Biran A, Rabie T, Schmidt W, Juvekar S, Hirve S, and Curtis V (2008). Comparing the Performance of Indicators of Hand-Washing Practices in Rural Indian Households. *Trop Med Int Hlth*. 13 (2):278-285

Abstract

Objectives: To compare the results obtained from 26 proxy indicators of domestic hand-washing practices with those obtained from direct, 'structured' observation of hand-washing in a sample of 387 households and to assess the potential of these indicators for use in the evaluation of hygiene promotion campaigns.
Methods: Fieldwork was carried out in a rural Indian setting between February 2005 and April 2006. Household-level data on hand-washing practices and the availability of soap and water were collected using 5 different data collection techniques. These were structured observation, questionnaire survey, pocket voting, hand-wash demonstration and an environmental check. Between them these techniques produced 27 binary indicators of hand-washing practices, each of which was used to classify households as 'hand-washing' or 'non-hand-washing. The Kappa statistic was used to assess the extent to which household classification based on each of 26 proxy indicators concurred with classification based on observation. The prevalence of households defined as 'hand-washing' according to each indicator was compared statistically with the prevalence according to structured observations by testing for a significant difference between two proportions
Results: Agreement between all the proxy indicators and the observation data was poor and all but 2 of the indicators produced estimates of hand-washing prevalence that were significantly different from that resulting from observation.
Conclusions: Although some interventions may be able to use proxy indicators as a guide to the magnitude and direction of their impact, these indicators do not provide an accurate guide to the actual practice / prevalence of hand-washing. Structured observation remains the best indicator of those tested.

Introduction

Diarrhoeal disease is a major public health problem that was estimated to have caused 2.2 million child deaths around the globe in 2000 (Kosek *et al.* 2003). Several reviews conclude that diarrhoeal infections can be prevented by the promotion of safe hygiene practices (Esrey *et al.* 1985, Huttly *et al.* 1997). A review of the available evidence concluded that amongst hygiene practices, hand-washing with soap in particular, could substantially reduce the risk of diarrhoeal infection (Curtis and Cairncross 2003). Recently, Luby *et al.* reported a 53% reduction in diarrhoea risk in a trial of a hand-washing intervention in a poor setting in Pakistan (Luby *et al.* 2004). There is mounting evidence that hand-washing with soap can also be effective in preventing respiratory infection (Luby *et al.* 2005; Rabie and Curtis 2005).

Whilst the promotion of hand-washing in the domestic setting has nominally figured in public health efforts around the world for many years, rates remain very low (typically 2-10% (Scott *et al.*. 2003)) and there is a growing realisation that this issue deserves more focused attention and investment. Programmes such as the Public-Private Partnership for Hand Washing with Soap (www.globalhandwashing.org), the

Hygiene Improvement Project (www.HIP.watsan.net) as well as work by UNICEF and others seek to substantially increase rates of hand-washing with soap among child carers, children and their families in resource poor settings around the world.

However, such programmes have great difficulty in measuring the impact they have had on the targeted behaviours. There is no simple, easily employed and reliable indicator of whether hands have been washed with soap at critical times or not (Curtis *et al.* 1993). A number of techniques have been employed, but all have drawbacks. Self-reports of hand-washing practices in response to verbal questioning produce rates that are inflated (Curtis *et al.* 1993, Cousens *et al.* 1996, Gittelson *et al.* 1997). Microbiological indicators of faecal contamination on hands are another alternative but they are expensive and difficult to execute in field settings, and the results are hard to interpret (Kaltenthaler *et al.*. 1991; Traoré *et al.*. 1994; Pinfold and Horan 1996).

Direct observation of hygiene practices (structured observation) is thought to be the most valid and reliable method currently in use for measuring hygiene behaviour (Cousens *et al.*. 1996). However, structured observation is extremely labour intensive. One fieldworker can only collect data on one household in one day. The method is also intrusive and tiring for the observer and for the observed and it requires well trained, supervised and motivated fieldworkers. If an alternative tool could be found that produces comparable results to structured observation this would be of great benefit in evaluating hygiene promotion interventions. In this paper we report the results of a study carried out in rural India which compares the performance of a wide variety of indicators of hygiene behaviour.

Methods

The study was conducted between April 2005 and February 2006 in the 22 villages in Pune District in the State of Maharashtra, India.

A simple random sample of 700 households was selected from a recent comprehensive census of the study site. To be eligible for recruitment a household had to have at least one child under the age of 18 months (the index child). Informed consent was sought from all 700 households. Households planning to emigrate from the study area within the period of the study were excluded, as well as those who, on visiting, were found not to fulfil the selection criteria. This left a total of 542 households.

Data collection was carried out by a team of 28 fieldworkers recruited from the villages in which the study was carried out. All had completed at least primary education and were able to speak and read Marathi. Fieldworkers received both in-class and in-field training in which they read, discussed and commented on the study consent form and the questionnaire and conducted mock practice interviews amongst themselves. Each fieldworker then carried out four supervised observation sessions of two and half hours duration over two days. During these observation sessions fieldworkers worked in houses in close proximity to each other allowing four fieldworkers to be supervised by one field supervisor. The four field supervisors had been trained in a similar way by one of the authors (A Biran). The training households were not part of the study sample.

Data relating to domestic hand-washing practices were collected using five different data collection techniques; structured observation, questionnaire survey, hand-wash demonstration, environmental check and pocket voting. This resulted in a total of 27 indicators of hand-washing practice. Each of the indicators could be scored in a binary fashion with a score of 1 indicating a hand-washing household and a score of 0 indicating a non-hand-washing household. The data collection techniques are described below and are summarised in table 1.

Structured observation:
Structured observation was carried out by means of a fieldworker sitting within a study household. Householders continued their normal activities and interaction between the fieldworker and householders was kept to a minimum. During the observation period the fieldworker observed and recorded the hand-washing practices of the primary caregiver following potential faecal contact. Householders were told that the fieldworker was observing child care and domestic activities but were not told the precise behaviours of interest.

The primary caregiver was defined as the person with primary responsibility for cleaning and feeding the index child as identified by the household. In most cases (93%) this was the child's mother. The primary caregiver was observed continuously while in, their house or compound. Observation was temporarily discontinued if the primary caregiver left the compound and went further away from the house. It was apparent when a carer went for defecation because she would take with her a small pot of water for anal cleansing. On her return after defecation, hand washing behaviour was observed.

All occurrences of hand-washing behaviour and the use of soap by the primary caregiver following potential faecal contact events were recorded. These events were defined as 'use of a latrine', 'use of a defecation site' and 'cleaning a child's bottom after defecation by the child'. After every potential faecal contact event the method of hand-washing (both hands, one hand, neither hand) and the cleansing material employed (soap, ash, mud, stone, water alone) were recorded.

Fieldworkers carried out observation between the hours of 06:00 am and 11:00 am. The timing and duration of observation were determined by practical considerations. Observation is tiring for the observer and the observed. For this reason, day-long observation was not possible and a 5-hour sample period was used. The time of day was chosen to maximise the likelihood of observing the events of interest since defecation and ablutions often occur in the early morning. An earlier start was not possible because of the logistical problems of getting the observers to households in the very early hours of the morning. The analysis assumes that the hand-washing practices of caregivers during the observed time period is representative of their usual practices.

Questionnaire survey
A questionnaire was administered verbally by fieldworkers on an occasion separate from the structured observation. This was usually done at the time of recruitment and

was always done few days prior to observation. There was no set time of day for administering the questionnaire. Respondents were female household members aged at least 16 years. Almost all respondents were mothers of the household index child. Thus the majority of respondents were the same individuals as those observed. Table 1 shows the hand-washing indicator questions that were asked.

Environmental check

The environmental check comprised a rapid visual inspection of the house and compound for the presence of soap and water. Environmental checks were carried out after completion of the structured observations. Fieldworkers checked three locations; in the house (houses were small and simple generally comprising only one or two rooms), in the compound and beside the latrine (if present). Within each location fieldworkers recorded whether there was a functioning water source, whether there was running water (i.e. could the water source be used by one person to create a running stream under which both hands could be washed simultaneously) and whether there was soap located in the vicinity of a functioning water source.

Hand-wash demonstration

After completion of the observation and environmental check fieldworkers asked the primary caregiver to demonstrate how she usually washed her hands following defecation. Fieldworkers recorded the method of hand-washing (both hands, one hand, neither hand), the cleansing material employed (soap, ash, mud, stone, water alone) and the presence or absence of soap at the hand-washing site.

Pocket voting

Pocket voting was developed as a participatory tool for assessing hygiene and sanitation practices at community level (Almedom 1997). Recent work in India demonstrated that the tool can be useful for investigating hand-washing practices in households (Cairncross et al 2005). Because the technique requires additional time, resources and fieldworker training it was used to collect information on hand washing behaviours on a sub-sample of 150 households on a separate 3rd occasion after the other data gathering techniques had already been applied. A convenience sample was weighted to reflect the relative contribution of different villages to the main sample.

Pocket voting participants were presented with a square matrix of 20 opaque ‘pockets’ (the pocket vote chart). Along the top of the chart was a series of pictures illustrating five options for washing hands (water, stone, ash, soap or not washed). Running up the side of the chart were illustrations of five different events that might trigger hand-washing (defecation, food preparation, eating, cleaning a child’s bottom and agricultural work). Participants were asked to ‘vote’ for their usual hand-washing practice for each trigger-event by placing a counter in the corresponding pocket. The chart was hung in such a way that neither the researchers nor the other householders could observe the actual placing of counters. This ensured that subjects’ responses remained anonymous.

The activity was explained to the householders, the chart was set up and all of those householders over the age of 10 who were present at the time of the visit voted in turn. Once all those present had voted the researchers took the chart and left the house, noted the responses and continued to the next house. Different coloured

counters were used for men, women and children to allow the data for women to be analysed separately.

Analysis

The purpose of the analysis was to compare the results obtained from the different indicators with those obtained from structured observation.

Indicators were compared with structured observation using the Kappa statistic. This statistic gives a measure of the extent to which two sets of ratings coincide over and above what would be expected by chance alone. A Kappa score of 0 indicates that agreement is no better than chance. A Kappa of 1 indicates a perfect match. Kappa scores between 0 and 1 indicate intermediate levels of agreement to which the following qualitative labels have been attached. Scores of >0-0.2 are considered to indicate slight agreement, 0.2-0.4 fair agreement, 0.4-0.6 moderate agreement, 0.6-0.8 substantial agreement and 0.8-1 near perfect, to perfect agreement (Viera and Garret 2005).

We also carried out a statistical comparison of the prevalence of 'hand-washing' households according to each indicator with the prevalence according to structured observations by testing for a significant difference between two proportions (Kirkwood 1988).

We decided a priori that to be considered an adequate indicator of hand-washing prevalence an indicator should show at least moderate agreement with structured observation as indicated by a Kappa score $\geq$ 0.4 and should show a prevalence of hand-washing not significantly different to that shown by structured observation.

The 26 proxy indicators of hand-washing each resulted in a binary score for each household. Households were rated as either hand-washing or non-hand-washing. However, structured observation resulted in a proportion of occasions on which hand-washing with soap occurred. For the analysis this proportion had to be converted into a binary rating of 'hand-washing' or 'non-hand-washing' for each household.

We classified households as 'hand-washing' if *all* primary caregiver, potential faecal contact events were followed by hand washing with soap and 'non-hand-washing' if not *all* potential primary caregiver faecal contact events were followed by hand washing with soap. Inconsistent hand-washing with soap was thus classified as non-hand-washing. These definitions of hand-washing and non-hand-washing were chosen because from a public health perspective hands should be washed following all potential faecal contact (Curtis et al 2000).

Households in which no potential faecal contact events were observed during structured observation were excluded from the analysis since there were no data with which to classify their hand-washing practices. This left a total of 387 households in the analysis.

Of these 387 households, data from the environmental check and hand-wash demonstration were available for all 387, verbal questionnaire data on hand-washing behaviour were available for 364 households and pocket vote data were available for 105 households. In 23 households the primary caregiver was not available to complete

the questionnaire survey and in 45 of the 150 households where pocket voting took place no potential faecal contact events were observed during structured observation.

We repeated our analysis using a dataset restricted to only those households in which hand-washing behaviours were consistent (i.e. 'hand-washing' households were those in which *all* primary caregiver potential faecal contact events were followed by hand washing with soap and 'non-hand-washing' households were those in which *no* primary caregiver potential faecal contact events were followed by hand washing with soap). This left a total of 330 households in the analysis (57 households in which the primary caregiver washed hands on some but not all occasions were excluded).

Results

On the basis of structured observation 105 households out of 387 (27%) were classified as 'hand-washing' households. The other 26 indicators reported prevalences of 'hand-washing' households ranging from 1% to 99%. The majority of indicators (19 out of 26) significantly overestimated the prevalence of hand-washing. Interestingly self-reported hand-washing with soap provided an underestimate. Two of the indicators (*observed soap beside the latrine* and *observed soap in the yard*), both from the environmental check tool, found prevalences of hand-washing that were similar to that found by structured observation (30% and 27% respectively). The prevalence of hand-washing found by all other indicators was significantly different from that found by structured observation.

None of the Kappa scores showed strong or even moderate agreement with the results of structured observation. The use of soap during the hand-wash demonstration showed a significant, fair agreement with structured observation (Kappa=0.2). However, the prevalence of soap use during the hand-wash demonstration was significantly higher at 67% than the prevalence of hand-washing as classified by structured observation (27%).

The results of this analysis are shown in Table 2.

Results for the restricted analysis gave slightly different prevalences due to the smaller number of non-hand-washing households included. However in all other respects these results were the same as those obtained from analysis of all households.

Discussion

A striking finding of this study is that all of the indicators that were tested performed poorly when compared with structured observation. No indicators achieved moderate or strong agreement with structured observation as assessed by their kappa scores, only one showed fair agreement and all but two gave a substantially different figure for handwashing prevalence. The implication is that none of those tested indicators provides a good guide to the actual prevalence of handwashing as assessed by structured observation in this setting. It might, however be possible to use proxy indicators as a guide to the direction and magnitude of programme impact, if not to the actual resultant changes in behaviour,

However, although this study treated structured observation as if it were a gold standard, this technique has a number of drawbacks. It is easier to use in some cultural and physical settings than others and the results are likely to be sensitive to the quality and supervision of field workers. Cousens *et al.* (1996) found that even within the same cultural setting, different behaviours and different individuals were influenced in different ways and to varying degrees by the presence of an observer.

A further problem with structured observation lies in its repeatability. The current study classified households as hand-washing or non-hand-washing on the basis of a single structured observation session with the primary caregiver and compared the performance of other indicators in terms of their agreement with this classification. Cousens *et al.* (1996) found that structured observation of mothers had poor repeatability at household level although at a population level the prevalence of hand-washing remained fairly constant across successive rounds of observation. The implication of this is that the majority of mothers are partial hand-washers who wash their hands with soap following some, but not all potential faecal contact events and that several observation sessions are needed to capture true hand-washing rates. Thus it is likely that, had our study carried out repeated observation, progressively fewer households would have been found to always practice hand-washing with soap and similarly progressively fewer to never practice hand-washing with soap and that most carers would have been observed to wash their hands with soap on some occasions.

Indicators of programme impact need to be robust enough to measure real changes in behaviour and not those which a household wishes to demonstrate to an observer having been exposed to the programme (courtesy bias). In the absence of a valid biological screening test for hand-washing behaviour, most of the above techniques are susceptible to this problem,

In contrast to clinical screening tools it is likely that the tools for evaluating hand-washing are influenced by the cultural context in which they are applied. Differences in the results of self-reported hand-washing behaviour were discussed above. Differences have also been reported in the validity of household-level pocket voting. This method was thought to work well in the Indian state of Kerala (Cairncross *et al.* 2005) but was found to perform poorly when used in Uganda (Shordt pers com Dec 2006). The background availability of convenient soap and water within households is also likely to vary across different locations and therefore these indicators will vary in the extent to which they point to a household having taken steps to facilitate their own hand-washing. Thus results found in Pune would not necessarily be replicated elsewhere and comparisons of hand-washing prevalence across different populations must be interpreted with care.

The indicator ‘reported soap use after potential faecal contact’, from the questionnaire tool resulted in significant under-reporting of soap use after faecal contact. This is a surprising result since self-report generally produces an over-estimate of ‘good practice’ as respondents, consciously or not, try to please the interviewers and/or portray a good image of themselves (e.g. Cousens *et al.* 1996). Why our study found under-reporting of ‘good practice’ is not clear. It may be because respondents did not easily recall all soap use or perhaps because they were hesitant about spontaneously discussing faecal contact events with an interviewer. This raises the possibility that the performance of the indicator could be improved with additional interviewer

training. However, use of this indicator in a small number of unpublished pilot interviews in another part of India by one of the authors (A Biran) found that on some occasions a considerable amount of probing and repeated questioning was needed to uncover hand-washing events and the reasons behind them (i.e. whether or not there may have been faecal contact prior to hand-washing). Thus the interview necessary to obtain good data from this indicator may not be sufficiently quick and simple for enumerators to implement on a large scale. Furthermore, self-report may be more likely to result in over-estimates of good practice *following* an intervention as the intervention is likely to sensitise respondents to the interviewer's notion of 'good practice'.

Conclusion

This study found no indicators with the potential to replace structured observation as the preferred measure of hand-washing behaviour. Previous work (Cousens et al 1996) has found problems relating to the repeatability of results obtained by this method and there are concerns over the susceptibility of this technique to a post-intervention courtesy bias. However, the results of structured observation remain the most straightforward to interpret, and though problematic this method still gives the best indication of hand-washing practices.

The results of the study call into question the idea that a single, universally applicable indicator of hand-washing practice is an achievable goal. Structured observation itself is likely to be influenced by context. The various proxy indicators tested are also likely to be context dependent and their relationships to true hand-washing rates are not known. This implies that comparisons of hand-washing rates across different settings should be interpreted with a good deal of caution and common sense since reported, detected or observed rates are not necessarily actual rates. It also implies that interventions need to develop indicators of success or combinations of indicators that are applicable to their specific context and content and that although these may provide an indication of the direction and magnitude of change achieved they will not necessarily provide accurate figures for hand-washing prevalence.

References

Almedom A.M, Blumenthal U & Manderson L (1997) *Hygiene Evaluation Procedures*. International Nutrition Foundation for Developing Countries.

Cairncross S, Shordt K, Zacharia S *et al.* (2005) What causes sustainable changes in hygiene behaviour? A cross-sectional study from Kerala, India. *Social Science and Medicine* **61**, 2212-2220.

Cousens S, Kanki B, Toure S *et al*. (1996) Reactivity and repeatability of hygiene behaviour: structured observations from Burkina Faso. *Social Science and Medicine* **43**, 1299-1308.

Curtis V, Cairncross S and Yonli R (2000) Domestic hygiene and diarrhoea – pinpointing the problem. Tropical Medicine and International Health, **5**, 22-32, January.

Curtis V, Cousens S, Mertens T *et al*. (1993) Structured observations of hygiene in Burkina Faso: validity, variability and utility. *Bulletin of the World Health Organization* **71**, 23-32.

Curtis V & Cairncross S (2003) Effect of washing hands with soap on diarrhoea risk in the community: a systematic review. *Lancet Infectious Diseases* **3**, 275-281.

Esrey SA, Feachem RG & Hughes JM (1985) Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities. *Bulletin of the World Health Organization* **63**, 757-772.

Gittelson J, Shankar AV, West KPJ *et al*. (1997) Estimating reactivity in direct observation studies of health behaviours. *Human Organisation* **56**, 182-189.

Huttly SRA, Morris SS & Pisani V (1997) Prevention of diarrhoea in young children in developing countries. *Bulletin of the World Health Organization*, **75**, 163-174.

Kaltenthaler E, Waterman R & Cross P (1991) Faecal indicator bacteria on the hands and the effectiveness of hand-washing in Zimbabwe. *Journal of Tropical Medicine and Hygiene* **94**, 358-363.

Kirkwood B (1988) *Essentials of medical statistics*. Blackwell, Oxford.

Kosek M, Bern C & Guerrant R. (2003) The magnitude of the global burden of diarrhoeal disease from studies published 1992-2000. *Bulletin of the World Health Organization* **81**, 197-204.

Luby S, Angboatwalla M, Painter J *et al*. (2004) Effect of Intensive Handwashing Promotion on Childhood Diarrhea in High-Risk Communities in Pakistan: A Randomized Controlled Trial. *Journal of the American Medical Association* **291**, 2547-2554.

Luby S, Agboatwalla M, Feikin D *et al*. (2005) Effect of handwashing on child health: a randomised controlled trial. The Lancet **366**, 225-233.

Pinfold JV, & Horan NJ (1996) Measuring the effect of a hygiene behaviour intervention by indicators of behaviour and diarrhoeal disease. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **90**, 366-371.

Rabie T & Curtis V (2005). Handwashing and Risk of Respiratory Infections: A Quantitative Systematic Review. *Tropical Medicine and International Health* **11**, 269-278.

Scott B, Curtis V & Rabie T (2003) Protecting children from diarrhoea and acute respiratory infections: the role of handwashing promotion in water and sanitation programmes. *WHO Regional Health Forum* **7**, 42-47.

Traoré E, Cousens S, Curtis V *et al*. (1994) Child defecation behaviour, stool disposal practices and childhood diarrhoea in Burkina Faso: results from a case-control study. *Journal of Epidemiology and Community Health* **48**, 270-275.

Viera JA and Garret JM (2005) Understanding interobserver agreement: the Kappa statistic. *Family Medicine* **37**, 360-3.

Acknowledgements
The study was a collaborative effort of the Hygiene Centre, London School of Hygiene and Tropical Medicine and the Vadu Rural Health Programme and KEM Hospital Research Centre, Pune. The authors acknowledge the contribution of the Vadu DSS (member of the INDEPTH Network) which made available its human and data resources for the study. Sally Pearson designed and produced the materials used to collect the pocket voting data and supervised the collection of these data. The study was funded by USAID and commissioned by the WHO. Lastly the authors acknowledge the cooperation of the Vadu area population for allowing our field investigators to intrude and observe hygiene related behaviours in their homes.

Statement of potential conflict of interest
Four of the authors (Biran, Curtis, Rabie and Schmidt) were partly funded by Unilever. This constitutes a potential conflict of interest since Unilever is a major commercial manufacturer of soap products and a promoter of hygiene practices.

Table 1: Summary of tools and indicators

Tool	*Instruction to enumerator* / question to respondent	Indicator	Scoring (1 = hand-washing household)
Structured observation	*Continuous observation and all occurrences recording for 3 hours, early morning.*	Primary caregiver observed washing both hands with soap following all potential faecal contact events occurring during observation period	Yes =1
Questionnaire survey	Do you have a functioning tap, pump or well in your house or compound? (Multiple answers acceptable)	Reported water source in house	Yes, in house = 1
	Do you have a functioning tap, pump or well in your house or compound? (*Multiple answers acceptable*)	Reported source water in yard	Yes, in yard = 1
	Do you have soap in your house today?	Reported soap present	Yes = 1
	Can you please bring me your soap?	Soap brought	Soap brought = 1
	Record time needed to bring soap.	Soap brought within 1 minute	Soap brought within 1 minute = 1
	Did you use soap today or yesterday?	Reported used soap today or yesterday	
	When you used soap today or yesterday what did you use it for? *Ask respondent, if she used soap today or yesterday what did she use it for.* *After each answer prompt by asking 'Was there anything else?'* *Don't specifically ask about any named activity. Record whether or not they are mentioned by the respondent.* *If respondent says they used soap for washing their hands as why they washed their hands.*	Reported soap use after potential faecal contact	Yes = 1 1 = respondent reports having used soap for washing their hands after defecating or after cleaning a child following defecation.
Hand-wash demonstration by primary caregiver	Please show me how you usually wash your hands following defecation.	Both hands washed with soap	Yes = 1
		Soap used	Yes = 1
		Soap available at demonstration site prior to demonstration	Yes = 1

Tool	*Instruction to enumerator* / question to respondent	Indicator	Scoring (1 = hand-washing household)
Pocket vote		Reported soap use for washing hands following defecation and cleaning a child's bottom by all present adult female householders	Yes = 1
Environmental check	*Carry out a spot-check for the presence of soap and water in the yard, house and by the latrine*	Stored water seen beside the latrine. Stored water seen in the house Stored water seen in the yard Stored water seen beside the latrine or in the house or in the yard Soap seen beside the latrine Soap seen in the house Soap seen in the yard Soap seen beside the latrine or in the house or in the yard Running* water seen beside the latrine. Running* water seen in the house Running* water seen in the yard Stored water seen beside the latrine or in the house or in the yard Pump, tap or well seen in the house Pump, tap or well seen in the yard Pump, tap or well seen in either house or yard	Yes = 1

* 'Running water' is defined as having the potential for one person to create a stream of water under which both hands could be washed simultaneously.

Table 2: Summary of results

Tool	Indicator	Prev. %	P (diff. between 2 proportions)	Kappa	95% CI lower	95% CI upper	Agree
Structured Observation	Observed both hands washed with soap by primary caregiver after all faecal contact events	27	-	-	-	-	-
Survey Questionnaire	Reported soap present	99	<0.0001	0.0097	-0.00461	0.024008	Slight
	Observed soap brought	98	<0.0001	0.0175	-0.00171	0.036708	Slight
	Observed soap brought within 1 minute	9	<0.0001	0.0335	0.00704	0.05996	Slight
	Reported used soap today or yesterday	99	<0.0001	0.0077	-0.00504	0.02044	Slight
	Reported soap use after potential faecal contact	20	0.018	0.0995	-0.00164	0.200636	Slight
	Reported water source in house	12	<0.0001	0.0902	-0.00133	0.181732	Slight
	Reported water source in yard	72	<0.0001	-0.0309	-0.09793	0.036132	Slight
Hand-wash Demonstration	Demonstrated hands washed with soap	1	<0.0001	-0.0153	-0.04352	0.012924	Slight
	Demonstrated soap used	67	<0.0001	0.2006	0.128668	0.272532	Fair
	Observed soap available at handwash site	78	<0.0001	0.1169	0.057904	0.175896	Slight
Pocket Vote	Reported soap use after potential faecal contact vote	52	<0.0001	0.1517	-0.02705	0.330452	Slight
Environmental check	Observed water by latrine^{+}	66	<0.0001	-0.0225	-0.17695	0.131948	Slight
	Observed water in house	86	<0.0001	-0.0063	-0.05334	0.04074	Slight
	Observed water in yard	84	<0.0001	-0.0292	-0.078	0.019604	Slight
	Observed water anywhere	100	<0.0001	-0.0052	-0.01147	0.001072	Slight
	Observed soap beside latrine^{+}	30	0.542	0.1312	-0.07617	0.338568	Slight
	Observed soap in house	58	<0.0001	0.0092	-0.07214	0.09054	Slight
	Observed soap in yard	27	0.935	-0.0553	-0.15487	0.044268	Slight
	Observed soap anywhere	94	<0.0001	0.0120	-0.02093	0.044928	Slight
	Observed running water by latrine^{+}	45	<0.0001	0.0830	-0.11143	0.277432	Slight
	Observed running water in house	17	<0.0001	0.0563	-0.03935	0.151948	Slight
	Observed running water in yard	58	<0.0001	-0.0222	-0.10334	0.058944	Slight
	Observed running water anywhere	66	<0.0001	-0.0311	-0.10421	0.042008	Slight

	Observed water source in house	15	<0.0001	0.0496	-0.0435	0.1427	Slight
	Observed water source in yard	61	<0.0001	-0.0497	-0.1279	0.028504	Slight
	Observed water source house or yard	66	<0.0001	0.0371	-0.03562	0.109816	Slight

[+] These data were only collected for those households that had latrines.

Note: H_0: Difference=0, therefore P-values more than 0.05 denote no evidence of statistically significant difference (fail to reject the null hypothesis) and those less than 0.05 denote statistically significant difference (reject the null hypothesis)