Published by Oxford University Press on behalf of the International Epidemiological Association *International Journal of Epidemiology* 2006;**35:**1469–1477

doi:10.1093/ije/dyl165

A simple index to measure hygiene behaviours

Amy L Webb,[1] Aryeh D Stein,[1,2] Usha Ramakrishnan,[1,2] Vicki S Hertzberg,[3] Miriam Urizar[2] and Reynaldo Martorell[1,2]*

Accepted 5 July 2006

Background Spot checks are becoming a popular method to assess hygiene behaviours; however, little is known about their repeatability or predictability. We evaluated the within-household repeatability of hygiene indices created from spot checks and their ability to predict incidence of diarrhoea in young Guatemalan children.

Methods We observed hygiene behaviours in 588 households in four rural Guatemalan communities over 36 months. Four indices related to drinking water (DWI; score = 0–3), food (FI; score = 0–3), personal hygiene (PHI; score = 0–3), and domestic household hygiene (DHI; score = 0–6) and one summary hygiene index (SHI; range 0–15) were created. Morbidity of 694 children aged birth to 36 months living in the study households was assessed using biweekly recall. Intraclass correlation coefficients were calculated to assess within-household repeatability; the generalized estimating equations approach was employed to analyse diarrhoea morbidity.

Results Households were observed a mean of 22.1 ± 11.2 times. All indices decreased with duration of follow-up (SHI = −0.67 ± 0.05 points/year; WI = −0.04 ± 0.01; FI = −0.07 ± 0.01; PHI = −0.21 ± 0.01; DHI = −0.37 ± 0.02; all $P < 0.05$). Intraclass correlations were low to moderate (SHI = 0.35–0.51; DWI = 0.17–0.21; FI = 0.16–0.18; PHI = 0.27–0.32; DHI = 0.27–0.38). Six separate spot checks would be needed to estimate a household's underlying level of hygiene within 20%. SHI and PHI scores were inversely associated with diarrhoea morbidity (both $P < 0.05$).

Conclusions Hygiene indices created using spot checks can be a rapid and efficient method for assessing hygiene and useful for predicting diarrhoea morbidity in young children. Multiple measures are required to accurately estimate the true hygiene pattern of a household.

Keywords Hygiene, rapid spot checks, diarrhoea, intraclass correlation coefficient, kappa, repeatability

Diarrhoea contributes to an annual 3.5 million deaths in children under the age of three, and diarrhoea-related illness and complications are associated with malnutrition, growth faltering, and compromised immunity.[1] The burden of diarrhoea morbidity lies largely in the developing world where water and living conditions remain poor. Interventions to improve hygiene behaviours, sanitation, and water quantity and quality can reduce all-cause morbidity by >20%,[2] and hygiene interventions specifically were reported to reduce diarrhoea morbidity by nearly 45%.[3] The World Bank has declared the promotion of improved hygiene to be the third most productive method for preventing diarrhoea-related morbidities.[4]

A variety of methods are currently utilized to capture sanitation and hygiene behaviours. Interview and questionnaire approaches, while easily modified to represent the community under study, may not be as easily standardized as other available methods and typically result in over-reporting of 'good' behaviours, thus, reducing their validity.[5–8] Structured observations are used to gather information on the occurrence of specific behaviours during an observational period at the

[1] Nutrition and Health Sciences Program, Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, GA 30322, USA.

[2] Hubert Department of Global Health, The Rollins School of Public Health of Emory University, Atlanta, GA 30322, USA.

[3] Department of Biostatistics, The Rollins School of Public Health of Emory University, Atlanta, GA 30322, USA.

* Corresponding author: Hubert Department of Global Health, The Rollins School of Public Health of Emory University, 1518 Clifton Road, N.E. Atlanta, Georgia 30322, USA. E-mail: rmart77@sph.emory.edu

home. While adequate for providing detailed information, structured observations tend to be time and labour intensive, difficult to standardize, and may result in behaviour changes due to the presence of the observer (reactivity).[9] Additionally, structured observations tend to be highly variable from visit to visit, thus reducing repeatability within subjects.[5,10] Spot checks are rapid and easily standardized alternatives to structured observations. Generally observers assess hygiene behaviours using a line-item checklist of behaviour proxies. For example, rather than observing a mother wash her hands with soap, a spot-check method will require that the cleanliness of the mother's hands and nails be observed and that observation will serve as an indicator of hand-washing behaviour. It is unnecessary for an observer to remain in the home until specific hygiene activities are observed, thus reducing burden on both the field worker and the subject. Ruel and Arimond recently concluded that spot-check methods provide a 'time-saving and economic alternative' to structured observations once the appropriate training and standardization have been completed.[9] Several studies have used observations from spot checks to create indices representing hygiene behaviours that cluster and may, therefore, better represent behaviours or potential routes of contamination.[11–13] Despite their apparent benefits related to ease of use in the field, little is known regarding the repeatability or predictability of spot-check methods. To date, only a single study evaluating the reactivity and repeatability of hygiene measures obtained from spot checks has been published; in that study reactivity appeared to be less than that observed with structured observations, while within-subject repeatability ranged from poor to good depending on the specific hygiene measure observed.[14]

We used data from a longitudinal study of maternal care giving behaviours, including hygiene behaviours collected by spot-check observations, to create four specific hygiene indices and, from these indices, a summary hygiene index (SHI). The objectives of this study were to assess the stability of these indices over time, the within-household repeatability of specific and summary hygiene indices and estimate their predictability for early childhood diarrhoea. We also estimated the number of days necessary to yield estimates of habitual hygiene likely to be within 20 and 50% of the household true mean.

Methods

Study population and data collection

Subjects for the current study were recruited from four villages of mixed Spanish-Mayan descent, located 40–110 km east of Guatemala City, Guatemala. The four villages have been part of a series of ongoing longitudinal community-based studies over the past 30 years. Detailed descriptions of the villages and their selection are published elsewhere.[15] From March 1996 to September 1999, a study of pregnancy and child growth was conducted in these villages by Emory University and the Institute of Nutrition of Central America and Panama (INCAP). Pregnant women and women with children younger than 36 months were eligible to participate, regardless of participation in earlier studies. Pregnancies were ascertained and followed through to delivery. Children under 36 months participated until they reached 36 months. Multiple children were eligible within households.

Serial measurements of hygiene and sanitation behaviours were collected for 613 households. We excluded from analysis data from 25 households because observations were available for only one visit. The 25 excluded households had significantly fewer family members residing in the household than did those included (4.4 vs 3.6; $P = 0.01$) but did not differ with respect to maternal education, number of children, maternal age, socioeconomic status, or the summary hygiene score. Households were visited monthly by trained field workers (mean of 22 visits per household, range 2–45 visits) and a checklist of individual and household hygiene behaviours was completed at each visit by rapid observation.

Guatemala has two distinct seasons, a wet season that runs from May to October and a dry season for the remaining months of the year. Twice a year, once each during the rainy and dry seasons, additional sanitation information was collected by questionnaire from the primary caregiver on sources, storage, and treatment of household water, the nature and condition of the household latrine, and the location for defaecation of children under 5 years. Data on household size, education, and socioeconomic factors were obtained by interview. Characteristics of home, possessions, and father's occupational status were used to create a composite score for socioeconomic status (SES) with lower scores representing lower SES.[16] All questionnaires and checklists were field checked by a supervisor and data were double entered. The study was approved by the institutional review boards of Emory University and INCAP.

Hygiene behaviour indices

The hygiene checklist, as administered, included 18 items. We excluded three items in the index creation: whether the child was wearing shoes, the types of animals present in/around the home, and the general perception of the field worker regarding the cleanliness of the individuals and the household. The first was excluded because of lack of variability, while the second and third were discarded because of their lack of contribution to a specific oral–faecal route for contamination. Four indices (Table 1) were created from the remaining 15 items; each represented a different pathway by which oral–faecal contamination and, subsequently, diarrhoea may occur: contamination of stored water (drinking water index, DWI) food (food index, FI), or hands during defaecation (personal hygiene index, PHI); and presence of animals or insects that can transmit faecal contamination to previously clean surfaces (domestic household hygiene index, DHI). Each item was scored as 0 or 1, with 1 representing a positive behaviour. The indices were calculated as the simple sum of the items. A SHI was calculated as the sum of the four individual indices.

Diarrhoea morbidity

Morbidity assessment was performed every 15 days using a 19 day recall method to ensure overlap of days.[17,18] Trained field workers visited the homes every 2 weeks and recorded information on precoded forms regarding signs, symptoms, and duration of illnesses. The beginning and ending dates of a symptom were always recorded. The morbidity recall used in this study has been previously validated in this population by a physician and demonstrates adequate sensitivity and

Table 1 Indicators of hygiene observed by monthly rapid spot-check observations that were used to develop indices of specific and summary hygiene behaviours

Drinking water (DWI)	Food (FI)	Personal (PHI)	Household (HHI)
Interior water storage container is covered	Clean dishes are covered	Mother is wearing shoes	No trash outside house
Exterior water storage container is clean	Clean dishes are kept high	Mother's hands are clean	No trash inside house
Pila[a] contains water	All food is covered	Index child's hands are clean	No unrestrained animal in patio or house
			No accumulation of dirty clothes
			Insignificant number of flies in house
			No standing water in patio or around house
Possible score 0–3	Possible score 0–3	Possible score 0–3	Possible score 0–6

A summary hygiene index was generated from the sum of the four specific indices and had a total of 15 points possible.
[a] A pila is a water storage container.

specificity for diarrhoea (66 and 99%, respectively).[17] Diarrhoea was defined as 3 or more loose/watery stools in a 24 h period preceded by 24 h of diarrhoea free time. Data were later summarized as incident episodes and duration of episodes.

All instruments were developed by the study investigators and field tested by project supervisors prior to training of field workers. Field workers who were experienced with ethnographic interviews were recruited and trained for 5 weeks prior to the start of the study. Training involved a series of field-based data collection and feedback sessions, including a series for standardizations, in which field workers were paired first with other field workers and then with supervisors. Field workers were approved by the project supervisors and study director in Guatemala before beginning data collection directly in the field. During data collection, field workers were assigned to a different village every month in a randomized manner to minimize bias. During the course of the study, all completed forms were checked for completeness and random in-field observations of data collection were performed daily by field supervisors. A subsample of households was also randomly selected for duplicate measures by another field worker. Retraining of the entire team occurred periodically throughout the study as needed, at least once per year.

Analytic methods

We computed summary statistics for the individual behaviours and each of the five indices. Pairwise associations among the four specific indices were estimated using Spearman's *rho*.

Stability of the indices over time was calculated by modelling the effect of time on each hygiene index using a generalized estimating equations approach with an exchangeable working matrix (SAS 9.0; Proc MIXED) to control for correlation introduced by repeated measures.[19,20] Repeatability of each individual indicator used in index creation was assessed using the kappa statistic. Owing to the presence of multiple measures, we compared kappas calculated from only the first two observations with those calculated using multiple observations. To create a balanced dataset of multiple measures that would theoretically span both seasons, we randomly selected five observations from each household that was followed for at least 1 year ($n = 462$).

To assess the repeatability of the created indices we utilized intraclass correlation coefficients (ICCs) as described by Shrout and Fleiss.[21] The 23 field workers, or observers, used in this study to collect hygiene information did not consistently visit specific households: some observers visited multiple households multiple times; while others visited multiple households only once; and still others may have visited one household many times. As such, of interest to this study was the inclusion of field worker in the ICC calculations. To address the issue of field worker, we calculated three ICCs. ICC1 was calculated with only consideration for household as a random effect using the formula: $\left[\sigma^2_{HH}/\left(\sigma^2_{HH}+\sigma^2_{RES}\right)\right]$ where the numerator represents the within-household variance and the denominator represents the sum of the household variance and the residual variance. A second and third ICC were calculated including both household and field worker random effects. ICC2 was calculated as $\left[\sigma^2_{HH}/\left(\sigma^2_{HH}+\sigma^2_{RES}\right)\right]$ where the numerator represents the within-household variance and the denominator represents the total variance less the within-field-worker variance; it should be noted that the inclusion of field worker level random effects changes the residual variance such that the formulas for ICC1 and ICC2, though similar, will yield different values. ICC2 represents the ICC given a specific field worker conducts all spot checks. ICC3 was calculated as $\left[\sigma^2_{HH}/\left(\sigma^2_{HH}+\sigma^2_{RES}+\sigma^2_{FW}\right)\right]$ and it can be seen that the denominator now includes the within-field-worker variance estimate $\left(\sigma^2_{FW}\right)$. ICC3 represents an ICC given that any one of a larger set of raters might conduct a given spot check (SAS 9.1; Proc Mixed).

Using the within-subject variance measures and the mean score values for each index we calculated coefficients of variance and estimated the number of days of observations necessary to obtain estimates of household hygiene likely (with alpha = 0.95) to be within 20 and 50% of a household's true scores for the five hygiene indices, using the approach of Beaton *et al.*[22]

Each index was further tested to assess its ability to predict diarrhoea in children less than 36 months in those households with at least one hygiene observation ($n = 613$). To facilitate interpretation, the SHI was categorized based on approximate tertiles: low ($\leqslant 5$), middle (6–9), and high (>9). The DHI was dichotomized at the median for ease of interpretation (median = 2; IQR = 1, 3). A child's odds for having an incident episode of diarrhoea during the month following a spot check were assessed using logistic regression. Additionally the number of episodes and number of days ill per month

Table 2 Descriptive characteristics of the 588 households with at least two hygiene observations

Maternal variables	*N*	Mean ± SD	25th%	Median	75th%
Number of home visits	588	22.1 ± 11.2	12	23	32
Maternal years of schooling (yr)	585	3.8 ± 2.7	2.0	4.0	6.0
Maternal age at start of study (yr)	585	25.7 ± 4.3	22.5	25.4	28.9
Number of children living in the household at the start of study	550	3.0 ± 2.2	1.0	2.0	4.0
Number of persons in the family at start of study	550	4.5 ± 1.8	3.0	4.0	5.0
Total number of persons living in the household at start of study	550	5.7 ± 2.4	4.0	5.0	7.0
SES factor score	535	−0.1 ± 0.9	−0.8	−0.1	0.5

were evaluated using Poisson regression applying a negative binomial distribution. We evaluated three models of predictability of hygiene scores for diarrhoea morbidity. We first examined crude associations between monthly hygiene indices and incident diarrhoea. We added terms for age and sex of the child, household SES, community and season as potential confounders. Inclusion of community and season did not influence parameter estimates or significance testing (data not shown); and consequently our results are presented for models that exclude these terms. The population mean for household SES was imputed when household SES was missing (n = 53 households). Logistic and Poisson regression analyses were performed using the GEE approach while applying an autoregressive type I working matrix (SAS 9.0; Proc GENMOD) to control for correlation introduced by repeated measures.[19,20]

Results

A total of 12 989 assessments were made on 588 households with at least two hygiene observations, representing 93.8% of scheduled visits. Selected characteristics of these households are presented in Table 2. Mean family size was 4.5 persons and most families had additional persons residing with them. About 83% of households had electricity, 65% had a radio, 56% had a television, and 10% had a refrigerator. Latrines were present in 74% of the homes. Mothers of the household were predominantly in their mid-twenties with between 1 and 3 children at the start of the study and a mean of 4 years of formal schooling.

The distributions of the 15 hygiene behaviours observed and the four corresponding specific indices are presented in Table 3 and Figure 1. The SHI was consistent with a normal distribution with a mean of 7.0 ± 2.9 points. About 35% of households had SHI scores ⩽ 5; 43% had scores between 6 and 9; and 22% had scores >9. The four specific indices were positively correlated with each other with Spearman rho ranging from 0.28 (FI and PHI) to 0.46 (DHI and FI). As expected, each specific index was strongly and positively correlated with the summary index ($r > 0.68$, $P < 0.05$). Diarrhoea morbidity has been observed to peak during the wet seasons in several developing countries.[23,24] Deteriorating hygiene behaviours may be one of several reasons for the observed increase.[25] While we did observe a trend for decreased hygiene index scores in the wet season (net difference −0.06; 95% CI −0.14, 0.02, P = 0.12), this difference was not statistically significant.

Small but significant decreases in scores were observed with each additional year of follow-up (SHI= −0.67 ± 0.05 per yr; DWI= −0.04 ± 0.01; FI = −0.07 ± 0.01; PHI = −0.21 ± 0.01; DHI = −0.37 ± 0.02; all $P < 0.05$).

Kappa coefficients for individual spot-check items calculated using the first and second observations were consistently higher than those calculated from the five randomly selected observations (Table 3). Kappa coefficients calculated from the first and second observations were all <0.4. Kappa coefficients for individual spot-check items calculated using five randomly selected observations were generally below 0.15. Only 4 indicators achieved kappa coefficients ⩾0.15: clean external water container (k = 0.27), mother wearing shoes (k = 0.15), no trash inside home (0.19), and no unrestrained animals in patio or home (k = 0.26).

Table 4 presents information on the repeatability of the indices created from the individual spot-check items. When only household level random effects were used, coefficients of variance ranged from 0.18 (DWI and FI) to 0.40 (DHI). Values for the coefficients of variance calculated with field worker and household level random effects were similar to those with only household random effects (Table 4). The three sets of ICCs were similar, though values for the ICC2 tended to be higher than those of ICC1 or ICC3. ICCs ranged from 0.16 (FI) to 0.51 (SHI) depending on the model (Table 4). For the SHI, to estimate a household's underlying level of hygiene within 20% would require six separate spot checks, regardless of whether field worker random effects were considered; the number of observations needed for specific indices varied from 3 to 15. To be within 50% of a household's true hygiene level, no more than two separate observations would be needed for any of the indices (Table 4).

Morbidity information was collected for 694 children less than 36 months representing 500 of the 613 households with at least one hygiene measure. The mean follow-up time for children was 839 ± 321 days; the mean number of incident episodes of diarrhoea was 3.0 ± 2.4, and the mean duration of episodes was 6.4 days. Associations between diarrhoea outcomes and hygiene index scores are presented in Table 5. A one-point increase in SHI and PHI scores was associated with a 2 and 6% decrease, respectively, in the odds of an incident episode of diarrhoea ($P < 0.05$). Those children in households scoring, on average, >9 on the SHI had a 14% reduction in the odds of an incident episode of diarrhoea in a given month compared with those scoring <5 on the SHI (OR =1.14; 95% CI 1.03, 1.30; P = 0.034). Similar findings were observed for

Table 3 Distribution of the hygiene spot-check observations used to create specific and summary hygiene indices at the first visit, second visit, and for all of the visits combined among those households with at least two observations

Index	% of first observations (n = 588)	% of second observations (n = 588)	% of total observations (n = 12 989)	Kappa (95% CI) first vs second observation[a]	Kappa (95% CI) from repeated observations[b]
Drinking water index (DWI)					
Interior water container water is covered	60.9	60.4	58.1	0.25 (0.17, 0.33)	0.05 (−0.014, 0.18)
Exterior water container is clean	48.8	55.8	48.0	0.38 (0.27, 0.43)	0.27 (0.18, 0.35)
Pila[c] contains water	89.17	90.2	93.6	0.38 (0.27, 0.50)	0.07 (−0.10, 0.19)
Food index (FI)					
Clean dishes are covered	12.2	11.5	8.5	0.14 (0.03, 0.25)	−0.03 (−0.09, 0.03)
Clean dishes are stored high	89.1	89.8	91.1	0.20 (0.09, 0.30)	−0.01 (−0.13, −0.02)
All food is covered	47.1	55.8	48.6	0.17 (0.09, 0.25)	0.09 (0.002, 0.19)
Personal hygiene index (PHI)					
Mother/caregiver is wearing shoes	68.4	64.8	64.5	0.39 (0.30, 0.47)	0.15 (0.04, 0.25)
Mother's/caregiver's hands are clean	65.3	67.9	55.1	0.29 (0.21, 0.38)	0.05 (−0.05, 0.14)
Index child's hands are clean	48.3	44.8	34.1	0.33 (0.25, 0.41)	−0.08 (−0.16, 0.01)
Domestic household hygiene index (DHI)					
No trash in yard	41.0	42.0	28.1	0.37 (0.31, 0.45)	−0.01 (−0.08, 0.05)
No trash inside home	57.1	59.9	47.6	0.33 (0.25, 0.41)	0.19 (0.10, 0.29)
No unrestrained animals	33.8	34.1	36.5	0.29 (0.21, 0.37)	0.26 (0.17, 0.35)
No dirty clothes accumulated in the home	53.4	67.9	61.6	0.20 (0.12, 0.27)	0.05 (−0.05, 0.14)
Insignificant quantity of flies inside the home	23.6	25.2	10.9	0.21 (0.16, 0.30)	−0.02 (−0.04, −0.01)
No standing water on patio of home	16.8	20.2	13.0	0.28 (0.18, 0.37)	−0.04 (−0.09, 0.01)

Kappa coefficients and 95% confidence intervals are presented for the first and second observation and for the combination of five randomly selected observations among those households followed for more than 1 year (n = 462).

[a] $N_{households}$ = 588; $N_{observations}$ = 1176.

[b] $N_{households}$ = 415; $N_{observations}$ = 2075.

[c] A pila is a general use water storage container.

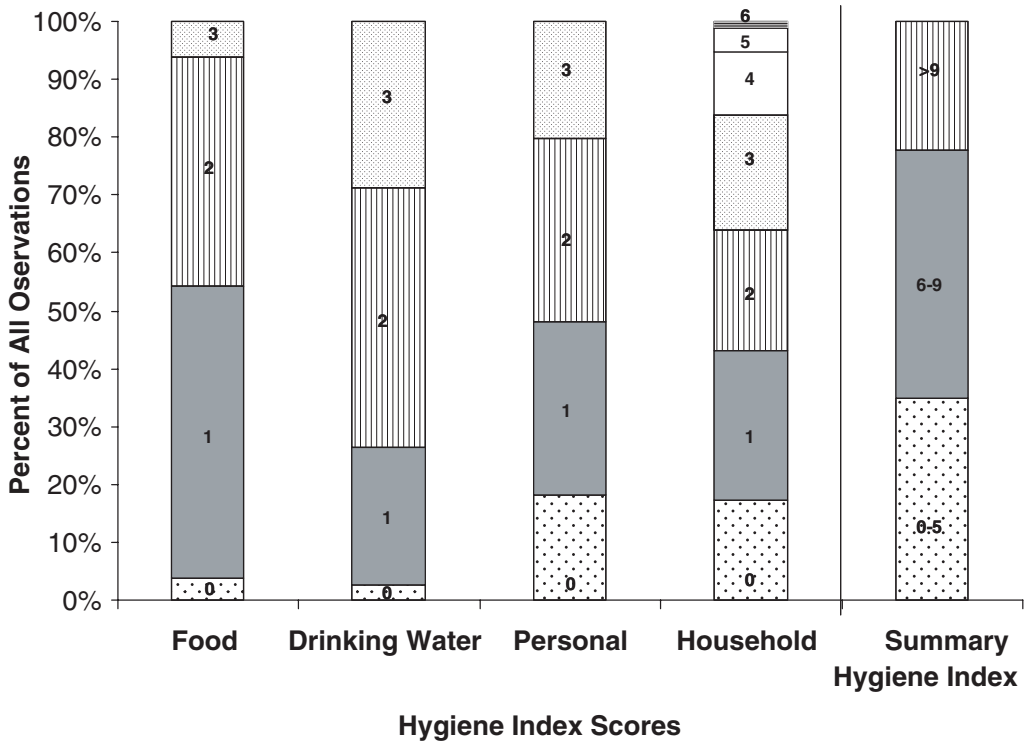

Figure 1 Distribution of scores for the specific hygiene indices and the SHI (n = 12 989 observations)

Table 4 Repeatability measures for monthly spot-check hygiene observations among 588 households in four rural Guatemalan villages

Hygiene index	Mean (x)	Repeatability values from variance estimates calculated with household (HH) as a random effect: Variance measure estimates[a] σ_{RES}	σ_{HH}	CV_w [b]	ICC1	Number of required observations[c] 20% of mean	50% of mean	Repeatability values from variance estimates calculated with household (HH) and field worker (FW) as random effects: Variance measure estimates[a] σ_{RES}	σ_{FW}	σ_{HH}	CV_w [b]	ICC2	ICC3	Number of required observations[c] 20%	50%
Summary	7.0 ± 2.9	5.2	3.1	0.3	0.4	6	1	2.8	2.5	2.9	0.2	0.5	0.4	6	1
Drinking water	2.0 ± 0.8	0.5	0.1	0.2	0.2	3	1	0.3	0.2	0.1	0.2	0.3	0.2	3	1
Food	1.5 ± 0.7	0.37	0.07	0.18	0.16	3	1	0.3	0.1	0.07	0.2	0.2	0.2	3	1
Personal	1.5 ± 1.0	0.7	0.3	0.4	0.3	12	2	0.6	0.2	0.3	0.4	0.3	0.3	12	2
Household	2.0 ± 1.5	1.5	0.6	0.4	0.3	15	2.5	0.9	0.6	0.6	0.4	0.4	0.3	13.9	2.2

Res, residual; HH, household; FW, field worker.

[a] Estimated from one-way and two-way random effects models; formulas provided in text.[21]

[b] $CV_w = (\sqrt{\sigma_w^2}/x)$.[22]

[c] N = the number of days required to be within 20% or 50% of the true mean score $[(Z_\alpha^* CV_w)/D_o]^2$ where $Z_\alpha = 1.96$ and $D_o = 0.2$ or 0.5.[22]

the PHI with respect to the number of episodes a child experienced per month ($OR_{1\ vs\ 3}$: 1.13; 1.01, 1.27; $P = 0.039$). Children in households that achieved a 3 on the PHI (all appropriate behaviours satisfied) were 16% less likely to have an incident episode of diarrhoea in the subsequent month compared with children in households with a PHI of zero ($OR_{0\ vs\ 3}$: 1.16; 1.00, 1.34; $P = 0.043$). Scoring a 3 vs a zero on the DWI for a given month was associated with a >30% reduction in the odds of an incident episode of diarrhoea (OR: 1.30; 1.00, 1.69; $P = 0.048$) and the number of episodes of diarrhoea (OR: 1.37; 1.03, 1.83; $P = 0.038$), though this relationship failed to remain significant in the adjusted model. A score of 3 vs a zero on the PHI was associated with a 26% reduction in the number of days ill with diarrhoea in the subsequent month (OR: 1.26; 1.01, 1.57; $P = 0.032$).

Discussion

We have described the stability over time and the repeatability of one summary and four specific hygiene indices that were derived from a 15 item spot check administered monthly over a 3 year period. Additionally, we have described the ability of each index to predict diarrhoeal illness in children <36 months.

In general, scores on the different hygiene indices deteriorated over time. The observed worsening in hygiene scores may have been a function of specific factors that only became evident because of the long period of follow-up time (3 years). To be eligible for enrolment in the study, women had to be pregnant or have a child <3 years of age. It is plausible that with increasing age and, thus, mobility of the child, the mother's diligence to hygiene decreased owing to greater energy expended on monitoring her child's activities. It is also plausible that mothers became more accustomed to the presence of observers (reactivity) as the study progressed and as such were not as likely to clean the home or improve their appearance prior to the arrival of the field worker. Reactivity has been reported as a problem when structured observations are employed to gather information on hygiene behaviour.[10] Studies employing spot-check methods have either not evaluated reactivity[11,12,26] or found that it was not a problem for domestic hygiene behaviours such as cleanliness of floors.[14] Reactivity of mothers to field-worker presence was not assessed in this study but could be important for understanding the influence of field worker presence on changes in hygiene practices over time. Additional investigations of potential household or personal factors that may influence hygiene practices are needed. For example, changes in family size or structure, changes in work status of the mother or the socioeconomic situation of the home, or increasing child age could all contribute to changes in hygiene practices over time. Findings from such studies could be beneficial to the development and assessment of hygiene promotion campaigns.

Reliability, also known as repeatability, is defined by Mehrens and Lehmann as 'the degree of consistency between two measures of the same thing'[27] and lack of repeatability is a common problem in studies attempting to characterize individuals' behaviours, including hygiene, due to day-to-day variability and inconsistency across observers. Ruel and Arimond suggest that spot checks may be less susceptible to variability because spot checks are designed to measure

Table 5 Odds ratios (OR) and rate ratios (RR), and 95% confidence intervals (CI) for the associations between hygiene scores and diarrhoea morbidity among 694 rural Guatemalan children <36 months

	Occurrence of an incident episode in a given month			Number of episodes of diarrhoea per month			Number of days ill with diarrhoea per month		
Score	OR[a]	95%	CI	RR[a]	95%	CI	RR[a]	95%	CI
Summary hygiene index									
Reduced model									
1	1.18	1.05	1.33	1.13	0.99	1.29	1.17	0.99	1.4
2	1.08	0.97	1.19	1.06	0.95	1.18	0.99	0.86	1.15
3	1.00			1.00			1.00		
Full model[b]									
1	1.18	1.03	1.34	1.13	1.00	1.27	1.18	0.97	1.43
2	1.13	1.01	1.26	0.98	0.99	1.30	0.92	0.77	1.09
3	1.00			1.00			1.00		
Drinking water index									
Reduced model									
0	1.30	1.00	1.69	1.37	1.03	1.83	1.33	0.90	1.97
1	1.12	0.99	1.27	1.11	0.98	1.27	1.08	0.90	1.30
2	0.98	0.88	1.09	1.00	0.89	1.12	0.95	0.82	1.11
3	1.00			1.00			1.00		
Full model[b]									
0	1.29	0.95	1.76	1.38	0.99	1.92	1.13	0.77	1.66
1	1.06	0.92	1.21	1.06	0.91	1.23	1.03	0.84	1.25
2	0.93	0.83	1.05	0.94	0.83	1.07	0.89	0.75	1.05
3	1.00			1.00			1.00		
Food index									
Reduced model									
0	1.28	0.98	1.67	1.17	0.88	1.56	0.95	0.63	1.41
1	1.04	0.85	1.27	1.03	0.83	1.27	0.84	0.61	1.14
2	0.99	0.81	1.21	1.00	0.81	1.24	0.77	0.57	1.04
3	1.00			1.00			1.00		
Full model[b]									
0	1.22	0.89	1.69	1.17	0.84	1.65	1.02	0.62	1.66
1	1.00	0.79	1.26	1.03	0.80	1.31	0.85	0.58	1.25
2	0.97	0.77	1.22	1.01	0.79	1.29	0.78	0.54	1.13
3	1.00			1.00			1.00		
Personal hygiene index									
Reduced model									
0	1.17	1.01	1.36	1.14	0.97	1.33	1.26	1.01	1.57
1	1.12	0.99	1.27	1.09	0.96	1.25	1.12	0.93	1.36
2	1.06	0.93	1.20	1.03	0.90	1.18	1.08	0.89	1.31
3	1.00			1.00			1.00		
Full model[b]									
0	1.18	1.01	1.39	1.13	0.94	1.35	1.27	0.99	1.62
1	1.16	0.99	1.34	1.12	0.96	1.31	1.17	0.94	1.46
2	1.06	0.92	1.23	1.03	0.88	1.2	1.07	0.85	1.34
3	1.00			1.00			1.00		
Household hygiene index									
Reduced model									
0	1.08	0.99	1.18	1.06	0.96	1.16	1.04	0.92	1.19
1	1.00			1.00			1.00		
Full model[b]									
0	1.05	0.95	1.16	1.03	0.93	1.15	1.02	0.88	1.18
1	1.00			1.00			1.00		

[a] The highest hygiene score or tertile is the referent category.
[b] Model adjusted for age and sex of index child and household SES.

'proxies' to behaviours rather than the actual behaviours themselves.[9] We did not compare practices assessed by spot checks to practices assessed by other measures and, thus, cannot refute or support this hypothesis. We did observe that ICCs and kappa coefficients for the hygiene indices and the specific indicators assessed in the spot checks were low. The estimates observed in this study are consistent with repeatability estimates observed by Gorter *et al.*[14] and suggest that indicators of hygiene practices assessed by spot checks are subject to substantial day-to-day variations within households over long periods. When practices that exhibit high variability are used to classify exposure status, such as specific hygiene practices, the use of a single observation may result in misclassification of exposure. Composite indices that use multiple indicators to capture aspects of hygienic behaviour may be more dependable than the use of one specific indicator of behaviour. We found this to be the case. For example, the specific indicator 'mother's hands are clean' had poor repeatability over time. However, the index to which this behaviour relates was substantially more stable. The increase in stability notwithstanding, however, multiple observations are necessary to reliably capture the true hygienic score of the home.

In this study, variation attributed to field workers was not a significant contributor to within-household variation as might be expected in a long-term follow-up study when it is likely that several field workers observed the same household at different time points. The field workers in this study had extensive training and continuous supervision. Hence error due to poor standardization of methods or field-worker-related error was potentially minimized.

In addition to being reliable and non-reactive, it is important that measured indicators be related to a given outcome (predictability). We found improved hygiene scores to be associated with reduced odds of diarrhoea. The reduction observed in this study is smaller than that observed in other studies.[14,28] The predictability of the indices for diarrhoea highlights their potential use in classifying homes into risk categories for diarrhoea. While the indices do not allow isolation of the effect of a specific practice, they do allow for the capture of behavioural tradeoffs that may influence risk for diarrhoea. For example, a mother may not consistently maintain clean hands but other aspects of her home hygiene practices may be appropriate and consistent (for example, maintenance of clean and covered water supplies). Risk assessment for diarrhoea that utilizes only the indicator of clean hands may inaccurately reflect risk in a home if that indicator is highly variable or if other aspects of the mother's home hygiene practices are appropriate but not used in risk assessment. Use of composite indices may potentially produce a more comprehensive and stable assessment of the hygienic state of the home and may be more appropriate for classifying homes into risk categories.

The indices used in the current study are not without limitations. Studies that have evaluated hand washing have consistently shown that soap is a critical component to the risk-reduction capabilities of hand washing.[29] Had we been able to identify soap use during hand washing, we can hypothesize that we probably would have observed greater reductions in diarrhoea-related morbidity. Though we did not assess soap use, hand cleanliness in this population appeared more strongly related to the risk of diarrhoea than were other items on the spot check. The PHI consists primarily of hand-cleanliness indicators and hand washing has repeatedly been associated with reduced diarrhoea incidence.[3,29] No other specific indices were significantly associated with diarrhoeal morbidity. Likewise, additional hygiene and sanitation factors such as the presence of latrines or the quantity of water delivered to the house each day may enhance predictability and usefulness of the indices. However, these are unlikely to be subject to a similar level of day-to-day variation that is within the mother's locus of control and as such were not assessed in this study.

In conclusion, spot-check methods for observing indicators of hygiene could potentially be a rapid and efficient method for assessing household level hygiene. The composite indices summarized specific hygiene indicators, were more stable over time compared with the individual indicators and were predictive of diarrhoea-related events in young children. However, multiple assessments would be required to approximate the true hygiene pattern of a household. It is unclear whether instability was the result of reactivity or other factors related to changing circumstances within households over time. Composite indices may provide for a more accurate and consistent representation of a household's true hygienic state, which could be of potential use in assessment of hygiene promotion programmes or for use in risk assessment for diarrhoea in homes with young children.

Acknowledgements

This study was made possible through the long-lasting collaboration of the people from the four-village longitudinal study. The authors acknowledge Dale McLerran of the Fred Hutchinson Cancer Research Center for statistical advice. We also gratefully acknowledge funding from the National Institutes of Health, USA, grants HD-29927 and HD 046125.

KEY MESSAGES

- Spot-check methods may be a rapid and efficient method for assessing domestic hygiene practices.
- Scores on composite indices of hygiene practices are associated with childhood diarrhoea.
- Scores on composite indices of hygiene were low in rural Guatemala.
- Owing to high daily variability, multiple assessments of domestic hygiene need to be performed to accurately capture a household's true hygiene practices.

References

[1] Lanata CF, Black RE. Diarrheal diseases. In: Semba RD, Bloem MW (eds). *Nutrition and Health in Developing Countries*. Totowa: Humana Press, 2001.

[2] Aiello AE, Larson EL. What is the evidence for a causal link between hygiene and infections? *Lancet Infect Dis* 2002;**2:**103–10.

[3] Fewtrell L, Kaufmann RB, Way D *et al*. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2005;**5:**42–52.

[4] Martines J, Phillips M, Feacham RGA. Diarrheal diseases. In: Jameson DT, Mosley WH, Measham AR, Bobadillo JL (eds). *Disease Control Priorities in Developing Countries*. Oxford: Oxford University Press, 1993.

[5] Curtis V, Cousens S, Mertens T *et al*. Structured observations of hygiene behaviours in Burkina Faso: validity, variability, and utility. *Bull World Health Organ* 1993;**71:**23–32.

[6] Manun'ebo MN, Cousens S, Haggerty PA *et al*. Measuring hygiene practices: a comparison of questionnaires with direct observations in rural Zaire. *Trop Med Int Health* 1997;**2:**1015–21.

[7] Odujinrin OM, Akitoye CO, Odugbemi T, Oyerinde JP, Esumeh FI. Ethnographic study on childhood diarrhoeal diseases in a rural Nigerian community. *West Afr J Med* 1993;**12:**185–88.

[8] Stanton BF, Clemens JD, Aziz KMA, Rahman M. Twenty-four hour recall, knowledge-attitude-practice questionnaires, and direct observations of hygiene behaviors in Burkina Faso: validity, variability, and utility. *Bull World Health Organ* 1987;**65:**217–22.

[9] Ruel M, Arimond M. Spot-check observational method for assessing hygiene practices: review of experiences and implications for programmes. *J Health Popul Nutr* 2002;**20:**65–76.

[10] Cousens S, Kanki B, Seydou T, Diallo I, Curtis V. Reactivity and repeatability of hygiene behavior: structured observations from Burkina Faso. *Soc Sci Med* 1996;**43:**1299–308.

[11] Amar-Klemesu M, Ruel MT, Maxwell MT, Levin CE, Morris SS. Poor maternal schooling is the main constraint to good child care practices in Accra. *J Nutr* 2000;**130:**1597–607.

[12] Bartlett AV, Hurtado E, Schroeder DG, Mendez H. Association of indicators of hygiene behavior with persistent diarrhea of young children. *Acta Paediatr Suppl* 1992;**381:**66–71.

[13] Kaltenthaler EC, Drasar BS. The study of hygiene behavior in Botswana: a combination of qualitative and quantitative methods. *Trop Med Int Health* 1996;**1:**690–98.

[14] Gorter AC, Sandiford P, Pauw J *et al*. Hygiene behaviour in rural Nicaragua in relation to diarrhoea. *Int J Epidemiol* 1998;**27:**1090–100.

[15] Habicht JP, Martorell R. Objectives, research design and implementation of the INCAP longitudinal study. *Food Nutr Bull* 1992;**14:**176–90.

[16] Stein AD, Barnhart HX, Wang M *et al*. Comparison of linear growth patterns in the first three years of life across two generations in Guatemala. *Pediatrics* 2004;**113:**e270–75.

[17] Martorell R, Habicht JP, Rivera JA. History and design of the INCAP longitudinal study (1969-1977) and its follow up (1988-1989). *J Nutr* 1995;**125:**1027S–41S.

[18] Kuklina EV, Ramakrishnan U, Stein AD, Barnhart HH, Martorell R. Growth and diet quality are associated with the attainment of walking in rural Guatemalan infants. *J Nutr* 2004;**134:**3296–300.

[19] Zeger SLL, Liang KY. Longitudinal data analysis for discrete and continous outcomes. *Biometrics* 1986;**42:**121–30.

[20] Hall DB, Severini TA. Extended generalized estimating equations for clustered data. *J Am Statist Assoc* 1998;**93:**1365–75.

[21] Shrout PE, Fleis JL. Intra-class correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**86:**420–22.

[22] Beaton GH, Milner J, Corey P *et al*. Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am J Clin Nutr* 1979;**32:**2456–559.

[23] Guerrant RL, Kirchhoff LV, Shields DS *et al*. Prospective study of diarrheal illnesses in northeastern Brazil: patterns of disease, nutritional impact, etiologies, and risk factors. *J Infect Dis* 1983;**148:**986–97.

[24] Rowland MG. The Gambia and Bangladesh: the seasons and diarrhoea. *Dialogue Diarrhoea* 1986;**26:**3.

[25] Mondal NC, Biswas R, Manna A. Risk factors of diarrhoea among flood victims: a controlled epidemiological study. *Indian J Public Health* 2001;**45:**122–27.

[26] Merchant SS, Udipi SA. Positive and negative deviance in growth of urban slum children in Bombay. *Food Nutr Bull* 1997;**18:**323–36.

[27] Mehrens WA, Lehmann IJ. *Measurement and Evaluation in Education and Psychology*. Philadelphia: Holt, Rhinehart and Winston, 1991.

[28] Dikassa L, Mock N, Magnani R *et al*. Maternal behavioural risk factors for severe childhood diarrhoeal disease in Kinshasa, Zaire. *Int J Epidemiol* 1993;**22:**327–33.

[29] Cairncross S, Curtis V. Effect of washing hands with soap on diarrhoea risk in the community: a systematic review. *Lancet Infect Dis* 2003;**3:**275–81.