International Journal of Environmental Health Research

ISSN: 0960-3123 (Print) 1369-1619 (Online) Journal homepage: http://www.tandfonline.com/loi/cije20

Comparison of structured observation and pictorial 24 h recall of household activities to measure the prevalence of handwashing with soap in the community

WP Schmidt, HE Lewis, K Greenland & V Curtis

To cite this article: WP Schmidt, HE Lewis, K Greenland & V Curtis (2018): Comparison of structured observation and pictorial 24 h recall of household activities to measure the prevalence of handwashing with soap in the community, International Journal of Environmental Health Research, DOI: 10.1080/09603123.2018.1511772

To link to this article: https://doi.org/10.1080/09603123.2018.1511772

Published online: 27 Aug 2018.

INTERNATIONAL JOURNAL OF ENVIRONMENTAL HEALTH RESEARCH
https://doi.org/10.1080/09603123.2018.1511772

Taylor & Francis
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

Comparison of structured observation and pictorial 24 h recall of household activities to measure the prevalence of handwashing with soap in the community

WP Schmidt, HE Lewis, K Greenland and V Curtis

Department of Disease Control, London School of Hygiene and Tropical Medicine, London, UK

ABSTRACT

This study compared structured observation with a 24 h pictorial recall of household activities ('sticker diary') to measure the prevalence of handwashing with soap (HWWS) in the community. The study was done within a cluster-randomised trial evaluating a handwashing promotion programme in Bihar, India. HWWS at key occasions in mothers and school children was measured by structured observation in 299 households from 32 villages. Sticker diaries recalling common activities, including personal hygiene, were used to measure HWWS in 299 households from a further 20 villages. Sticker diary HWWS prevalence estimates were about 13% points higher than structured observation estimates, but the differences varied by the type of handwashing occasion. This study confirms structured observation as the method of choice for the study of handwashing behaviours. The sticker diary method may be useful in large-scale surveys. Sticker diaries may overestimate HWWS at important occasions, but probably less so than conventional questionnaire tools.

ARTICLE HISTORY
Received 19 June 2018
Accepted 10 August 2018

KEYWORDS
Handwashing; observation; recall

Background

The measurement of common health-related behaviours such as compliance with regular medications (Zhang et al. 2014), exclusive breastfeeding (WHO 2008), condom use (Doyle et al. 2010), preparation of oral rehydration solutions, latrine use (Sinha et al. 2016) and hand hygiene (in medical (Santosaningsih et al. 2017) or community settings (Cousens et al. 1996; Curtis et al. 2001)) remains a challenge for the evaluation of public health interventions. Methods of measurement include self-report, the use of proxy markers for the target behaviour or direct observation of the target behaviour itself.

Direct observation may be carried out by an observer who is present at the site or by video recording. Uniquely among the three approaches, direct observation attempts to measure the target behaviour as it occurs, potentially providing data that closely reflects the outcome of interest. However, direct observation has a number of drawbacks. Firstly, the observation of intimate behaviours such as breastfeeding, condom use or defecation may be inappropriate. Secondly, the target behaviour needs to occur frequently enough for a sufficient number of events to be observed during field worker visits. Thirdly, observation may cause study subjects to change their behaviour because they know that they are being observed, which may lead to an overestimate of socially desirable behaviours (reactivity). This is a particular problem in trials when reactivity may be higher in the intervention than in the control arm. Reactivity is least when subjects are unaware that they are being observed. If this is not feasible, then the observed need to

CONTACT WP Schmidt Wolf-Peter.Schmidt@lshtm.ac.uk Department of Disease Control, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT London, UK

be kept unaware of the exact purpose of the observation, in an attempt to keep reactivity non-differential across arms. In particular, subjects in the intervention arm should be kept unaware that the observation is linked to the intervention they have previously been exposed to, for example by providing them with a cover story (Biran et al. 2014). Finally, the use of direct observation for studies in community settings is often limited by logistical issues, requiring a large number of well-trained staff or video-equipment.

Proxy-markers of behaviour can take a variety of forms. Clinical studies often use the number of pills left in a bottle as a proxy of actual consumption (Zhang et al. 2014). Studies on handwashing and latrine use have applied electronic sensors to monitor the use of soap (Ram et al. 2010), soap dispensers (Judah et al. 2009) and the movement of persons inside a latrine (Sinha et al. 2016). These proxy markers may be subject to reactivity unless their use remains hidden (for example by wiring up soap dispensers in a public toilet (Judah et al. 2009)). A proxy may be the presence of a piece of equipment without which the target behaviour cannot be carried out. For example, the presence of soap and water is commonly used as an indicator of handwashing with soap (Biran et al. 2008). Ultimately, the validity of a proxy depends on how closely related it is to the target behaviour. The number of pills left in a pill dispenser may (or may not (Czobor and Skolnick 2011)) closely reflect the taking of pills, while the presence of soap and water is only a prerequisite to handwashing with soap; handwashing with soap is unlikely to take place without it, but may still not always take place when facilities are in place.

The most common, but potentially the least reliable, way of measuring health behaviour relies on self-reported or carer-reported data. Questionnaire surveys are straightforward, cheap, and require only limited expertise. However, respondents are likely to exaggerate behaviours that are socially desirable or that have been promoted in a recent intervention (Curtis, Cousens et al. 1993; Manun'Ebo et al. 1997; Biran et al. 2008). The difference between true and reported behaviour is often large enough to prevent any meaningful interpretation. For example, the reported prevalence of handwashing with soap after defecation commonly reaches 90% while observed behaviour in comparable settings rarely exceeds half of this figure (Biran et al. 2008; Scott et al. 2008; Rajaraman et al. 2014). While there may be some correlation between reported and actual behaviour, it is difficult to tell whether an intervention affected actual behaviour, or was limited to influencing the way in which participants responded to questions. Nevertheless, reported behaviour remains the most commonly used method to assess health behaviours. Assessing exclusive breastfeeding (one of the most important behaviours to prevent child death in low-income settings) relies almost exclusively on self-report (WHO 2008). In many questionnaire surveys, efforts are made to reduce the tendency of respondents to over-report desirable behaviours, by making the questions less leading, by avoiding prompting and by reducing the risk of recall error. For example, surveys on breastfeeding specifically ask for behaviours in the last 24 h (WHO 2008).

In an attempt to further reduce the over-reporting of handwashing with soap, the Lifebuoy brand of Unilever Ltd and the consumer research company Nielsen have developed a survey methodology where participants are given a set of pictorial representations of common daily activities and are asked to arrange them in the sequence as they have performed them over a set period of time. Handwashing is just one of many options the respondent can choose from. Over-reporting of handwashing is expected to be minimised, as the respondent is kept unaware of the purposes of the survey. In this study, we compared this new approach with direct observation.

Materials and methods

Study setting and population

The study was conducted in the context of a cluster-randomised trial (CRT) to measure the effect of Unilever's 'School of 5' campaign (So5) in the Indian state of Bihar. The campaign and CRT have been described elsewhere (Lewis, et al. 2018). Briefly, So5 is a school-based intervention to increase handwashing with soap (HWWS) in school children and their families at 5 'key occasions': after defecation, while

taking a bath (ie using soap whilst bathing), and before breakfast, lunch and dinner. The intervention consists of four school visits. A team of two facilitators engage the students in activities to generate awareness and commitment, followed by activities to reinforce messages and a reward ceremony. Parents are invited to join the activities on one day.

So5 has been rolled out across parts of India, Indonesia and several African countries. The campaign in Bihar was co-funded by the Children's Investment Fund Foundation (CIFF), a UK-based charity, which engaged the London School of Hygiene and Tropical Medicine (LSHTM) to carry out an impact evaluation. The campaign was planned to last for three years, covering 9 million school children. However, as the CRT demonstrated only a minimal increase in HWWS, the campaign was halted after one year.

The CRT to measure the effect of the campaign was carried out in three blocks (an administrative unit) across two neighbouring districts of Vaishali (Desari Block) and Samastipur (Bibhutipur Block and Rosera Block). Within the three blocks, all villages with schools eligible to receive So5 (public schools with at least 150 children) were randomised to intervention and control groups. All eligible schools in a village received the same allocated treatment. The intervention was the same in all three blocks, except for a modification in Rosera Block where a single session at childcare centres (Angawadi) during which mothers of young children in the village discussed handwashing was added.

The pre-specified primary outcome of the study was the proportion of So5 key occasions that were associated with handwashing with soap or soap use for bathing in children attending a school eligible to receive So5 and their mothers. The outcome measure was a composite of actual handwashing with soap at pre-defined occasions and soap use for bathing. For simplicity, we refer to all primary outcome behaviours as HWWS. Within study villages, households were eligible to be included for observation or diary if: 1) they had a child that regularly attended a school eligible to receive So5 in the last 3 months, and if: 2) that child had a sibling under 5 years. Households were recruited through house-to-house search, until the intended number of households in that village was enrolled, or no further eligible households could be found. Most villages contained less than 20 eligible households.

We used two different methods to measure HWWS, structured observation (henceforth 'observation') (Curtis, Cousens et al. 1993) and Unilever's 'sticker diary' method (henceforth 'diary'). Observation served as the primary outcome, the diary method as additional outcome. Importantly, to avoid interference between the two methods potentially causing reactivity, we decided to employ observations and the diary method in separate blocks. In two blocks (Desari and Bibhutipur), we used observation. In the third block, Rosera, we used diaries.

Structured observation

Structured Observations took place from 05:00 to 07:30 (Curtis, Cousens et al. 1993; Cousens et al. 1996; Biran et al. 2014). We employed female observers who had no connection with the intervention. The observers used coded sheets to record their observations and wrote a short description for each observation. Observations were done in 32 villages (16 in each of the two blocks). Within each block, a random sample of 16 (8/8) villages from the pool of villages randomised to intervention and control (27 in Desari, 41 in Bibhutipur) was selected. No baseline observation data were collected to decrease the risk of reactivity. Villages were visited about 8–10 weeks after the end of the intervention to identify eligible households, introduce the observers to the household and obtain consent. All observations in that village were conducted the following day. As a cover story, households were told that this was a study of domestic water use.

Sticker diary

Respondents were asked to fill in a diary sheet using a set of stickers showing photographs of different tasks selected from a comprehensive set that reflected activities that they may have undertaken that day (shown in Table 1). The respondent was asked to paste the relevant sticker in

the diary in the appropriate time slot in the order that the activities happened over the past 24 h with the help of the field staff. No specific hints as to the purpose of the exercise were given. The mother and the school child in a household were interviewed simultaneously by two enumerators to avoid one influencing the responses of the other.

Diaries were used in 20 villages in Rosera Block (10 intervention and 10 control villages) drawn at random from the 40 originally randomised villages in that block. In contrast to the observation villages, a baseline diary survey was done in the diary villages approximately 6–8 weeks before the intervention, as a baseline survey was part of Unilever's methodology. The follow-up diary survey was conducted 10–12 weeks after the intervention. Enrolment and application of the diaries in one village were all carried out on a single day.

Statistical methods

The primary outcome included HWWS before each meal and after defecation as well as soap use during bathing by children attending an eligible school and their mothers. As a further outcome we calculated the prevalence of handwashing with or without soap (HW) before each meal and after defecation. The comparison of handwashing prevalence estimates resulting from diaries and observation was restricted to the control arm and to morning time so that the diary data coincided with the observation data. Since following the programme modification in Rosera the intervention

Table 1. Sticker diary content.

School-children	Mothers
personal care	personal care
Urinate	urinate
Defecate	defecate
have bath with water only	have bath with water only
have bath with soap	have bath with soap
wash hand with water only	wash hand with water only
wash hand with water and soap	wash hand with water and soap
wash hand with ash, mud, soil	wash hand with ash, mud, soil
brushing teeth	brushing teeth
applying powder	applying cream
work	work
go to school	go to field
coming from school to home	clean kitchen/house
help father/mother in the field	milking/feeding/giving bath to cow or goats
help mother in house hold work	prepare tea
doing home work or going to tuition	prepare food
food/drink	washing dishes
eat	wash clothes
drink water	serving food
drink milk/tea, coffee	child care
leisure	give bath to child with water only
play	give bath to child with soap
listen to radio	take child for defecation
watch tv	feeding child
sleeping or resting	baby care (not feeding or bathing)
pray	food/drink
other	eat
	drink water
	drink milk/tea/coffee
	leisure
	listen to radio
	watch tv
	sleeping or resting
	going out (doctor, market, relative's house, etc)
	pray
	other

effect among mothers may not be comparable between Rosera and the other two blocks, we excluded mothers from the analysis comparing the effect estimates between diary and observation.

Prevalence of HWWS was calculated using STATA's *svy: proportion* command, adjusted for clustering at village-level using linearised standard errors. Prevalence differences between diary and observation, and intervention and control were calculated using binomial regression analysis (binomial distribution, identity link). Clustering at the village-level was accounted for by using Generalised Estimating Equations (GEE) and robust standard errors. P-values for the comparison of socio-demographic characteristics across the three blocks were calculated using linear regression (continuous variables), logistic regression (binary variables) and ordinal logistic regression (ordered categorical variables) within STATA's *svy* commands. All analyses were done in STATA 12.0.

The study was approved by LSHTM's ethics committee and Hindustan Unilever's Independent Ethics Committee. The trial was registered at Clinicaltrials.gov (NCT02424812).

The 'School of 5' campaign in Bihar was funded by Unilever plc and the Children's Investment Fund Foundation (CIFF), a UK-based charity. This evaluation was funded by CIFF. The funders of this campaign, and the study had no role in the data collection, data analysis, data interpretation, writing of the report or decision to submit for publication.

Results

For the observation study, we enrolled 338 households from 32 villages (16 per block and per study arm). We recruited 10.6 households per village (SD 1.2, range 9–12). For the diary study, we originally enrolled 341 households from 20 villages (one block only, 10 per study arm). We restricted the analysis to 299 households for which complete socio-demographic data were available, resulting in 15.0 households per village (SD 2.9, range 7–19).

There were some socio-demographic differences among the three blocks (Table 2). Rosera Block had a higher percentage of Muslim households, lower levels of paternal and maternal education, less access to electricity, and less solid housing (pukka) compared to the observation blocks Desari and Bibhutipur. Access to water was similar across blocks. The prevalence of open defecation was around 90% in Rosera and Bibhutipur but less common in Desari.

Overall, HWWS prevalence was 31% (Table 3) in the sticker diary baseline survey (Rosera), and somewhat higher in the follow-up survey (35%). The overall HWWS prevalence in the structured observation study ranged between 20% (Bibhutipur) and 25% (Desari).

Table 4 compares handwashing with or without soap (HW) and HWWS prevalence estimates from diary (follow-up only) and observation surveys by type of key occasion. For both children and mothers, diary and observation estimates were similarly low for HWWS before breakfast. Prevalence estimates from diaries were about twice as high as for observation estimates for HWWS after defecation. By contrast, diary estimates for soap use during bathing were lower than observation estimates. HW was almost universal after defecation for both person groups and type of method.

Figure 1 shows the differences in HWWS prevalence between diary (follow-up only) and observation, adjusted for maternal and paternal education (none vs other) and sanitation (open defecation vs other). Other variables unevenly distributed among blocks (caste, electricity, house type, water access) were dropped from the model as they did not affect the estimates. Diary HWWS prevalence estimates were about 13% points higher than observation estimates. HWWS proportions after defecation were 24% points higher for the diary method than for observation (children and mothers combined), while soap use during bathing was lower for diary than for observation.

Figure 2 compares how the two methods performed as effect measures (children only). Most effect estimates (overall, soap use during bath) were lower for the diary method than for observation. For HWWS after defecation, the estimated effect size was slightly higher for diary than for observation All effect sizes were well below the pre-specified effect size of interest for the sample size calculation (15% prevalence difference) and were characterised by wide confidence intervals.

Table 2. Socio-demographic characteristics of enrolled households (intervention and control).

	Sticker diary	Structured observation		
	Rosera (N = 299)	Bibhutipur (N = 173)	Desari (N = 165)	P value
Household size, mean (SD)	7.0 (2.3%)	7.0 (2.8)	7.2 (2.3)	0.770
Children under 5 years, mean (SD)	1.3 (0.6%)	1.5 (0.9)	1.4 (0.6)	0.09
Caste, n (%)				0.378
General	23 (7.7%)	14 (8.1)	12 (7.3)	
Other Backward caste	117 (39.1%)	77 (44.5)	76 (46.1)	
Scheduled caste	119 (39.8%)	70 (40.5)	71 (43.0)	
Scheduled Tribe	3 (1.0%)	9 (5.2)	3 (1.8)	
Muslim	37 (12.4%)	3 (1.7)	2 (1.2)	
Other	0 (0.0%)	0 (0%)	1 (0.6%)	
Father's education, n (%)				0.007
None	130 (43.5%)	63 (36.4%)	58 (35.2)	
Some primary	54 (18.1%)	20 (11.6%)	19 (11.5)	
Primary completed	30 (11.7%)	25 (14.5%)	13 (7.9)	
Some secondary	29 (9.7%)	13 (7.5%)	16 (9.7)	
Secondary completed	14 (4.7%)	20 (11.6%)	16 (9.7)	
Higher	37 (12.4%)	31 (17.9%)	43 (26.1)	
Unknown	0 (0.0%)	1 (0.6%)	0 (0.0)	
Mother's education, n (%)				0.06
None	223 (74.6%)	115 (66.5%)	101 (61.2)	
Some primary	25 (8.4%)	12 (6.9%)	14 (8.5)	
Primary completed	12 (4.0%)	16 (9.3%)	11 (6.7)	
Some secondary	16 (5.4%)	9 (5.2%)	10 (6.1)	
Secondary completed	11 (3.7%)	9 (5.2%)	10 (6.1)	
Higher	12 (4.0%)	11 (6.4%)	19 (11.5)	
Unknown	0 (0%)	1 (0.6%)	0 (0.0)	
Electricity, n (%)	141 (47.2)	94 (54.3%)	122 (73.3%)	< 0.001
Motorbike, n (%)	24 (8.0)	17 (9.8%)	13 (7.9%)	0.808
House type, n (%)				< 0.001
Pukka (concrete)	72 (24.1%)	55 (31.8%)	64 (38.8)	
Semi-pukka (half concrete)	86 (28.8%)	36 (20.8%)	59 (35.8)	
Kuccha (mud)	141 (47.2%)	82 (47.4%)	42 (25.5)	
Drinking water source, n (%)				0.776
Private tubewell	196 (65.6%)	104 (60.1%)	103 (62.4)	
Public tap	31 (7.0%)	22 (12.7%)	11 (6.7)	
Public tubewell	65 (21.7%)	43 (24.9%)	46 (27.9)	
Dug well	10 (3.3%)	2 (1.2%)	4 (2.4)	
Other	7 (2.3%)	2 (1.2%)	1 (0.6)	
Location of water source, n (%)				0.427
Inside house	73 (24.4%)	51 (29.5%)	54 (32.9)	
Inside compound	83 (27.8%)	43 (24.9%)	39 (23.8)	
Outside compound	143 (47.8%)	79 (45.7%)	71 (43.3)	
Sanitation, n (%)				< 0.001
Pour flush latrine	20 (6.7%)	16 (9.3%)	36 (21.7)	
Pit latrine with slab	5 (1.7%)	0 (0.0%)	4 (1.8)	
Pit latrine without slab	5 (1.7%)	0 (0.0%)	5 (3.0)	
None	269 (90.0%)	157 (90.8%)	122 (73.5)	

Discussion

Sticker diaries are a new survey tool that attempt to measure handwashing with soap – an important health behaviour – by using unprompted recall of past activities pasted as stickers into a diary sheet. We found marked differences in estimates of HWWS prevalence between direct observation and sticker diaries. Both methods produced results that suggested that the effect of the intervention on HWWS had been small. If direct observation is used as the reference, then it appears that the diary method may overestimate HWWS, especially after defecation.

The diary method has a number of advantages. In particular, the number of recorded events in a household is higher than for observation, as diary covers a 24 h period. Further, in contrast to observation, one enumerator can survey several households per day. The diary was also able to capture intimate behaviours that are difficult to observe by observation, such as bathing in mothers.

Table 3. Estimated prevalence of handwashing with soap/soap use for bathing (control arm) by block and method.

	Sticker diary			Structured observation					
	Rosera (10 villages, 157 HHs)			Bibhutipur (8 villages, 87 HHs)			Desari (8 villages, 82 HHs)		
	N*	Pre-valence	95% CI	N*	Pre-valence	95% CI	N*	Pre-valence	95% CI
Baseline survey (diary)									
HWWS									
Overall	854	31.0%	(21.2, 40.1)						
Children	466	31.1%	(21.4%, 40.8%)						
Mothers	388	30.1%	(20.5%, 41.3%)						
Post-intervention survey									
HWWS									
Overall	740	34.7%	(27.2, 42.2)	191	19.9	(13.0%, 26.8%)	218	24.8%	(18.0%, 31.5%)
Children	428	32.0%	(23.9%, 40.2%)	124	20.2%	(10.2%, 30.1%)	138	18.8%	(12.7%, 25.0%)
Mothers	312	38.5%	(31.0%, 46.0%)	67	19.4%	(8.4%, 30.4%)	80	35.0%	(19.1%, 50.9%)

*Number of observed key occasions (after defecation, before each meal, soap use during bathing); NB: no baseline was conducted for structured observation.

Table 4. Estimated prevalence of handwashing with soap/soap use for bathing (control arm, restricted to follow-up survey) by type of occasion.

	Sticker diary (10 villages)			Structured observation (16 villages)		
	N*	Prevalence	95% CI	N*	Prevalence	95% CI
Children						
Handwashing ± soap						
before breakfast	157	9.5%	(3.9%, 15.4%)	156	11.5%	(5.9%, 17.1%)
after defecation	146	96.6%	(93.9%, 99.2%)	41	100.0%	- -
Handwashing with soap						
before breakfast	157	0.0%	- -	156	0.0%	- -
after defecation	146	50.7%	(34.9%, 66.4%)	41	24.4%	(11.7%, 37.1%)
Soap use during bath	125	50.4%	(40.2%, 60.6%)	65	63.1%	(48.8%, 77.4%)
Mothers						
Handwashing +/- soap						
before breakfast	88	3.4%	(0.0%, 7.1%)	7	14.3%	(0.0%, 52.4%)
after faecal contact	137	98.5%	(96.4%, 100.0%)	132	98.5%	(96.3%, 100.0%)
Handwashing/bathing with soap						
before breakfast	88	1.1%	(0.0%, 3.6%)	7	0.0%	- -
after faecal contact	137	47.4%	(31.3%, 63.6%)	132	25.8%	(15.8%, 35.7%)
Soap use during bath	84	64.3%	(56.3%, 72.3%)	8	87.5%	(56.0%, 100.0%)

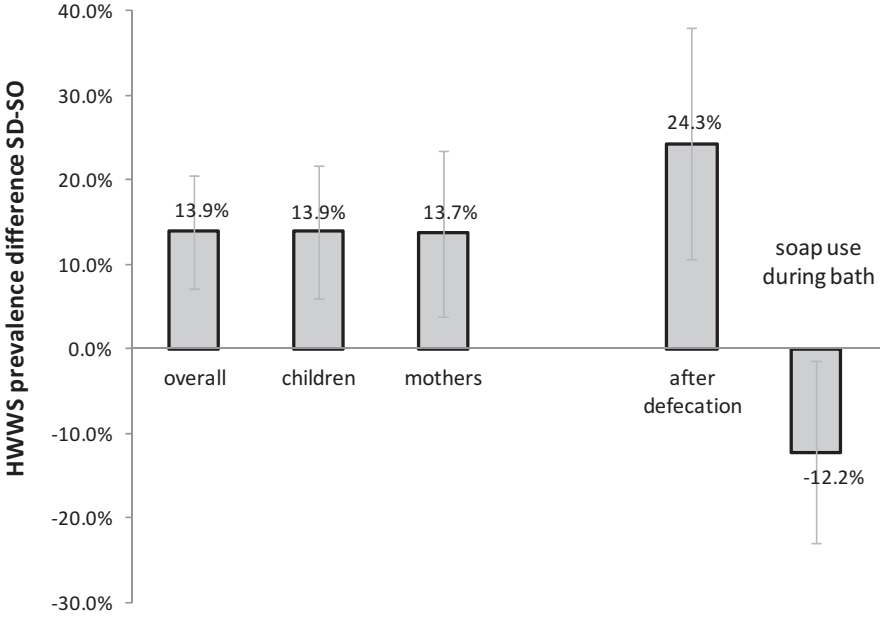

Figure 1. Difference in handwashing with soap (HWWS) prevalence between sticker diary and structured observation. Differences and 95% CI calculated using binomial regression with identity link/GEE, adjusted for mothers' and fathers' education, and sanitation access.

The analysis of diary data is made difficult by the uncertainty with regards to the exact time at which activities occur. If for example, a respondent places the sticker for HWWS or bathing with soap ahead of eating, then it is unclear how much time has actually passed and whether potentially contaminating behaviours that are not captured by the diary occurred in between. Diaries appear to overestimate HWWS especially post-defection, presumably because this behaviour already seems to be the social norm with a strong potential for over-reporting. However, over-reporting of post-defecation HWWS appears lower with sticker diaries than for self-reported HWWS using conventional questionnaire techniques, which often produce estimates of 90% or higher (Manun'Ebo et al. 1997; Biran et al. 2008; Scott et al. 2008).

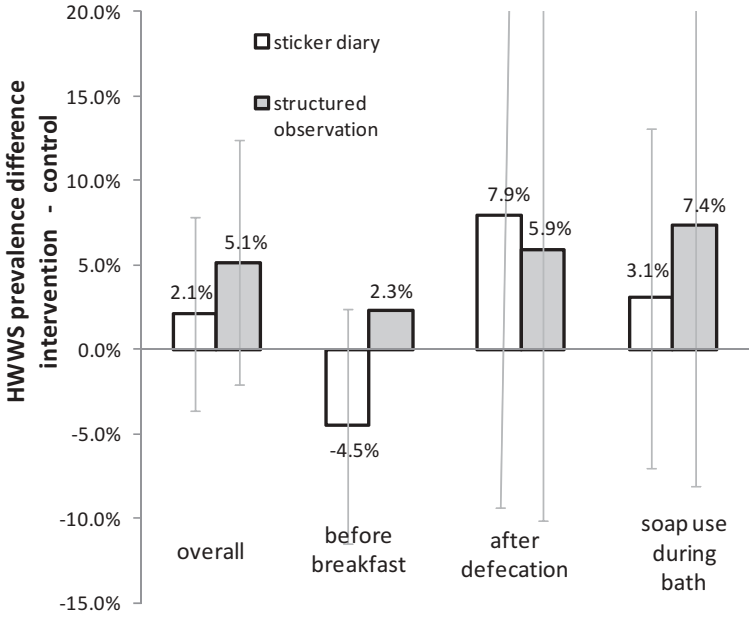

Figure 2. Effect of the intervention on handwashing with soap (HWWS) at key occasions in school-age children as measured by observation and sticker diary. Differences and 95% CI calculated using binomial regression with identity link/GEE. (95% CI could not be calculated for before breakfast as the model did not converge.)

In this analysis, we used structured observation as the reference to compare the performance of sticker diaries. Observation has its own disadvantages and may not qualify as a 'gold standard' in the strict sense. Observation carries a risk of reactivity, intimate activities are difficult to observe and HWWS before eating is hard to define as people may eat intermittently. Also, breakfast may be preceded by defecation and HWWS. It is often up to the observer to define whether HWWS after defecation can be counted as HWWS before eating in this case, depending on how much time has passed. However, the problem of timing of activities is not as pronounced as for diary as the observer is able to record exact times. Observation also allows the simultaneous observation of more household members, including men, grandmothers (who may take an active role in child care), and other children, as all can often be observed at the same time.

It is not clear from our study whether a true change in HWWS prevalence following an intervention would affect diary and observation estimates differently, or whether, say an increase in 10% points measured by diary is equivalent to a 10% point increase in HWWS measured by observation. Only a study with a larger increase in HWWS than found in our trial (Figure 2) would allow us to explore this issue further.

The most important limitation of this study, apart from using observation as an imperfect 'gold standard', lies in the fact that the study was not originally designed to compare the two methods. The primary outcome of the CRT was HWWS observed by observation. We chose not to use the two methods in the same households, villages, or even the same blocks, so as to minimise the risk of reactivity. Blocks differed somewhat in their socio-economic characteristics. However, it seems unlikely that true handwashing prevalences would have differed much among blocks as shown by the similar observation estimates for HWWS in Desari and Bhibutipur (Table 3). Adjusting for socio-economic differences produced somewhat larger differences between diary and observation (Figure 1) but did not materially alter the overall pattern. Unlike observation, which was only carried out after the intervention to avoid respondents linking the research to the intervention, diary surveys were done at baseline and follow-up. HWWS estimates from diaries were about 4% points higher at follow-up than at baseline which is a small difference, but which may be due to

repeated questioning leading to more over-reporting. As for observation, we recommend restricting diary surveys to the post-intervention period, leaving perhaps some 8–10 weeks between intervention and survey, which may contribute to minimise over-reporting and reactivity as the link between intervention and survey may be less clear to participants.

In conclusion, this study corroborates observation, if not as the gold standard, then as the method of choice for the study of handwashing behaviours. While the use of observation may be preferred to evaluate randomised controlled trials, the diary method may have a role in large-scale surveys where observation may be infeasible. Sticker diaries may overestimate HWWS at important occasions, but less so than conventional questionnaire tools. They require considerably less staff training and expertise and a smaller sample size of households to reach the same precision as observation surveys. Sticker diaries may also be a useful tool for measuring other sensitive health behaviours such as open defecation, or breast feeding. Both diary and observation methods need to be carefully adapted to the social setting of the target population and require thorough piloting. By modifying the stickers offered to respondents and by combining the diary data with other questionnaire variables such as the presence of a handwashing place in the house, it may be possible to further improve the validity of the estimates.

Acknowledgments

We thank the study participants and field staff for their valuable contribution.

Disclosure of interest

Val Curtis, Katie Greenland and Wolf-Peter Schmidt have received research funding including salary costs from Unilever Ltd. This can be seen as a conflict of interest as this study evaluated a Unilever campaign. Unilever had no role in study design, study execution, data analysis, the writing of the manuscript or the decision to publish.

Funding

This work was supported by the Children's Investment fund foundation [NA];

References

Biran A, Rabie T, Schmidt W, Juvekar S, Hirve S, Curtis V. 2008. Comparing the performance of indicators of hand-washing practices in rural Indian households. Trop Med Int Health. 13(2):278–285.

Biran A, Schmidt WP, Varadharajan KS, Rajaraman D, Kumar R, Greenland K, Gopalan B, Aunger R, Curtis V. 2014. Effect of a behaviour-change intervention on handwashing with soap in India (superAmma): a cluster-randomised trial. Lancet Glob Health. 2(3):e145–154.

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):1299–1308.

Curtis V, Cousens S, Mertens T, Traore E, Kanki B, Diallo I. 1993. Structured observations of hygiene behaviours in Burkina Faso: validity, variability, and utility. Bull World Health Organ. 71(1):23–32.

Curtis V, Kanki B, Cousens S, Diallo I, Kpozehouen A, Sangare M, Nikiema M. 2001. Evidence of behaviour change following a hygiene promotion programme in Burkina Faso. Bull World Health Organ. 79(6):518–527.

Czobor P, Skolnick P. 2011. The secrets of a successful clinical trial: compliance, compliance, and compliance. Mol Interv. 11(2):107–110.

Doyle AM, Ross DA, Maganja K, Baisley K, Masesa C, Andreasen A, Plummer ML, Obasi AI, Weiss HA, Kapiga S, et al. 2010. Long-term biological and behavioural impact of an adolescent sexual health intervention in Tanzania: follow-up survey of the community-based MEMA kwa Vijana Trial. PLoS Med. 7(6):e1000287.

Judah G, Aunger R, Schmidt WP, Michie S, Granger S, Curtis V. 2009. Experimental pretesting of hand-washing interventions in a natural setting. Am J Public Health. 99(Suppl 2):S405–411.

Lewis HE, Greenland K, Curtis V, and Schmidt WP. 2018. Effect of a School-Based Hygiene Behavior Change Campaign on Handwashing with Soap in Bihar, India: Cluster-Randomized Trial. Am J Hyg Trop Med, https://doi.org/10.4269/ajtmh.18-0187.

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):1015–1021.
Rajaraman D, Varadharajan KS, Greenland K, Curtis V, Kumar R, Schmidt WP, Aunger R, Biran A. 2014. Implementing effective hygiene promotion: lessons from the process evaluation of an intervention to promote handwashing with soap in rural India. BMC Public Health. 14:1179.
Ram PK, Halder AK, Granger SP, Jones T, Hall P, Hitchcock D, Wright R, Nygren B, Islam MS, Molyneaux JW, et al. 2010. Is structured observation a valid technique to measure handwashing behavior? Use of acceleration sensors embedded in soap to assess reactivity to structured observation. Am J Trop Med Hyg. 83(5):1070–1076.
Santosaningsih D, Erikawati D, Santoso S, Noorhamdani N, Ratridewi I, Candradikusuma D, Chozin IN, Huwae TE, Van Der Donk G, Van Boven E, et al. 2017. Intervening with healthcare workers' hand hygiene compliance, knowledge, and perception in a limited-resource hospital in Indonesia: a randomized controlled trial study. Antimicrob Resist Infect Control. 6:23.
Scott BE, Schmidt WP, Aunger R, Garbrah-Aidoo N, Animashaun R. 2008. Marketing hygiene behaviours: the impact of different communication channels on reported handwashing behaviour of women in Ghana. Health Educ Res. 23(3):392–401.
Sinha A, Nagel CL, Thomas E, Schmidt WP, Torondel B, Boisson S, Clasen TF. 2016. Assessing latrine use in rural India: a cross-sectional study comparing reported use and passive latrine use monitors. Am J Trop Med Hyg. 95 (3):720–727.
WHO. 2008. Indicators for assessing infant and young child feeding practices. part i: definitions. Geneva: WHO.
Zhang Z, Peluso MJ, Gross CP, Viscoli CM, Kernan WN. 2014. Adherence reporting in randomized controlled trials. Clin Trials. 11(2):195–204.