hands were washed after cleaning a child's bottom	11	1.46	0.64	2.5	0.48	33	1.54	0.69	3.33	0.54	0.75
Hands were washed while bathing	40	1.49	0.66	3.33	0.48	4	1.77	0.82	2.67	0.97	0.44
Hands were washed before cooking	15	1.45	0.6	3.33	0.9	29	1.56	0.71	2.67	0.49	0.62
Structured Observation Variables											
Any use of soap during the structured observation	35	1.52	0.69	3.33	0.48	9	1.54	0.61	2.5	0.9	0.92
Any use of soap during defecation-related events	19	1.41	0.54	2.67	0.48	18	1.75	0.77	3.33	0.58	0.13
Any use of soap during water-related events	2	1.02	0.03	1.05	1	31	1.59	0.68	3.33	0.54	0.26
Any use of soap before food preparation or serving	10	1.56	0.85	3.33	0.58	34	1.51	0.63	2.67	0.48	0.86
Any use of soap before eating	17	1.39	0.69	2.67	0.48	27	1.6	0.66	3.33	0.6	0.32
Any use of soap before feeding the child	11	1.29	0.63	2.67	0.54	29	1.62	0.66	3.33	0.58	0.16
Any use of soap after returning to the compound from outside	7	1.42	0.66	2.48	0.54	26	1.47	0.64	2.67	0.58	0.87
Any use of soap by the primary caregiver during the structured observation	28	1.5	0.68	3.33	0.48	15	1.58	0.68	2.5	0.54	0.71
Any use of soap by the primary caregiver during defecation-related events	15	1.45	0.55	2.67	0.48	19	1.93	0.74	3.33	0.58	0.16
Any use of soap by the primary caregiver during water-related events	1	1	.	1	1	23	1.74	0.66	3.33	0.6	0.28
Any use of soap by the primary caregiver before food preparation or serving	10	1.56	0.85	3.33	0.58	32	1.51	0.64	2.67	0.48	0.86
Any use of soap by the primary caregiver before feeding the child	10	1.37	0.61	2.67	0.6	29	1.59	0.69	3.33	0.54	0.37

per capita sensor soap use on pre-observation days. For this analysis, Groups 1 and 2 were combined. Mothers who washed their hands with soap before giving the interviewer a glass of water to drink had a lower mean 24-hour per capita sensor soap use (mean 1.0, SD 0.4) than mothers who did not wash hands with soap before giving the interviewer a glass of water to drink (mean 1.6, SD 0.7) ($p = .04$). We found no significant associations between any other spot-check measures of handwashing behavior and 24-hour per capita sensor soap use (Table 5).

Association Between Socioeconomic Status and Sensor Soap Use

Using analysis of variance methods, we examined the relationship between socioeconomic status, as defined by SES quartiles based on PCA analysis presented above and found differences between the second poorest and the wealthiest quartiles with respect to sensor soap use. The mean number of 24-hour per capita sensor soap-use events was 1.9 (SD 0.7) among those in the second poorest quartile and 1.1 (SD 0.4) among those in the wealthiest SES quartile ($p < .05$). There were no other significant differences when comparing any of the other quartiles to each other (Table 5).

Association Between Observation of Soap Use During Structured Observation and Sensor Soap Use

Since we had the opportunity to detect soap usage using both sensor soap and structured observation during the five-hour structured observation block, we compared results from the two methods using paired t-tests. For this analysis, we combined households observed on the fourth full day and the eighth full day of sensor soap placement for a total of 45 households. Sensor soap data was not available for the remaining five households due to logger malfunction. The mean number of soap-use events recorded by the sensor soap was 6.3 (range 0–18) and the mean number of observations of soap use by the observer was 1.8 (range 0–7) ($p < .0001$). Although the mean number of soap-use events detected by the sensor soap during the five hours of the structured observation was 6.3, observers only recorded a mean of 1.4 observations of Lifebuoy soap. In many cases, observers informed study supervisors that they were unable to determine from their vantage points whether the soap observed was Lifebuoy or another brand.

We found no association between the observation of soap use, by any household member, or by the primary caregiver in particular, at any time or during specific critical times, and 24-hour per capita sensor soap use measured on pre-observation days (Table 5).

Study Question 1: How reactive is a subject, with respect to HW behavior, when observed during an extended structured observation?

Reactivity to Structured Observation

Using paired t-tests, we examined whether there was reactivity to the presence of the observer during SO time block, or the five-hour time block of the extended structured observation. Data for this analysis was available for 21 households in the group that had the structured observation on the fourth full day after sensor soap placement and 24 households in the group that had the structured observation on the eighth full day after sensor soap placement. The mean number of soap-use events, as measured by the sensor soap and that occurred during the SO time block on Days 1–3 for the group that had the structured observation on the fourth full day after sensor soap placement was 3.4 events (range 0–6). On Day 4, the day of the structured observation, the mean number of soap-use events was 4.8 (range 0–10) ($p = 0.012$). The mean number of soap-use events that occurred during the SO time block on Days 1–7 for the group that had the structured observation on the fourth full day after sensor soap placement was 5.3 events (range 2–10). On Day 8, the day of the structured observation, the mean number of soap-use events was 7.6 (range 1–18) ($p = 0.011$).

In order to describe those households that were particularly reactive to the presence of the observer, we divided all 50 households into quartiles of reactivity, with those in the top quartile being defined by having the highest absolute differences in sensor soap-use events between the SO time block on the SO day and the preceding days. We compared these most reactive households to the lower those in the lower three quartiles of reactivity. In the group of households observed on the fourth full day of sensor soap placement, which had six reactors and 15 non-reactors, we found several statistically significant predictors of reactivity. Reactors had a significantly higher number of years of education for both the respondent $p = .03$) and the household head ($p = .02$) (Table 6). Additionally,

TABLE 6: PREDICTORS OF REACTIVITY TO THE STRUCTURED OBSERVATION AMONG HOUSEHOLDS OBSERVED ON THE FOURTH FULL DAY AFTER SENSOR SOAP PLACEMENT, BANGLADESH, 2007

Variable	Reactors (N = 6)	Non-Reactors (N = 15)	P-Value
Mean difference in number of soap-use events between structured observation day and mean of pre-observation days	4.4 (3−6.3, SD 1.3)	0.2 (–2.7−2.7, SD 1.5)	<.0001
Mean number of soap-use events during the SO time block on pre-observation days	3.6 (2.7−4.3, SD 0.6)	3.3 (0.3−6.3, SD 1.4)	.66
Child <5 with cough	2 (33%)	0 (0%)	.07
Report that diarrhea can be prevented by covering food containers	6 (100%)	7 (50%)	.05
Number of years education for the primary caregiver	8.5 (6–10, SD 1.6)	4.6 (0–10, SD 3.7)	.03
Number of years education for household head	8.3 (0–14, SD 4.5)	3.9 (0–10, SD 3.9)	.02
Land ownership (other than homestead land, units in "decimals")	80.5 (0–165, SD 72)	24 (0–132, SD 42)	.07
Watch	6 (100%)	8 (53%)	.06
Mobile phone	5 (83%)	5 (33%)	.06
Stores water in a clay or aluminum pot	4 (80%)	3 (23%)	.05
Do not share a defecation place with other households	5 (83%)	4 (29%)	.05
HW station inside or within three steps of the toilet	6 (100%)	5 (33%)	.01
Socioeceonomic status quartile			
• Poorest quartile	0 (0%)	4 (31%)	.05
• Quartile 2	1 (17%)	5 (38%)	
• Quartile 3	2 (33%)	3 (8%)	
• Wealthiest quartile	3 (50%)	4 (23%)	
Soap logger weight <65 grams at the time of collection	0 (0%)	5 (50%)	.26
Logger visible at the time of collection	1 (17%)	2 (15%)	1.0
Participated in hand microbiology assessment	2 (33%)	6 (40%)	1.00

*p-value calculated based on Mantel-Haenszel chi-square, t-test, or Wilcoxon rank sum, as appropriate.

reactors were more likely to have a handwashing station inside or within three steps of the toilet than non-reactors (p = .01). Additional predictors were not having to share a defecation place with other households (p = .05), storing water in a clay or aluminum pot (p = .05), ownership of a mobile phone (p = .06), and ownership of a watch (p = .06). There was no significant difference between reactors and non-reactors with respect to assignment to SES quartiles. Obvious visibility of the logger and soap weight <65 grams, a sign that the logger might be visible through the soap, were also not predictive of

TABLE 7: PREDICTORS OF REACTIVITY TO THE STRUCTURED OBSERVATION AMONG HOUSEHOLDS OBSERVED ON THE EIGHTH FULL DAY AFTER SENSOR SOAP PLACEMENT, BANGLADESH, 2007

Variable	Reactors (N = 6)	Non-Reactors (N = 18)	P-Value
Mean difference in number of soap-use events between structured observation day and mean of pre-observation days	7.9 (6.6−10.3, SD 1.3)	0.4 (−3.71−3.3, SD 2.4)	<.0001
Mean number of soap-use events during the SO time block on pre-observation days	5.9 (2.7–10.6, SD 3.1)	5.2 (2.0−10.6, SD 2.6)	.63
Considers child's stools dirty	3 (50%)	0 (0%)	.01
Stores water in a clay/ aluminum pot	1 (17%)	11 (69%)	.06
Mobile phone	6 (100%)	9 (50%)	.05
Natural walls	6 (100%)	11 (61%)	.08
Number of rooms for sleeping in the house	3.7 (2–6, SD 1.4)	2.6 (1–7, SD 1.9)	.10
Do not share a defecation place with other households	6 (100%)	8 (47%)	.05
Child defecates in potty (as opposed to "no specific place")	2 (33%)	0 (0%)	.05
Child's palms observed to be "clean"	5 (83%)	7 (39%)	.05
Washed both hands upon request to demonstrate usual HW	6 (100%)	9 (60%)	.05
Socioeceonomic status quartile • Poorest quartile • Quartile 2 • Quartile 3 • Wealthiest quartile	1 (17%) 0 (0%) 4 (67%) 1 (17%)	4 (22%) 3 (17%) 5 (28%) 6 (33%)	.83
Soap logger weight <65 grams at the time of collection	4 (67%)	9 (50%)	.65
Logger visible at the time of collection	3 (50%)	10 (56%)	1.0
Participated in hand microbiology assessment	2 (33%)	11 (61%)	.36

*p-value calculated based on Mantel-Haenszel chi-square, t-test, or Wilcoxon rank sum, as appropriate.

reactivity. Whereas two (33%) of the six reactors participated in hand rinses for quantitation of hand contamination, six (40%) of non-reactors did so ($p = 1.00$). Notably, reactors and non-reactors had similar numbers of soap-use events during the SO time block on pre-observation days.

In the group of households observed on the eighth full day of sensor soap placement, believing that a child's stools are dirty ($p = .01$) was a significant predictor of reactivity (Table 7). Additional predictors included ownership of a mobile phone ($p = .05$), not sharing a defecation place with other households (p = .05), child defecating in a potty (p = .05), and washing both hands upon request to demonstrate usual handwashing practice (p = .05). There was no significant difference between reactors and non-reactors with respect to assignment to SES quartiles. Obvious visibility of the logger and soap weight <65 grams were not predictive of reactivity. Whereas two (33%) of the six reactors participated in hand rinses for quantitation of hand contamination, 11 (61%) of non-reactors did so (p = .36). Reactors and non-reactors had similar numbers of soap-use events during the SO time block on pre-observation days.

Due to the small sample sizes, we have not conducted multivariate modeling to assess for confounding between the several predictors of reactivity.

Study Question 2: What is the optimum duration of a structured observation for the purposes of measuring HW behavior?

Comparison of Extended Observation Group to the Abbreviated Observation Group

Fifty households were enrolled into the extended observation group and 50 households into the abbreviated observation group. All respondents were the mothers of children <2 years old, with mean ages of 26 years in each group. The mean age of the youngest child was 0.8 years in the extended group and 1.0 year in the abbreviated group (p = .10). The extended and abbreviated observation groups differed significantly with respect to mobile phone ownership (56% and 36%, respectively, p = .05), material used for the walls of the home (details in Table 1, p = .05), frequency of cough in the youngest child during the previous two weeks (8% and 42%, respectively, p = .0001) and diarrhea in the youngest child during the previous two weeks (10% and 26%, respectively, p = .04) (Table 8). There was no difference between the extended and abbreviated observation groups with respect to other demographic characteristics, SES quartile, education of the mother or the head of household, or self-reported or spot-check measures of handwashing behavior (Table 1, Table 3, Table 9, Table 10, Table 11, Table 12).

Interviewers were instructed to maintain the primary caregiver of the child younger than two years old as the primary focus of the structured observation. All other household members were to be observed as they came into the line of sight of the interviewer. Opportunities to observe handwashing included the following critical times for washing hands with soap: defecation-related events, water-related events, food preparation and serving, eating, feeding a child, and after returning to the compound from outside. There was at least one observation of a handwashing opportunity *by any household member* in all of the extended and abbreviated observation households. The mean of the total number of opportunities to observe handwashing behavior by any household member was 22.7 (range 11–37) in the extended group and 7.5 (range 2–24) in the abbreviated group (p < .0001) (Table 13). Defecation-related opportunities were observed in 42 (84%) extended observation households and 14 (28%) abbreviated observation households (p < .0001). Similarly, extended observation households were significantly more likely than abbreviated observation households to have at least one opportunity to observe handwashing during a water-related event, food preparation event, eating event, child-feeding event, and upon return to the compound. In households that did not have any event for a specific critical time, we coded the number of opportunities for witnessing that critical time as zero for that household. The mean number of opportunities in total and during specific critical times was significantly greater in the extended observation households. The ratio of the mean number of opportunities in the extended observation group to the abbreviated observation group ranged from 2.5 (water-related events and before feeding a child) to 5.3 (defecation-related events).

TABLE 8: DISEASE EXPERIENCE IN YOUNGEST CHILD DURING TWO WEEKS PRECEDING INTERVIEW

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Cough	25%	8%	42%	.0001	0.2 (0.1–0.5)
Difficulty breathing	15%	12%	18%	.40	0.7 (0.3–1.7)
Diarrhea	18%	10%	26%	.04	0.4 (.1–1.0)

*p-value calculated based on Mantel-Haenszel chi-square.

TABLE 9: SELF-REPORTED HANDWASHING AND SOAP USE

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Number of times hands washed "yesterday" (mean, range, SD)	14.9 (3–28, SD 5.2)	14.4 (4–25, SD 4.7)	15.4 (3–28, SD 5.7)	0.49	
Number of times hands washed with soap "yesterday" (mean, range, SD)	4.0 (0–15, SD 2.1)	3.8 (0–8, SD 1.7)	4.2 (0–15, SD 2.4)	0.64	
Why Soap Was Used "Yesterday"					
I was going to eat	2%	0%	4%	0.49	0.2 (0–4.1)
I was going to toilet	56%	50%	62%	0.23	0.8 (0.6–1.1)
I was feeding my child	6%	6%	6%	1.0	1.0 (0.2–4.7)
I had cleaned my child's bottom	31%	26%	36%	0.28	0.7 (0.4–1.3)
I was cooking	25%	30%	20%	0.25	1.5 (0.7–3.0)
I just felt like I needed to wash them	12%	14%	10%	.54	1.4 (0.5–4.1)
I felt like my hands were dirty	9%	8%	10%	1.9	0.8 (0.2–2.8)
I was bathing	90%	88%	92%	0.74	1.0 (0.8–1.1)
Washing dishes	50%	52%	48%	0.69	1.1 (0.7–1.6)
Cleaning/washing clothes	42%	34%	50%	0.11	0.7 (0.4–1.1)
Other	13%	14%	12%	0.77	1.2 (0.4–3.2)

*p-value calculated based on Mantel-Haenszel chi-square, t-test, or Wilcoxon rank sum, as appropriate.

TABLE 10: OBSERVATION OF HANDWASHING STATIONS

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Specific place to wash hands after using the **toilet**	97	50 (100%)	47 (94%)		
Water observed at HW station	89 (92%)	48 (96%)	41 (87%)	0.15	1.1 (1.0–1.2)
Bar soap observed at HW station	16	9 (18%)	7 (15%)	0.68	1.2 (0.5–3.0)
Laundry soap observed at HW station	12	6 (12%)	6 (13%)	0.91	0.9 (0.3–2.7)
Mud observed at HW station	15	10 (20%)	5 (11%)	0.20	1.9 (0.7–5.1)
Specific place to wash hands **before cooking, eating, or feeding child** (different from toilet HW location)	74	38 (76%)	36 (72%)		
Kitchen HW station distinct from Toilet HW station	70	35 (70%)	35 (70%)	1.0	1.0 (0.8–1.3)
Water observed at HW station	37 (53%)	18 (51%)	19 (54%)	0.81	0.9 (.6–1.5)
Bar soap observed at HW station	1	0%	1 (3%)	1.0	0.3 (0–7.9)
Laundry soap observed at HW station	3 (4%)	1 (3%)	2 (6%)	1.0	0.5 (0–5.3)
Mud observed at HW station	7 (10%)	3 (9%)	4 (11%)	1.0	0.8 (0.2–3.1)

*p-value calculated based on Mantel-Haenszel chi-squares.

TABLE 11: OBSERVATION OF MOTHER'S AND CHILD'S HANDS

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value
Child's nails				.68
Visible dirt	66%	64%	68%	
Unclean appearance	19%	20%	18%	
Clean	15%	16%	14%	
Child's palm				.91
Visible dirt	29%	30%	28%	
Unclean appearance	29%	28%	30%	
Clean	42%	42%	42%	
Child's fingerpads				.90
Visible dirt	30%	30%	30%	
Unclean appearance	31%	32%	30%	
Clean	39%	38%	40%	
Mother's nails				.11
Visible dirt	33%	40%	26%	
Unclean appearance	20%	20%	20%	
Clean	47%	40%	54%	
Mother's palm				.52
Visible dirt	8%	10%	6%	
Unclean appearance	18%	18%	18%	
Clean	74%	72%	76%	
Mother's fingerpads				.19
Visible dirt	11%	16%	6%	
Unclean appearance	16%	12%	20%	
Clean	73%	72%	74%	

*p-value calculated based on chi-squares.

TABLE 12: DEMONSTRATION OF HANDWASHING BY MOTHER

Variable	Total (N = 100)	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk (95% CI)
Used only water	9%	14%	4%	.08	3.5 (0.76–16.03)
Used soap	59%	56%	62%	.54	0.9 (0.6–1.3)
Used ash	10%	8%	12%	.51	0.67 (0.2–2.2)
Used mud	28%	24%	32%	.37	0.75 (0.4–1.4)
Washed both hands	58%	56%	60%	.69	0.9 (0.6–1.4)
Duration of rubbing hands with soap (seconds)	10.4 (2–26, SD 5.5)	9.3 (3–21, SD 4.9)	11.3 (2–26, SD 5.9)	.13	
Dried hands on clothing she was wearing	84%	84%	84%	1.0	

*p-value calculated based on Mantel-Haenszel chi-square, t-test, or Wilcoxon rank sum, as appropriate.

TABLE 13: OPPORTUNITIES TO OBSERVE RESPONDENTS (ALL MEMBERS OF THE HOUSEHOLD) DURING THE STRUCTURED OBSERVATION

Variable	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk	Ratio of Opportunities (Extended : Abbreviated)
Mean number of opportunities observed during SO	22.7 (range 11–37, SD 5.9)	7.5 (range 2–24, SD 3.7)	<.0001		3.0
Any opportunity to observe defecation-related events: after cleaning child's anus after defecation after toileting	84%	28%	<.0001	3.0 (1.9–4.8)	
Mean number of opportunities to observe defecation-related events	2.1 (range 0–8, SD 1.8)	0.4 (range 0–3, SD 0.8)	<.0001		5.3
Any opportunity to observe water-related events	74%	50%	0.014	1.5 (1.1–2.0)	
Mean number of opportunities to observe water-related events	1.5 (range 0–5, SD 1.3)	0.6 (range 0–2, SD 0.7)	.0011		2.5
Any opportunity to observe cooking, preparing, or serving food	100%	76%	<.0001	1.3 (1.1–1.5)	
Mean number of opportunities to observe cooking, preparing, or serving food	4.1 (range 1–10, SD 1.9)	1.6 (range 0–6, SD 1.4)	<.0001		2.6
Any opportunity to observe handwashing before eating	100%	90%	0.02	1.1 (1.0−1.2)	
Mean number of opportunities to observe handwashing before eating	6.7 (range 1–15, SD 3.2)	2.3 (range 0–8, SD 1.8)	<.0001		2.9
Any opportunity to observe handwashing before feeding the child	92%	64%	0.001	1.4 (1.1−1.8)	
Mean number of opportunities to observe handwashing before feeding the child	2.5 (range 0–6, SD 1.5)	1.0 (range 0–5, SD 1.0)	<.0001		2.5
Any opportunity to observe handwashing after returning to the compound	76%	52%	0.01	1.5 (1.1–2.0)	
Mean number of opportunities to observe handwashing after returning to the compound	1.8 (range 0–8, SD 1.7)	0.6 (range 0–3, SD 0.7)	.0002		

*p-value calculated based on Mantel-Haenszel chi-squares, t-test, or Wilcoxon rank sum, as appropriate.

There was at least one opportunity to observe handwashing behavior by the primary caregiver in 49 (98%) extended observation households and 50 (100%) abbreviated observation households (Table 14). Observation of the primary caregiver was not possible in the 50th extended observation household because the caregiver was away due to illness in the family. There was a significant difference in the opportunity to observe the primary caregiver during defecation-related events (76% of extended observation households vs. 12% of abbreviated observation households, $p < .0001$). Opportunities to observe handwashing behavior during food preparation ($p < .0001$), before eating

TABLE 14: OPPORTUNITIES TO WITNESS HANDWASHING BEHAVIOR AMONG PRIMARY CAREGIVERS

Variable	Extended (N = 50)	Abbreviated (N = 50)	P-Value	Relative Risk	Ratio of Opportunities (Extended: Abbreviated)
Any opportunity to observe the primary caregiver	98%	100%	0.31	1 (0.9–1.0)	
Mean number of opportunities to observe the primary caregiver	10.3 (range 0–23, SD 4.0)	3.3 (range 1–13, SD 1.9)	<.0001		3.1
Any opportunity to observe defecation-related events	76%	12%	<.0001	6.3 (2.9–13.6)	
Mean number of opportunities to observe defecation-related events	1.2 (range 0–4, SD 0.9)	0.2 (range 0–2, SD 0.5)	<.0001		6.0
Any opportunity to observe water-related events	56%	38%	0.07	1.5 (1.0–2.3)	
Mean number of opportunities to observe water-related events	1.0 (range 0–4, SD 1.2)	0.4 (range 0–2, SD 0.6)	<.0001		2.5
Any opportunity to observe cooking, preparing, or serving food	96%	70%	<.0001	1.4 (1.1–1.7)	
Mean number of opportunities to observe cooking, preparing, or serving food	3.2 (range 0–9, SD 1.7)	1.4 (range 0–4, SD 1.2)	<.0001		2.3
Any opportunity to observe handwashing before eating	82%	20%	<.0001	4.1 (2.3–7.2)	
Mean number of opportunities to observe handwashing before eating	1.3 (range 0–4, SD 0.9)	0.2 (range 0–1, SD 0.4)	<.0001		6.5
Any opportunity to observe handwashing before feeding the child	90%	58%	0.0003	1.6 (1.2–2.0)	
Mean number of opportunities to observe handwashing before feeding the child	2.3 (range 0–6, SD 1.5)	0.9 (range 0–5, SD 1.0)	<.0001		2.6
Any opportunity to observe handwashing after returning to the compound	0%	0%	—	—	

*p-value calculated based on Mantel-Haenszel chi-squares, t-test, or Wilcoxon rank sum, as appropriate.

(p <.0001), and before feeding a child (p = .0003) were fewer in abbreviated observation households. The mean number of opportunities in total and during specific critical times was significantly greater in the extended observation households. The ratio of the mean number of opportunities in the extended observation group to the abbreviated observation group ranged from 2.3 (food preparation) to 6.5 (before eating).

We hypothesized that, if soap were used consistently during different opportunities for handwashing, the need to witness specific critical times for handwashing would be reduced. Thus, we examined whether soap use during a specific critical time was associated with soap use during other critical times, using observation of soap use during the structured observation. Among all household members, soap use during a defecation-related event was significantly associated with soap use before eating (p=.008, RR = 3.2, 95% CI = 1.3–7.9). There was no other significant association in soap use between different opportunities for handwashing (Table 15). Among primary caregivers, soap use during a defecation-related event was significantly associated with soap use before feeding the child (p=.02, RR 7.9, 95% CI = 1.0–60.4). There was no other significant association in soap use between different opportunities for handwashing (Table 16).

Study Question 3: What is the change in soap-use behavior during repeated structured observations within the same household?

We conducted structured observations for five hours on each of three consecutive days. All of these households had the first day of structured observation on the eighth full day that sensor soap had been in the home. As a reminder, sensor soap was provided to these households on Day 0, collected on Day 4 and replaced with a new sensor soap, with subsequent structured observation on Day 8. Thus, the serial structured observations were done on Days 8, 9, and 10. The sensor soap was retrieved at the conclusion of the structured observation on Day 10. A summary of the results of structured observation and sensor soap-use detection is presented in Table 17. There was no significant difference in the number of opportunities to observe HW behavior during the five-hour structured observation across the three days. The mean number of soap-use events recorded by the observer were 1.6 on Day 8, 2.7 on Day 9, and 0.9 on Day 10 (overall effect p=.02). The differences in the number of soap-use events recorded by the observer on Day 8 and Day 9 (p=.10), and Day 8 and Day 10 (p=.14) were not statistically significant. The mean number of sensor soap-use events detected were 8.1 on Day 8, 7.6 on Day 9, and 3.5 on Day 10 (overall effect p=.05). The differences in the

TABLE 15: ASSOCIATION IN SOAP USE DURING DIFFERENT TYPES OF HW OPPORTUNITIES AMONG ALL HOUSEHOLD MEMBERS

	Defecation-Related Event	Water-Related Event	Before Food Preparation	Before Eating	Before Feeding the Child	After Returning from Outside the Compound
Defecation-related event	—	p = .55 RR = 2.9	p = .48 RR = .56	p = .008 RR = 3.2	p = .26 RR = 2.4	p = 1 RR = .85
Water-related event		—	p = .44 RR = 2.0	p = .54 RR = 1.5	p = 1 RR = .83	p = .27 RR = 3.9
Before food preparation			—	p = .13 RR = 2.0	p = 1 RR = .88	p = 1 RR = 1.1
Before eating				—	p = .52 RR = 1.4	p = .41 RR = 1.7
Before feeding the child					—	p = .64 RR = 1.5
After returning from outside the compound						—

TABLE 16: ASSOCIATION IN SOAP USE DURING DIFFERENT TYPES OF HW OPPORTUNITIES AMONG PRIMARY CAREGIVERS

	Defecation-Related Event	Water-Related Event	Before Food Preparation	Before Eating	Before Feeding the Child	After Returning from Outside the Compound
Defecation-related event	—	p = .17 RR = 6.7	p = .44 RR = 0.5	*	p = .02 RR = 7.9	**
Water-related event		—	p = .35 RR = 2.9	*	p = 1 RR = .84	**
Before food preparation			—	*	p = 1 RR = .83	**
Before eating				—	*	**
Before feeding the child					—	**
After returning from outside the compound						—

number of sensor soap-use events detected on Day 8 and Day 9 (p = .77), and Day 8 and Day 10 (p = .08) were not statistically significant.

There was no significant difference in the number of opportunities to observe HW behavior by the primary caregiver during the five-hour structured observation across the three days (Table 17). The mean number of soap-use events recorded by the observer were 0.7 on Day 8, 1.3 on Day 9, and 0.7 on Day 10 (overall effect p = .07). The difference in the number of soap-use events recorded by the observer on Day 8 and Day 9 was statistically significant

TABLE 17: SOAP USE DURING THREE CONSECUTIVE DAYS OF OBSERVATION BY STRUCTURED OBSERVATION AND SENSOR SOAP

Variable / N=9 Households	Day 8	Day 9	Day 10	Overall Treatment Effect (Unadjusted P-Value*)	Day 8 to Day 9 Comparison (Unadjusted P-Value*)	Day 8 to Day 10 Comparison (Unadjusted P-Value*)
All Household Members						
Number of opportunities to observe HW behavior	23.3	25	22.4	0.64	0.65	0.70
Number of soap-use events observed	1.6	2.7	0.9	0.02	0.10	0.14
Proportion of soap-use events to number of opportunities	0.07	0.11	0.04	0.03	0.10	0.18
Number of sensor soap-use events detected	8.1	7.6	3.5	0.05	0.77	0.08
Primary Caregivers						
Number of opportunities to observe HW behavior	9.2	10.4	9.4	0.67	0.52	0.84
Number of soap-use events observed	0.7	1.3	0.7	0.07	0.02	1.00
Proportion of soap-use events to number of opportunities	0.07	0.13	0.06	0.22	0.19	0.81

*p-value calculated based on analysis of variance.

TABLE 18: SUMMARY OF RESULTS OF HAND RINSE SAMPLING FROM MOTHERS

Variable	N	Minimum	Maximum	Arithmetic Mean (CFU / 100 ml)	Standard Deviation	Geometric Mean (CFU / 100 ml)
Random (fecal coliforms)	55	2	2320000	45228	312616	326
Random (*E. coli*)	55	0.5	3620	262	676	24
Time 1 (fecal coliforms)	49	0.5	112000000	2773047	16185012	2785
Time 1 (*E. coli*)	49	0.5	5400	282	784	53
Time 2 (fecal coliforms)	24	0.5	300000	18017	61840	554
Time 2 (*E. coli*)	24	0.5	1600	195	414	28
Recontamination assessment (fecal coliforms)	25	10	1416000	58552	282837	494
Recontamination assessment (*E. coli*)	25	0.5	3580	197	710	17

($p=.02$). The difference in the number of soap-use events recorded by the observer on Day 8 and Day 10 was not statistically significant ($p=1.00$). Sensor soap data was only available at the household level and, thus, we cannot report on variations in sensor soap use by the primary caregiver across the three observation days.

Study Question 4: Does hand contamination measured at random times predict hand contamination at times critical to pathogen transmission?

Hand rinses from primary caregivers were taken in a total of 55 households at random, in 49 households during one critical time, hereafter labeled as Critical Time 1, and in 24 households during a second critical time, hereafter labeled as Critical Time 2. The summary results, including the geometric means, are presented in Table 18. Levels of hand contamination, as measured by the number of colony-forming units (CFU)/100 ml of fecal coliforms and *E. coli*, were greatest in the Critical Time 1 sampling, compared to Random and Critical Time 2 sampling.

Hand rinses from 49 primary caregivers were taken for Critical Time 1 during the following opportunities for handwashing: food preparation (57%), feeding a child (18%), eating (2%), and drinking water (2%). In 10 (20%) households, no specific opportunity for handwashing was observed and, hence, the Critical Time 1 sampling was actually taken at random just before the completion of the structured observation. All households that had the Critical Time 1 sampling taken at random were in the abbreviated structured observation group. From the 49 primary caregivers that had Critical Time 1 sampling, the geometric mean of fecal coliforms taken at random was 371 CFU/100 ml, compared to 2785 CFU/100 ml during Critical Time 1. The geometric mean of *E. coli* taken at random was 24 CFU/100 ml, compared to 53 CFU/100 ml during Critical Time 1. There was no correlation between the log-transformed results of fecal coliforms ($R = .16$, $p=.26$) or *E. coli* ($R = -0.13$, $p = .38$) testing of the Random and Critical Time 1 samples (Figure 2, Figure 3).

We calculated the mean absolute difference between the log-transformed results of the # CFU/100 ml of fecal coliforms or *E. coli* at Random and at Critical Time 1 to examine the degree of difference between the two sample times. The mean absolute difference in log-transformed results of fecal coliforms between Random and Critical Time 1 sampling, was 3.5 (SD 1.5). The Bland-Altman plot demonstrates increasing difference between Random and Critical Time 1 sampling as the mean result increases (Figure 4). The mean absolute difference in log-transformed results of *E. coli* between Random and Critical Time 1 sampling, was 2.1 (SD 0.8) (Figure 5).

Hand rinses from 24 primary caregivers, all in the Extended Observation Group, were taken for Critical Time 2 during the following opportunities for handwashing: food preparation (42%), feeding a child (25%), eating (17%), and handling

FIGURE 2: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT RANDOM AND CRITICAL TIME 1 (FECAL COLIFORMS)

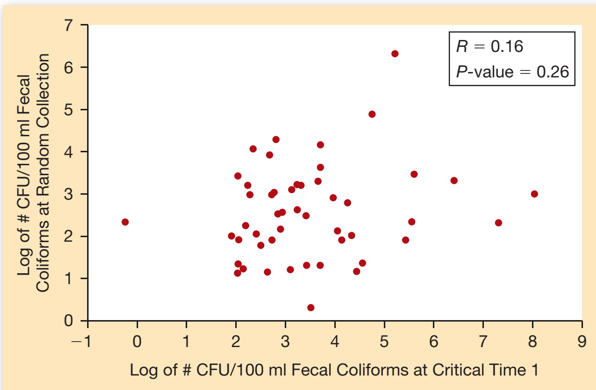


FIGURE 3: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT RANDOM AND CRITICAL TIME 1 (*E. COLI*)

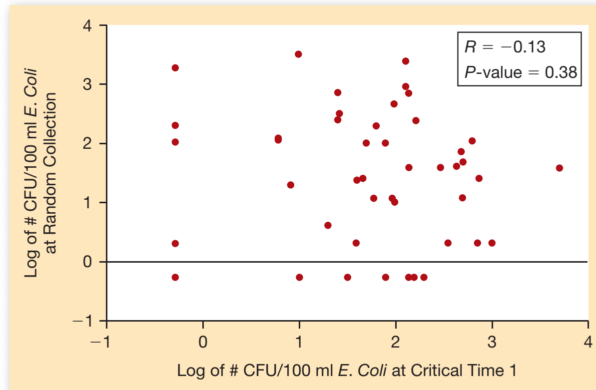


FIGURE 4: BLAND-ALTMAN PLOT OF LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT RANDOM AND CRITICAL TIME 1 (FECAL COLIFORMS)

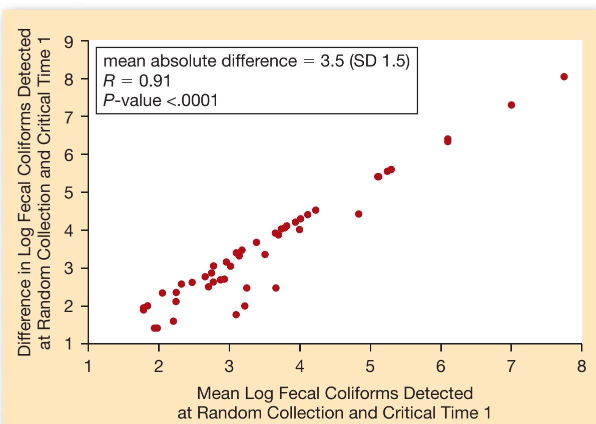
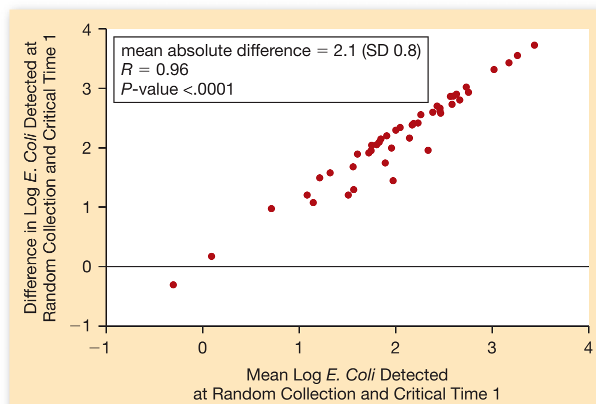


FIGURE 5: BLAND-ALTMAN PLOT OF LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT RANDOM AND CRITICAL TIME 1 (*E. COLI*)



water for storage (13%). In three (13%) households, no specific opportunity for handwashing was observed such that a second critical time hand rinse sample could be obtained, and thus, the Critical Time 2 sampling was taken at random just before the completion of the structured observation. From these 24 primary caregivers, the geometric mean of fecal coliforms taken at random was 233 CFU/100 ml, compared to 554 CFU/100 ml during Critical Time 2. The geometric mean of *E. coli* taken at random was 16 CFU/100 ml, compared to 28 CFU/100 ml during Critical Time 2. There was no correlation between the log-transformed results of fecal coliforms (R = −0.4, p = .84) or *E. coli* (R = .03, p = .90) testing of the Random and Critical Time 1 samples (Figure 6, Figure 7).

The mean absolute difference in log-transformed results of fecal coliforms between Random and Critical Time 2 sampling, was 2.9 (SD 1.4). The Bland-Altman plot demonstrates increasing difference between Random and Critical Time 2 sampling as the mean result increases (Figure 8). The mean absolute difference in log-transformed results of *E. coli* between Random and Critical Time 2 sampling was 1.9 (SD 0.9) (Figure 9).

FIGURE 6: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT RANDOM AND CRITICAL TIME 2 (FECAL COLIFORMS)

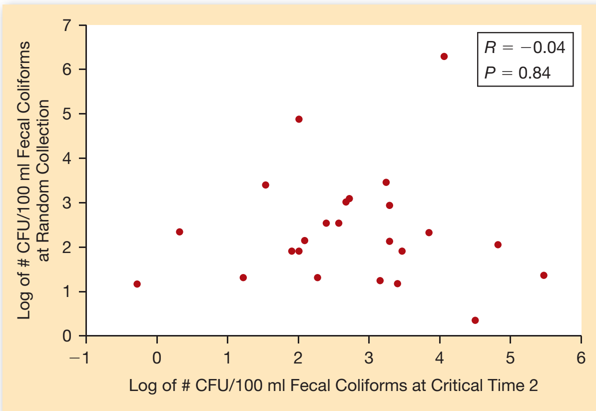


FIGURE 7: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT RANDOM AND CRITICAL TIME 2 (*E. COLI*)

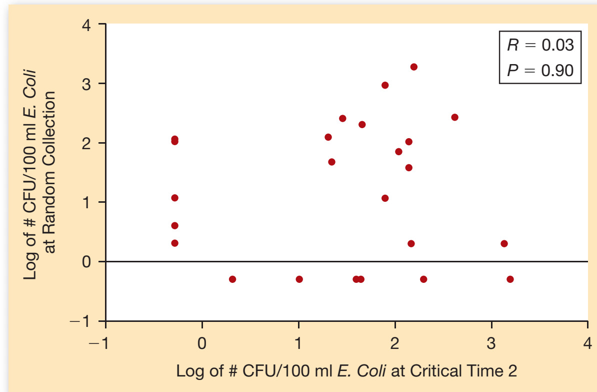


FIGURE 8: BLAND-ALTMAN PLOT OF LOG-TRANSFORMED RESULTS OF RANDOM AND CRITICAL TIME 2 SAMPLING (FECAL COLIFORMS)

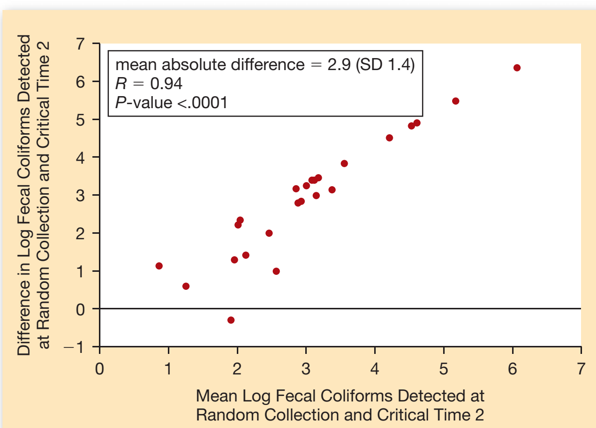
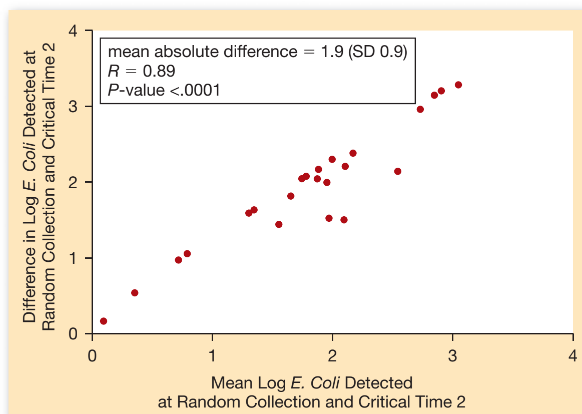


FIGURE 9: BLAND-ALTMAN PLOT OF LOG-TRANSFORMED RESULTS OF RANDOM AND CRITICAL TIME 2 SAMPLING (*E. COLI*)



In 24 households, we had the opportunity to compare results of hand rinses collected at Critical Time 1 and Critical Time 2. The frequencies of the specific opportunities for hand rinse sampling are detailed in Table 19. The geometric mean of fecal coliforms taken at Critical Time 1 was 1997 CFU/100 ml, compared to 554 CFU/100 ml during Critical Time 2. The geometric mean of *E. coli* taken at Critical Time 1 was 52 CFU/100 ml, compared to 28 CFU/100 ml during Critical Time 2. The correlation in the results of Critical Time 1 and Critical Time 2 was significant for fecal coliforms (R = .45, p = .03) and *E. coli* (R = .47, p = .02) (Figures 10 and 11).

The mean absolute difference in log-transformed results of fecal coliforms between Critical Time 1 and Critical Time 2 sampling, was 3.4 (SD 1.5) (Figure 12). The mean absolute difference in log-transformed results of *E. coli* between Critical Time 1 and Critical Time 2 sampling was 1.8 (SD 1.1) (Figure 13).

Next, we assessed whether observed use of soap by a primary caregiver during the structured observation was associated with reduced hand contamination. Of the various times at which hand rinse sampling was done, we found an

TABLE 19: FREQUENCIES OF OPPORTUNITIES FOR HAND RINSE SAMPLING FOR CRITICAL TIME 1 AND CRITICAL TIME 2

Opportunity for Hand Rinse Sampling	Critical Time 1 (N=24)	Critical Time 2 (N=24)
Food preparation	63%	42%
Feeding a child	29%	25%
Eating	4%	17%
Handling water for storage	0%	13%
Drinking water	4%	0%

FIGURE 10: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT CRITICAL TIME 1 AND CRITICAL TIME 2 (FECAL COLIFORMS)

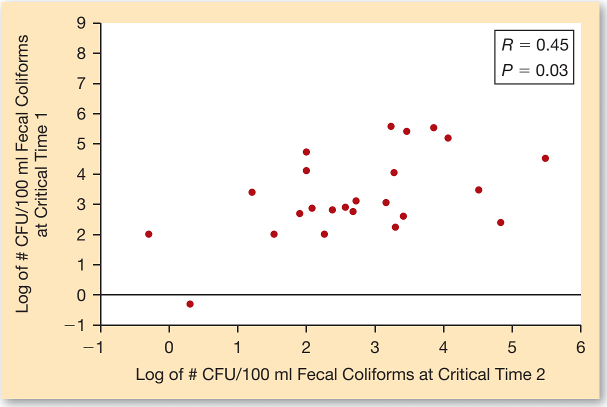


FIGURE 11: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT CRITICAL TIME 1 AND CRITICAL TIME 2 (*E. COLI*)

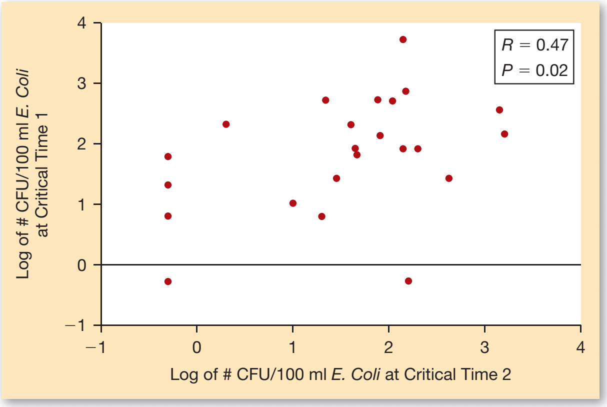


FIGURE 12: BLAND-ALTMAN PLOT OF LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT CRITICAL TIME 1 AND CRITICAL TIME 2 (FECAL COLIFORMS)

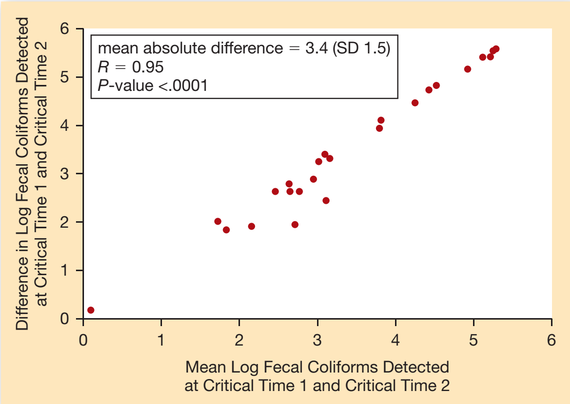


FIGURE 13: BLAND-ALTMAN PLOT OF LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT CRITICAL TIME 1 AND CRITICAL TIME 2 (*E. COLI*)

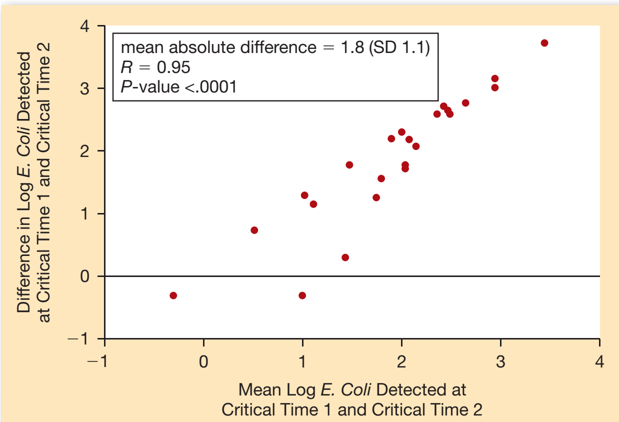


TABLE 20: MICROBIOLOGICAL CONTAMINATION OF THE PRIMARY CAREGIVER'S HANDS ACCORDING TO OBSERVATION OF HER SOAP USE DURING STRUCTURED OBSERVATION

	Soap Use Observed at Least Once		Soap Use Not Observed		
Variable	N	Geometric Mean (CFU/100 ml)	N	Geometric Mean (CFU/100 ml)	P-Value
Random (fecal coliforms)	20	406	34	300	0.96
Random (*E. coli*)	20	29	34	21	0.55
Time 1 (fecal coliforms)	18	1722	31	3682	0.27
Time 1 (*E. coli*)	18	77	31	43	0.3
Time 2 (fecal coliforms)	14	273	10	1497	0.28
Time 2 (*E. coli*)	14	9	10	142	0.01
Recontamination assessment (fecal coliforms)	4	117	21	650	0.21
Recontamination assessment (*E. coli*)	4	9	21	19	0.53

*p-value calculated based on Wilcoxon rank sum.

TABLE 21: MICROBIOLOGICAL CONTAMINATION OF THE PRIMARY CAREGIVER'S HANDS ACCORDING TO OBSERVATION OF HER SOAP USE AFTER A DEFECATION-RELATED EVENT DURING STRUCTURED OBSERVATION

	Soap Use Observed at Least Once		Soap Use Not Observed		
Variable	N	Geometric Mean (CFU/100 ml)	N	Geometric Mean (CFU/100 ml)	P-Value
Random (fecal coliforms)	11	942	10	65	0.07
Random (*E. coli*)	11	38	10	10	0.24
Time 1 (fecal coliforms)	9	4311	9	3628	1
Time 1 (*E. coli*)	9	141	9	23	0.17
Time 2 (fecal coliforms)	8	457	9	974	0.6
Time 2 (*E. coli*)	8	24	9	34	0.63

*p-value calculated based on Wilcoxon rank sum.

association between observed soap use and reduced hand contamination during Critical Time 2 (*E. coli*) (Table 20). There was no other association between observed soap use and hand contamination. The only association between observed soap use by the primary caregiver during a defecation-related event and hand contamination was in the Random sampling (fecal coliforms) (Table 21).

We found that there was no correlation between the mean number of per capita 24-hour sensor soap-use events during pre-observation days and the results of hand rinse sampling at Random, Critical Time 1, and Critical Time 2 (Table 22).

TABLE 22: CORRELATION BETWEEN LOG-TRANSFORMED RESULTS OF HAND RINSE SAMPLING AT VARIOUS TIMES AND 24-HOUR PER CAPITA SENSOR SOAP-USE EVENTS

	N	R	P-Value
Random (fecal coliforms)	21	0.13	0.57
Random (*E. coli*)	21	0.04	0.88
Time 1 (fecal coliforms)	20	0.09	0.71
Time 1 (*E. coli*)	20	−0.31	0.18
Time 2 (fecal coliforms)	20	0.03	0.92
Time 2 (*E. coli*)	21	0.15	0.53

*p-value calculated based on Wilcoxon rank sum.

Study Question 5: What is the rate of recontamination among the target population at two to three hours following a thorough handwashing?

We requested 25 mothers to perform thorough handwashing with soap under the supervision of our fieldworkers. At approximately two hours after the thorough handwashing, we found that all mothers had detectable fecal coliforms, with a geometric mean of 494 CFU/100 ml. There were no detectable *E. coli* on the hands of 5 (20%) mothers. The geometric mean of *E. coli* was 17 CFU/100 ml, with quartile cut-offs at 6 (25%), 22 (50%), and 76 (75%). The highest concentration of *E. coli* was 3580 CFU/100 ml.

We compared mothers in the highest quartile of *E. coli* contamination during the post-supervised handwashing sampling to mothers in the lower three quartiles and found no differences with respect to demographic factors, SES quartile, or self-reported or spot-check measures of handwashing behavior. Since the sensor soap was not deployed in the abbreviated observation group, from which mothers were requested to take part in the recontamination assessment, we do not have sensor soap data in order to assess for differences in soap use in this group.

4. Discussion

KEY POINTS

- **Using sensor soap, we found that many households altered their soap-use practices during the structured observation.**
- **There was substantial loss in the ability to observe handwashing behavior during defecation-related events among households observed for only 90 minutes.**
- **We found no correlations between levels of fecal coliform or *E. coli* contamination of hands at random and hands at critical times.**

This study addressed several questions that are fundamental to developing valid and useful measures of handwashing behavior. We integrated a novel technology, the sensor soap, into a comprehensive study of various methods of measuring handwashing behavior, including data collected using questionnaires for self-reported behavior, spot checks and hand microbiology for proxy measures, and structured observations for observed soap use.

Our data demonstrate that the number of daily sensor soap-use events per household does not vary significantly, suggesting that this is a reliable measure of daily household soap use. We found no significant difference between the number of uses of sensor soap on Day 1, the first full day that the sensor soap was in place, and the mean of the number of uses on subsequent pre-observation days. Although we found this to be true with adequate power in the group that had extended structured observations done four days after the sensor soap was placed in the home, we had less power to detect this equivalence in the group that had extended structured observations done eight days after the sensor soap was placed in the home. It will be important to confirm this lack of reactivity to the sensor soap in larger sample sizes. Still, based on the experience in group that had extended structured observations done four days after the sensor soap was placed, our data allow us to recommend that sensor soap be used to estimate daily household soap use in evaluating handwashing programs.

No measures of handwashing behavior, collected by self-report, spot checks, or structured observations were positively associated with per capita 24-hour sensor soap use on pre-observation days. The two significant relationships identified were contrary to expectation and logic, with fewer sensor soap-use events in households of mothers who washed their hands with soap before giving the interviewer a glass of water compared with those households where the mother did not wash hands with soap, and among households in the highest SES quartile compared with those in the second poorest SES quartile. There was a significant difference between the number of soap-use events detected by the sensor soap and the number recorded by the observer during the five-hour structured observation. This may be explained by the fact that sensor soap could have been used by any member of the household and that the observer was instructed to keep the primary caregiver as her main focus of observation. It is possible that the sensor soap is detecting events that do not represent true soap use for hand or body washing, but rather movement of soap for other purposes and, thus, overestimating soap use. The lack of significant positive relationships between expected predictors of good handwashing practices, and per capita daily sensor soap use, calls into question the utility of frequently used measures, such as self-report and spot checks. Alternatively, given that sensor soaps were used by all household members, and not just by primary caregivers, sensor soap data may not be the best comparison base for primary caregivers' self-reported handwashing practices. However, since sensor soaps were used by all household members, sensor soap data may represent an appropriate comparison base for spot-check measures that attempt to assess handwashing behavior at the household level, for example through the observation of soap at designated handwashing stations. It will be important to confirm these findings, as opportunities allow, through additional study among larger sample sizes.

Study Question 1: How reactive is a subject, with respect to HW behavior, when observed during an extended structured observation?

The sensor soap enabled us to identify those households in which the presence of an observer for structured observation resulted in substantial reactivity, or change in behavior from natural practices. The differences in the number of sensor soap-use events during the SO time block between pre-observation days and the observation day were substantial, and statistically significant, in both the group that had structured observation on the fourth full day after sensor soap placement and the group that had structured observation on the eighth full day after sensor soap placement. We do not yet have an explanation for why reactors in the group that had structured observation on the eighth full day after sensor soap placement had so much greater a mean difference (7.8 events) than reactors the group that had structured observation on the fourth full day after sensor soap placement (4.4 events). However, it is clear that sizable proportions of each group altered their soap-use practices in the presence of the observer for the structured observation. On univariate analysis, we found several predictors of reactivity in the groups that had structured observation on the fourth full day and the eighth full day after sensor soap placement. Consistent between the two groups was the finding that a greater proportion of reactors, than non-reactors, did not have to share a defecation place with other households, which was significantly associated with SES status in the larger group of 100 households (Table 2). Other notable predictors included owning a mobile phone and higher education status for the respondent and the head of household. Ownership of a mobile phone might be suggestive of early adopter behavior, in that these may be individuals who accept a new trend and thus may be receptive to handwash promotion campaigns, or may again be indicative of socioeconomic status. Higher education may result in increased awareness of the importance of hand hygiene, which may or may not translate into improved handwashing behavior. Despite the difference in education, and its potential impact on handwashing practices, we found no difference in sensor soap use during the SO Time Block on pre-observation days according to education. Thus, reactors altered their behavior and substantially increased their soap use in the presence of the observer but they were not naturally "better handwashers." Although reactivity was noted only in a subset of households, it would be difficult to detect which households were reactive and which ones allowed measurement of true behavior, without the simultaneous deployment of sensor soap.

Whereas sensor soap provides useful information regarding daily household-level soap use, it does not provide detail regarding context-specific handwashing behavior, such as after defecation or before feeding a child. Based on local beliefs or traditions, reactivity-dependent soap use may occur more or less frequently than usual in some opportunities. It is conceivable that mothers who believe that hands ought to be washed with soap after defecation or who live in a society where that would be a sign of good motherhood would increase post-defecation soap use while being observed more than mothers who either had no such knowledge or whose society did not encourage it. If mothers do not believe that it is important to wash hands before feeding a child, it is possible that they may not demonstrate reactivity-based increases in handwashing with soap before feeding a child during the course of a structured observation. Still, context-specific information regarding soap use and other handwashing behaviors is provided by structured observation and cannot be provided by sensor soap.

We found that about one-third of households were particularly reactive, with the rest having minimal difference in the number of soap-use events in the SO Time Block between the structured observation day and pre-observation days. Among those households that did not demonstrate reactivity in the presence of the observer, the structured observation would provide useful information regarding soap use and other handwashing practices during specific critical times.

Study Question 2: What is the optimum duration of a structured observation for the purposes of measuring HW behavior?

Although structured observation provides potentially useful context-specific information about handwashing practices, lengthy observation durations, ranging in the literature from three to seven hours, contribute to impractically high costs for evaluation of most handwashing promotion programs. In this study, we examined the data loss that might result from limiting the observation duration to 90 minutes, using a five-hour observation as the basis of comparison. At least one observation of any household member was possible in all 90-minute abbreviated observation households. Based on the ratio of five hours to 90 minutes (3.3), we would have expected to see a similar ratio in terms of the number of specific opportunities to observe handwashing behavior.

However, there was substantial loss of ability to observe handwashing behavior during defecation-related events, with the ratio of defecation-related events in extended versus abbreviated households at 5.3 among all household members. Among all primary caregivers, that figure was even higher (6.0), suggesting that mothers and other household members may have been reluctant to carry out defecation-related activities in the presence of the observer who was in the house for a limited period of time. In the extended observation households, it is conceivable that mothers became more comfortable with the presence of the observer over the course of the five-hour observation period and allowed themselves to carry out defecation-related activities. A more detailed analysis of the specific timing of defecation-related events would elucidate this further. Events that may not be perceived as private, such as water-handling, or that are time-sensitive, such as feeding a child, were more frequently observed than defecation-related events in the abbreviated observation households. Among primary caregivers, the ratio of extended to abbreviated households was greatest for observation of handwash behavior by the primary caregiver before eating. This may be culturally related, since mothers would be unlikely to want to eat when the observer, a guest in the home, has not yet eaten.

Observation of handwash behavior during specific critical times, such as after defecation or before eating, is important to understanding whether behavior varies across critical times. During the course of analysis, we hypothesized that if soap use were consistent across critical times, there may not be a need to observe each critical time. Rather, if a sufficient number of critical times could be observed, the mother's handwashing behavior during other critical times could be inferred. However, we found few consistencies in soap use across different critical times, for either any household member or for the primary caregiver. Such analysis should be attempted in various cultural settings, since handwashing behavior at different critical times may be mediated by local beliefs regarding purity and pollution associated with defecation, eating, and taking care of young children, as well as by exposure to hand hygiene promotion programs carried out in the area. Moreover, it is worth questioning and examining whether changing handwashing behavior at specific critical times has greater health impact than simply increasing overall use of soap for washing hands in the course of a day. Examining the effectiveness of different types of handwashing promotion campaigns (e.g., social marketing versus community-level mobilization) for changing handwashing behavior at specific critical times, versus simply increasing overall soap use, would be very relevant to this discussion. In handwashing promotion programs that aim to increase soap use overall, and that do not set out to change event-specific behaviors, examining handwashing behavior at critical times might not be relevant and, thus, shorter durations of structured observations would be acceptable.

Study Question 3: What is the change in soap-use behavior during repeated structured observations within the same household?

Previous authors[20] have suggested that conducting repeated structured observations in the same household might result in decreasing reactivity and, hence, greater validity of observational data collected after the first day or two. We did not find significant differences in soap-use behavior, as measured either by structured observation or by sensor soap, on the third day of serial observations, compared to the first day of serial observations. The utility of sensor soap data on the third day of serial observations is called into question because of the high likelihood that the soap component of the sensor soap may have been so worn down that it became unusable. Even among our households that only had the sensor soap in place for a total of five days, including the day of deployment, the loggers became visible in a majority. In these households that had a single sensor soap for a total of seven days, including the day of deployment, it is quite possible that there was very little soap left to use by household members. It is important to note that we were limited by a small sample size of nine households, but this limitation faces most under-funded handwashing promotion programs, which are unlikely to be able to afford serial observations in much larger sample sizes, particularly to demonstrate pre-post differences in handwashing behavior.

[20] Curtis et al. and Cousens et al.

Study Question 4: Does hand contamination measured at random times predict hand contamination at times critical to pathogen transmission?

Conducting structured observations allows the collection of hand rinse samples for estimating hand contamination at critical times. Given the costs and other logistical challenges associated with structured observations, we sought to examine whether hand contamination measured at random would be predictive of hand contamination at times when pathogens might be expected to be transmitted from hands to other persons or to food or water. We found no correlations between levels of fecal coliform or *E. coli* contamination of hands at random and hands at critical times. Moreover, we found large absolute differences between levels of contamination at random and at critical times, suggesting that microbiology testing of the same subject's hands can yield very variable results. The Bland-Altman plots demonstrated that these differences are particularly great when hands are heavily contaminated. Based on the evidence available to us currently, we cannot know whether the variability in results has implications for disease risk. We did find correlations between results of hand contamination of two critical time samples. Similarly, we found correlations between the results of the random hand rinse taken before the structured observation was begun and, the hand rinse taken two hours after a supervised thorough handwashing. Despite these correlations, the absolute differences between the two random samples, and the two critical time samples, were actually quite large, again pointing to the lack of repeatability of hand microbiology testing.

Variability in hand microbiology results may result from any of the following factors: duration since last use of soap or mud for handwashing, duration since last defecation or contact with a child's feces, and overall fecal contamination of the subject's environment. Although we have not examined factors related to duration since last fecal contact or overall fecal contamination of the subject's environment, we did examine the relationship between observed soap use, either by the interviewer during the structured observation or via the sensor soap (per capita 24-hour soap use during pre-observation days), and the level of contamination on the primary caregiver's hands and found no association. Thus, hand contamination did not appear to be associated with direct measures of soap use. Irrespective of the explanation of variable hand contamination in an individual, the degree of variability identified in this study calls into question the validity of a single random hand rinse sample or even critical time-based hand rinse sampling as proxy measures for the quality of handwashing behavior.

Study Question 5: What is the rate of recontamination among the target population at two to three hours following a thorough handwashing?

Following the thorough handwashing with soap, we found fecal coliforms on all respondents' hands but noted that a sizable proportion (20%) of primary caregivers did not have any detectable *E. coli* contamination. Moreover, 75% of primary caregivers had 76 or fewer CFU/100 ml. A study by Luby, et al. found that the risk of diarrheal disease was higher among households in which the primary caregiver had > 4000 CFU/100 ml fecal coliforms on hands, compared to households in which primary caregivers hand ≤ 4000 CFU/100 ml fecal coliforms on hands.[21] The disease risk associated with <100 CFU/100 ml of *E. coli* on hands, as detected in our study, is unclear. Further study may elucidate the health implications of various levels of hand contamination. The assessment of recontamination, and the degree of recontamination, may yet prove useful in measuring the quality of handwashing behavior. If the sensor soap could be kept in place in the home following a supervised thorough handwashing with soap, it might be possible to examine the relationship between the number of soap-use events detected following the thorough handwashing and the degree of contamination found on the hands several hours later.

Limitations

This study faced two principal limitations. First, the sample size was small and prevented full exploration of some of the study questions. However, the questions of greatest importance regarding reactivity to structured observations and whether results of random hand rinses could predict

[21] Luby, et al. 2007.

results of critical time hand rinses were answered with adequate power. The remaining questions were exploratory in nature and we did not anticipate having sufficient power to answer all of them. The second limitation was the intense flooding that occurred during the course of data collection for this study. Some households were enrolled after floods had occurred and this may have altered their risk perceptions and handwashing practices. Because of the limited sample size, we did not feel that we could exclude households enrolled post-flooding. Since this study is not about actual handwashing behavior, but rather the measurement of whichever behavior exists, we feel that it is valid to include post-flood households in the data analysis.

5. Conclusions

Despite the limitations outlined above, this study confirmed the utility of sensor soap, a novel technology for measuring soap use at the household level. We identified substantial reactivity to structured observation among a subset of households, with socioeconomic and education markers as potential explanatory markers for reactivity. Among those households that did not display reactivity, structured observation may provide useful information regarding context-specific soap use and other handwashing behaviors. Our findings call into question the validity of hand microbiology, given the great variability of results from the same person's hands over the course of just a few hours. Notably, we found little correlation between self-report, spot check, structured observation, or hand microbiology indicators of handwashing behavior, and sensor soap data on soap use.

Based on our findings, we have the following recommendations, in brief:

1) Evaluate utility and feasibility of sensor soap in large-scale research studies and evaluations of handwashing promotion programs.
2) Deploy sensor soaps among households participating in structured observation, in order to assess for reactivity. Among households found not to demonstrate reactivity, structured observation data should be analyzed for context-specific soap-use behaviors.
3) Confirm that knowledge regarding sensor soap construction, deployment, retrieval, and data download can be transferred speedily and accurately to in-country personnel in various countries. A major barrier to wider use of sensor soap in monitoring and evaluation of handwashing promotion programs is the current focus of knowledge in one institution, Unilever. As the sensor soap technology becomes more widely recognized, accepted, and sought after, it will be critical to transfer technology in order to make the sensor soap more widely available and usable to handwashing promotion program evaluators.
4) Confirm our findings regarding the lack of correlation between sensor soap data on per capita daily soap use and other measures of handwashing behavior. In particular, the relationship between spot checks and sensor soap data should be elucidated since both measure handwashing behavior at the household level. This is important to be done in larger sample sizes than ours in order to inform smaller-scale and lower budget handwashing promotion programs, as well as large nationally representative surveys such as the Demographic and Health Surveys and Multi-indicator Cluster Surveys. Such programs and surveys and need to be advised on how best to measure handwashing behavior in the absence of sensor soap and structured observation methods, which are at present, costly and person-intensive methods to measure handwashing behavior.
5) Examine consistency in soap use during different critical times for handwashing among households found not to be reactive to the presence of the observer. Such analysis can be done with the current dataset, and should be done when larger datasets with both sensor soap and structured observation data become available.
6) Examine the relationship between hand contamination following a supervised thorough handwashing with soap to soap use, as measured by the sensor soap during the time between the supervised handwashing and the collection of the hand rinse samples. This analysis is not possible in the current dataset because sensor soap was not deployed among those households that had the recontamination assessment.
7) Although not purely questions of handwashing measurement, the following question arose during our review of study findings: What is the relative impact on behavior, and if possible, health, of promoting handwashing with soap at specific critical times versus general promotion of handwashing with soap in order to increase overall soap use in a day?

References

Bentley, M., Boot, M., Gittelsohn, J. & Stallings, R. 1994. The Use of Structured Observations in the Study of Health Behaviour. The Hague, The Netherlands, IRC International Water and Sanitation Centre.

Bland, J.M. & Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet. 1* (8476): 307–10.

Cousens, S., Kanki, B., Toure, S., Diallo, I. & Curtis, V. 1996. Reactivity and repeatability of hygiene behaviour: structured observations from Burkina Faso. *Soc Sci Med. 43* (9): 1299–308.

Curtis, V., Cousens, S., Mertens, T., Trarore, E., Kanki, B. & Diallo, I. 1993. Structured observations of hygiene behaviors in Burkina Faso: validity, variability, and utility. *Bull WHO. 71* (1): 23–32.

Gorter, A.C., Sandiford, P., Pauw, J., Morales, P., Perez, R.M. & Alberts, H. 1998. Hygiene behaviour in rural Nicaragua in relation to diarrhoea. *Int J Epidemiol. 27* (6): 1090–100.

Hoque, B.A., Mahalanabis, D., Alam, M.J. & Islam, M.S. 1995. Post-defecation handwashing in Bangladesh: practice and efficiency perspectives. *Public Health. 109* (1): 15–24.

Kaltenthaler, E.C. & Pinfold, J.V. 1995. Microbiological methods for assessing handwashing practice in hygiene behaviour studies. *J Trop Med Hyg. 98* (2): 101–6.

Luby, S.P., Agboatwalla, M., Billhimer, W. & Hoekstra, R.M. 2007. Field trial of a low cost method to evaluate hand cleanliness. *Trop Med Int Health. 12* (6): 765–71.

Luby, S.P., Agboatwalla, M., Feikin, D.R., Painter, J., Billhimer, W., Altaf, A. et al. 2005. Effect of handwashing on child health: a randomised controlled trial. *Lancet. 366* (9481): 225–33.

Luby, S., Agboatwalla, M., Painter, J., Altaf, A., Billhimer, W. & Hoekstra, R. 2004. Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan: a randomized controlled trial. *JAMA. 291* (21): 2547–2554.

Luby, S.P., Agboatwalla, M., Raza, A., Sobel, J., Mint, E.D., Baier, K. et al. 2001. Microbiologic effectiveness of hand washing with soap in an urban squatter settlement, Karachi, Pakistan. *Epidemiol Infect. 127* (2): 237–44.

Manun'Ebo, M., Cousens, S., Haggerty, P., Kalengaie, M., Ashworth, A. & Kirkwood, B. 1997. Measuring hygiene practices: a comparison of questionnaires with direct observations in rural Zaire. *Trop Med Int Health. 2* (11): 1015–21.

Pinfold, J.V. 1990. Faecal contamination of water and fingertip-rinses as a method for evaluating the effect of low-cost water supply and sanitation activities on faeco-oral disease transmission. I. A case study in rural north-east Thailand. *Epidemiol Infect. 105* (2): 363–75.

Pinfold, J.V. and Horan, N.J. 1996. Measuring the effect of a hygiene behaviour intervention by indicators of behaviour and diarrhoeal disease. *Trans R Soc Trop Med Hyg. 90* (4): 366–71.

Ruel, M.T. & Arimond, M. 2002. Spot-check observational method for assessing hygiene practices: review of experience and implications for programmes. *J Health Popul Nutr. 20* (1): 65–76.

Rusin, P., Orosz-Coughlin, P. and Gerba, C. 1998. Reduction of faecal coliform, coliform and heterotrophic plate count bacteria in the household kitchen and bathroom by disinfection with hypochlorite cleaners. *J Appl Microbiol. 85* (5): 819–28.

Sobel, J., Mahon, B., Mendoza, C.E., Passaro, D., Cano, F., Baier, K. et al. 1998. Reduction of fecal contamination of street-vended beverages in Guatemala by a simple system for water purification and storage, handwashing, and beverage storage. *Am J Trop Med Hyg. 59* (3): 380–7.

Stanton, B.F., Clemens, J.D., Aziz, K.M. & Rahman, M. 1987. Twenty-four-hour recall, knowledge-attitude-practice questionnaires, and direct observations of sanitary practices: a comparative study. *Bull World Health Organ. 65* (2): 217–22.

Strina, A., Cairncross, S., Barreto, M.L., Larrea, C. & Prado, M.S. 2003. Childhood diarrhea and observed hygiene behavior in Salvador, Brazil. *Am J Epidemiol. 157* (11): 1032–8.

Vyas, S. and Kumaranayake, L. 2006. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy Plan. 21* (6): 459–68.

water and
sanitation program

SANITATION FOR ALL