Estimating the Longitudinal Prevalence of Diarrhea and Other Episodic Diseases

Continuous Versus Intermittent Surveillance

Wolf-Peter Schmidt, Stephen P. Luby,† Bernd Genser,‡ Mauricio L. Barreto,‡ and Thomas Clasen**

Background: Longitudinal prevalence (ie, the proportion of time with the disease) is used to describe morbidity from diarrhea and other episodic conditions. The aim of this analysis was to compare estimates of longitudinal prevalence based on intermittent sampling at regular intervals with 24- or 48-hour recall, with estimates based on continuous surveillance.

Methods: Based on 2 real datasets from Brazil and Guatemala, we developed a simulated dataset representing the diarrhea morbidity of 10,000 individuals followed over 365 days.

Results: Both the model and the real datasets showed that the standard deviation of the longitudinal prevalence increases with decreasing numbers of days sampled, so that a study sampling only a fraction of days would require a larger sample size. However, due to the correlation of diarrhea between consecutive days, sampling at 7- to 14-day intervals results in relatively small loss of precision and power compared with daily morbidity records, especially when the average diarrheal episode is long. A study based on morbidity data for every seventh day may require only a 5%-24% larger sample size than a study with daily records, depending on the average duration of episodes. Using a recall period of 48 hours instead of 24 hours increases power if the average episode is short.

Conclusions: The results question the necessity of continuous surveillance to estimate longitudinal prevalence. In addition to savings in cost and staff time, intermittent sampling of morbidity may improve validity by minimizing recall error and reducing the influence of surveillance on participants' behavior.

(*Epidemiology* 2007;18: 537–543)

Longitudinal prevalence (ie, the proportion of time with the disease in an individual) is an increasingly common outcome measure to assess diarrhea morbidity in observational and intervention studies.[1–5] Longitudinal prevalence has been suggested as a better predictor of nutritional status and mortality than diarrhea incidence.[6] Further, measuring longitudinal prevalence avoids analytic problems in defining disease episodes[7,8] and the occurrence of repeated episodes that arise when using disease incidence as an outcome measure.[6] However, little work has been done on the question of whether longitudinal prevalence should be estimated based on continuous morbidity surveillance or on sampling days of observation at intervals.[9]

Using continuous diarrhea surveillance is problematic. Diarrhea has been shown to be underreported if the recall period is longer than 2 or 3 days.[10–12] Some investigators have used frequent household visits (2 to 3 per week) to shorten the recall period and minimize the potential for under-reporting.[2,13,14] However, household visits are highly resource-intensive. Also, close disease surveillance by frequent visits may affect the risk and reporting behavior of the study participants (ie, the Hawthorne effect). Where surveillance includes treatment or referral to treatment for ethical purposes, it may change the natural course of disease (ie, prevent mild episodes from becoming severe). The burden of participation may also compromise the willingness of households to cooperate. A number of studies have noted a decline of disease occurrence over time that could not fully be explained by seasonal variation, suggesting that motivation to report disease may decrease during the course of a study.[15–17] Reporting fatigue may be particularly pronounced if the method of disease assessment is time-consuming for the participants.

Morris and colleagues[9] pointed out that it may not be necessary to obtain a record of all days of observation to estimate the longitudinal prevalence. They found that a sample of 72 days of observation of a total of 365 days actually observed was adequate to reliably estimate the longitudinal prevalence in an individual. However, most recent intervention studies of diarrheal illness using the longitudinal prevalence as the outcome measure have calculated only the population-level longitudinal prevalence, without looking at the individual level. Our aim was to explore the effect of sampling a fraction of days on the precision of the population

Submitted 27 September 2006; accepted 10 February 2007.

From the *Infectious and Tropical Diseases Department, London School of Hygiene and Tropical Medicine, London, United Kingdom, †International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Bangladesh, ‡Instituto de Saude Coletiva, Federal University of Bahia, Salvador, Bahia, Brazil.

Supported by Unilever Ltd.

Correspondence: Wolf-Peter Schmidt, Infectious and Tropical Diseases Department, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom. E-mail: Wolf-Peter.Schmidt@lshtm.ac.uk.

ISSN: 1044-3983/07/1805-0537
DOI: 10.1097/EDE.0b013e318093f3ce

estimate of the longitudinal prevalence and the implications for the sample size.

METHODS

We addressed the study objectives by developing a simulated dataset based on the methods published by Morris and colleagues.[8] The simulated dataset reflected the diarrhea occurrence in a population, with episodes being randomly distributed over time, but highly clustered in individuals, with the majority of individuals having a small number of episodes and few people experiencing many episodes. The model was parameterized based on 2 real datasets with daily morbidity data over the course of at least 1 year. The first dataset was based on 2 longitudinal cohort studies conducted in 1997–1999 and 2000–2002 in Salvador de Bahia in Brazil in children under the age of 5.[14,18,19] Diarrhea was assessed by twice-weekly household visits, ie, recall periods of 3–4 days. We combined the 2 cohorts into one dataset, matched by calendar date to preserve the seasonal pattern of diarrhea occurrence. The dataset was reduced to 365 days of observation to enable comparability with the simulation. The study represented an open cohort. The number of days of observation varied among individuals. In the reduced dataset (n = 1839), the mean number of days under observation was 263 days (range, 2–365). The weighted mean number of episodes per individual in the reduced dataset (weighted by the number of days of observation) was 3.7. The mean episode duration was 2.5 days, whereas the weighted mean longitudinal prevalence (weighted by the number of days of observation) was 2.9%.

The second dataset was from a randomized-controlled trial studying the effect of different household water treatment techniques on diarrhea occurrence in rural Guatemala.[20] This study was a closed cohort (n = 2982) with equal follow-up time of just over 365 days, with disease being assessed in all age groups by weekly visits (ie, recall periods of around 7 days). We reduced the dataset to 365 days of observation, again ensuring that each day in the dataset represented one calendar day. All 5 study arms were combined. The mean number of episodes per individual in the reduced dataset was 1.7. The mean episode duration was 5.1 day, and the mean longitudinal prevalence was 2.5%. Thus, the 2 real datasets differed in the distribution of the number of episodes and the episode duration (presumably due to differences in age range, study setting or procedures) while being similar with regard to the mean longitudinal prevalence.

Development of the Simulated Dataset

We generated a dataset representing the daily diarrhea experience of 10,000 individuals over a period of 365 days. First, we created a variable that defined the number of diarrhea episodes per individual following a gamma distribution (a distribution suitable to represent highly skewed data such as diarrhea incidence).[8] The parameters of the gamma distribution were chosen so that the distribution of the number of episodes was within the bounds of the 2 real datasets (Fig. 1A); the parameters chosen for this default model were $\alpha = 0.48$ (shape parameter) and $\beta = 6$ (stretch parameter). This resulted in a mean number of 2.9 episodes per individual, which was less than assumed by Morris et al[8] based on data from Peru (mean number of episodes 9.0, $\alpha = 1.5$, $\beta = 6$).

In the next step we distributed the episodes of each individual randomly over the period of 365 days. Each episode was then randomly allocated an episode duration following an exponential distribution $y = \exp(kx)$ with $k = -0.32$, which resulted in a distribution of episode durations between the 2 real datasets (Fig. 1B), with a mean duration of 3.8 days and a mean longitudinal prevalence of 2.7%.

Different episodes were allowed to overlap. Thus, the average "observed" number of episodes in the simulated dataset was slightly lower than the number of episodes as given by the gamma distribution (2.6 vs. 2.9). Likewise, the average "observed" episode duration was slightly longer then determined by the negative exponential distribution (3.8 vs. 3.6).

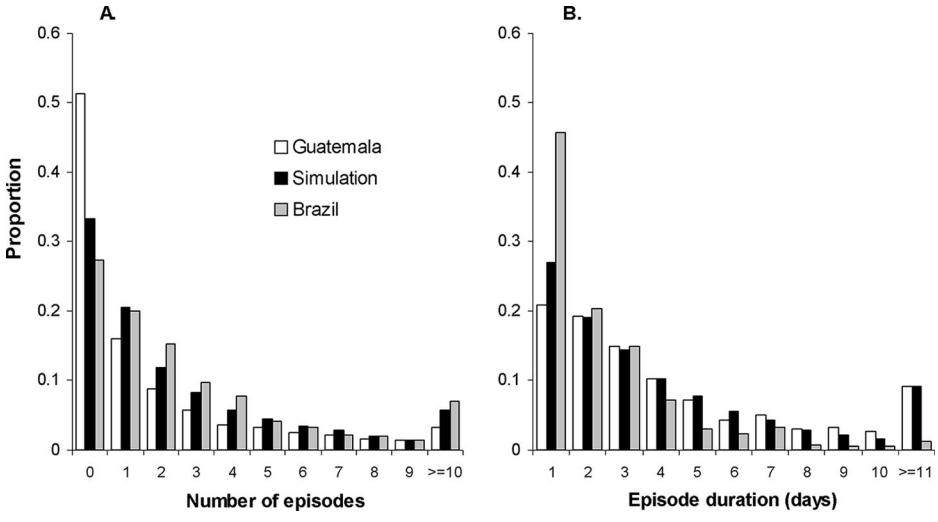

FIGURE 1. Distribution of the (A) number and (B) duration of episodes in the default model and for Brazil and Guatemala. For comparison, episodes were regarded as distinct if separated by at least 1 day; for the number of episodes the simulation assumes a gamma distribution with $\alpha = 0.48$ (shape) and $\beta = 6$ (stretch parameter); the simulated episode durations follow an exponential distribution ($y = \exp(-0.32x)$).

Choice of Outcome Measures

Unlike the conventional definition of prevalence, the longitudinal prevalence is a continuous variable. An individual can experience a longitudinal prevalence between 0% and 100%. For diarrhea, the distribution of the individual values of the longitudinal prevalence is highly skewed, with most individuals experiencing a longitudinal prevalence of up to one percent (Fig. 2, highest bars in each panel), and few individuals having a high longitudinal prevalence. Thus, the mean longitudinal prevalence may at first sight not seem to be an appropriate outcome measure to describe the longitudinal prevalence on population level. However, the mean longitudinal prevalence in a population is equal to the average diarrhea point prevalence over the study period (ie, the mean of the daily point prevalences) if each point prevalence is based on the same number of individuals. The mean longitudinal prevalence is also equal to the proportion of diarrhea days among all days observed in the study population, provided that all individuals were followed for the same number of days. If the follow-up time varies among participants, then the weighted mean longitudinal prevalence (weighted by the number of days of observation) is equal to the mean point prevalence over time weighted by the number of individuals contributing to the daily point prevalence. This again is equal to the proportion of days with diarrhea in the population, which certainly is an outcome measure of interest.

To describe precision, we used the standard deviation (SD), which allows straightforward sample size calculations for studies using the longitudinal prevalence as an outcome measure. With decreasing days of observation per individual, the SD of individual longitudinal prevalence values is expected to rise, indicating the loss of precision with fewer days of observation per individual. The increase of the SD with fewer numbers of visits is proportional to the increase of the standard error of the mean longitudinal prevalence ($SD/\sqrt{n}$), since the overall number of individuals remains unchanged. The loss of precision of the longitudinal prevalence estimate has implications for the required sample size of a study using the longitudinal prevalence as an outcome measure. We applied a standard formula for the comparison of 2 means ($n = (0.84 + 1.96)^2 (2*SD^2)/(mean1 - mean2)^2$), because most diarrhea studies compare 2 or more groups. For illustration, we assumed a 30% reduction of the mean longitudinal prevalence, 80% power and $\alpha = 0.05$.

Simulation Procedures

We simulated the assessment of the longitudinal prevalence by assuming increasing intervals between visits over 365 days. At each visit, we simulated a 24-hour and a 48-hour recall period, which have been suggested as the optimum periods to achieve a high level of recall.[10–12] For simplicity, we assumed perfect recall.

To determine the robustness of our findings to changes in the model assumptions, we varied the parameters of the default simulation by assuming plausible high and low values for number of episodes and duration of illness as suggested by the 2 datasets from Brazil and Guatemala and by the parameterization used by Morris et al.[8] We also explored the effect of seasonal variation of illness incidence by relocating one-third of the episodes at random from the second half of the simulated time period to the first half, resulting in a diarrhea incidence that was twice as high during the first half of the year as in the second half. Simulations and analyses were performed with STATA 9.0 (StataCorp, College Station, TX).

RESULTS

Figure 3 shows the association between number of days of observation and SD of the longitudinal prevalence for the simulation model. It demonstrates the loss of precision and power that occurs in the default model when decreasing numbers of days of observation are sampled. As the whole

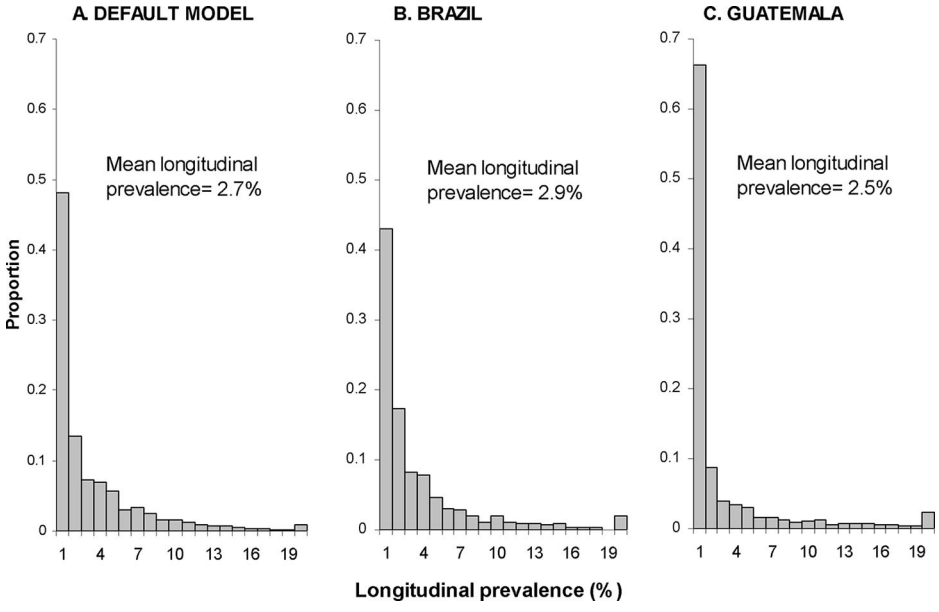

FIGURE 2. Distribution of individual values for the longitudinal prevalence in (A) the default model, (B) Brazil, and (C) Guatemala. LP, longitudinal prevalence; mean LP for Brazil weighted by the number of days of observation.

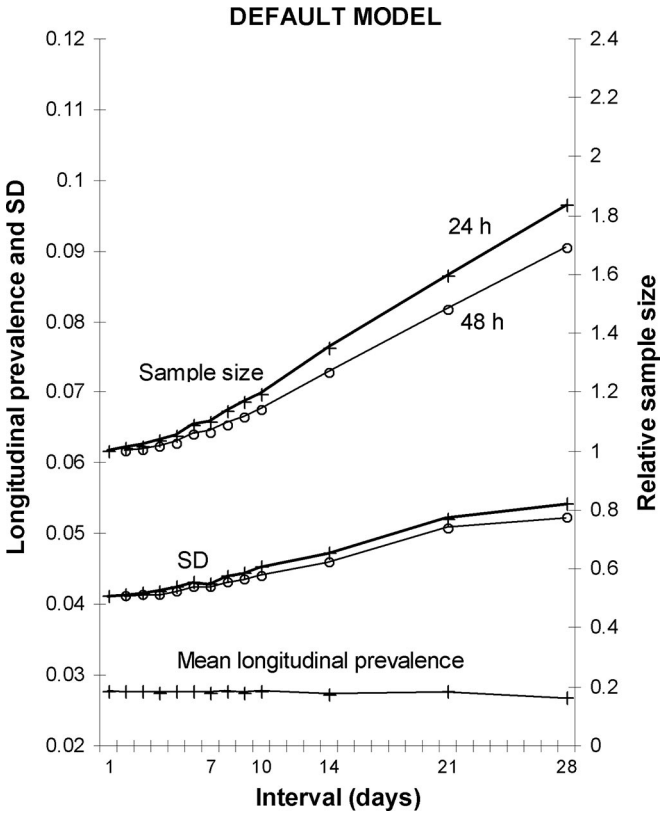

FIGURE 3. Mean longitudinal prevalence, standard deviation and relative sample size for different sampling intervals over 365 days for the simulated dataset (default model). LP, longitudinal prevalence; SD, standard deviation of the mean LP; relative sample size (right y-axis) indicates increase compared with baseline sample size if all days are sampled (n = 390 per arm); sample size calculation for comparison of 2 mean LP estimates.

sampling period covers 365 days, a 7-day surveillance interval amounts to 53 visits overall, a 14-day interval to 27 visits, and so on. The left y-axis reflects the mean longitudinal prevalence, which remains constant, and the SD of the mean longitudinal prevalence, which rises with longer surveillance intervals (and decreasing number of visits). As the SD enters into the sample size formula for the comparison of 2 means to the square, the effect of reducing the number of days sampled on the relative increase of the sample size is more pronounced (right y-axis). For example, a study recording daily disease prevalence over 365 days would require a sample size of 390 persons per arm (baseline) in the default model simulation. Recording only every 28th day (14 visits) increases the SD by the factor 1.3 (from 0.041 to 0.054) and the sample size by the factor 1.8 (from 390 to 716 participants per arm). In this model there seems to be a slight benefit of applying a 48-hour recall period instead of a 24-hour period (sample size 659 vs. 716 for 28-day interval).

In the dataset from Brazil (Fig. 4A.) the rise of the SD (factor 1.5 for 28-day interval) and the sample size (factor 2.3 for 28-day interval) with decreasing numbers of days sampled is steeper than in the simulated dataset, although the increase in the sample size when relying on a 7-day interval is still limited. Using a recall period of 48 hours instead of 24 hours substantially reduces the need to increase the sample size.

Figure 4B shows the data from Guatemala, where the average duration of illness was much longer. For a sampling interval of 28 days, the SD increases by the factor 1.2, and the required sample size by a factor of just 1.6 compared with all days sampled. There is little benefit of sampling at intervals shorter than 10 days. Using a 48-hour recall offers a slight advantage only for long sampling intervals.

We explored the apparent association between illness duration and the loss of precision with decreasing numbers of visits by varying the input parameters of the simulated dataset. Figure 5A confirms that fitting the illness duration in the model to the distribution observed in Brazil (approximately an exponential curve with k = −0.6 and mean episode duration 2.4 days) results in a steep rise in the sample size that is reduced by using a 48-hour recall. For a 28-day interval, the sample size goes up by a factor of 2.6 (compared with sampling all 365 days), which is similar to the Brazil dataset (factor of 2.4). Likewise, increasing the illness duration to what was observed in Guatemala (k= −0.25; mean episode duration = 4.6 days) leads to a sample size increase equal to the Guatemala dataset (factor 1.6 for a 28-day interval). In contrast, varying the incidence of diarrhea by fitting the parameter for the gamma distribution to the data from Brazil (approximately $\alpha = 0.56$, $\beta = 6$) and Guatemala ($\alpha = 0.3$, $\beta = 6$), while leaving the illness duration constant, does not have a strong effect on the proportional increase of the sample size (Fig. 5B). Although the overall sample sizes are highly dependent on the disease incidence, the relative increase in the required sample size for decreasing numbers of surveillance visits changes little, even if assuming a very high disease incidence based on the Ghana/Peru model ($\alpha = 1.5$, $\beta = 6$, mean longitudinal prevalence = 8.3%). The main findings for weekly and fortnightly visits with the key parameters for illness duration and incidence are summarized in Table 1.

Assuming a 2-fold higher incidence in the first half year to simulate seasonality while keeping the number of episodes unchanged resulted in a slightly steeper curve compared with the default simulation model (factor 2.0 for 28 days interval instead of 1.8), indicating that seasonal variation has only a limited impact on our findings.

Finally, we explored the implications for the precision and the sample size if the number of visits remains constant, while varying the length of the sampling intervals between visits. In other words, we simulated a situation where a fixed number of visits per household is spread over different overall durations of a study. Figure 6 shows that for 25 visits in the default model and the long episodes model (k = −0.25) with 24-hour recall, the power of a study can be maximized by applying at least a 10–14 day interval (equivalent to a study duration of 241–337 days). If the average duration of illness is smaller, as in the Brazil dataset (k = −0.6), the surveillance intervals can be shorter (around 7

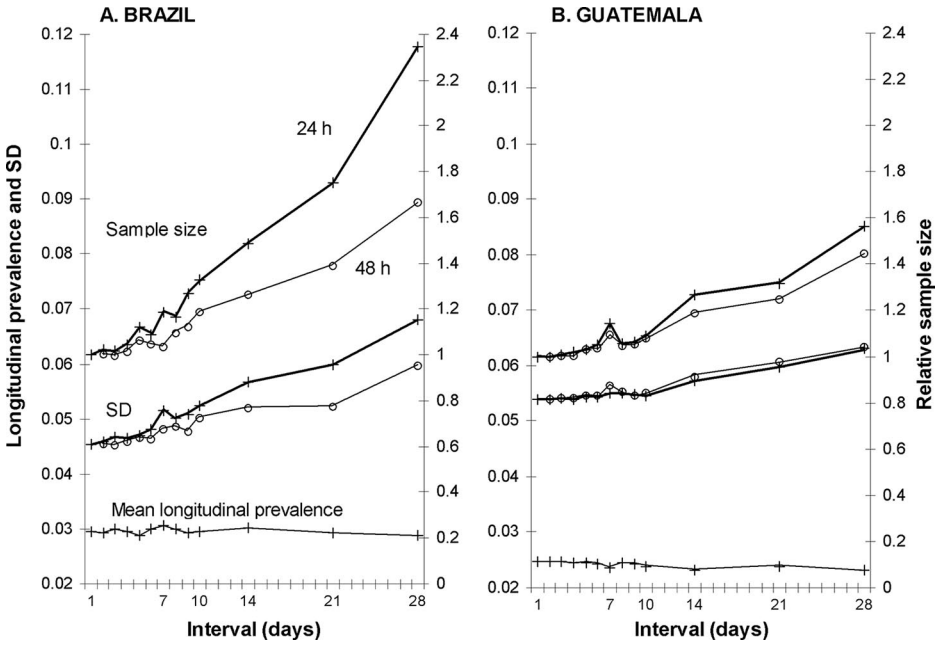

FIGURE 4. Mean longitudinal prevalence, standard deviation, and sample size for different sampling intervals over the course of 365 days for the (A) Brazil and (B) Guatemala datasets. Mean LP and SD for Brazil weighted by the number of days of observation; relative sample size (right y-axis) indicates increase compared with baseline sample size if all days are sampled (Brazil: 415; Guatemala: 833 per arm); sample size calculation for comparison of 2 mean LP estimates.

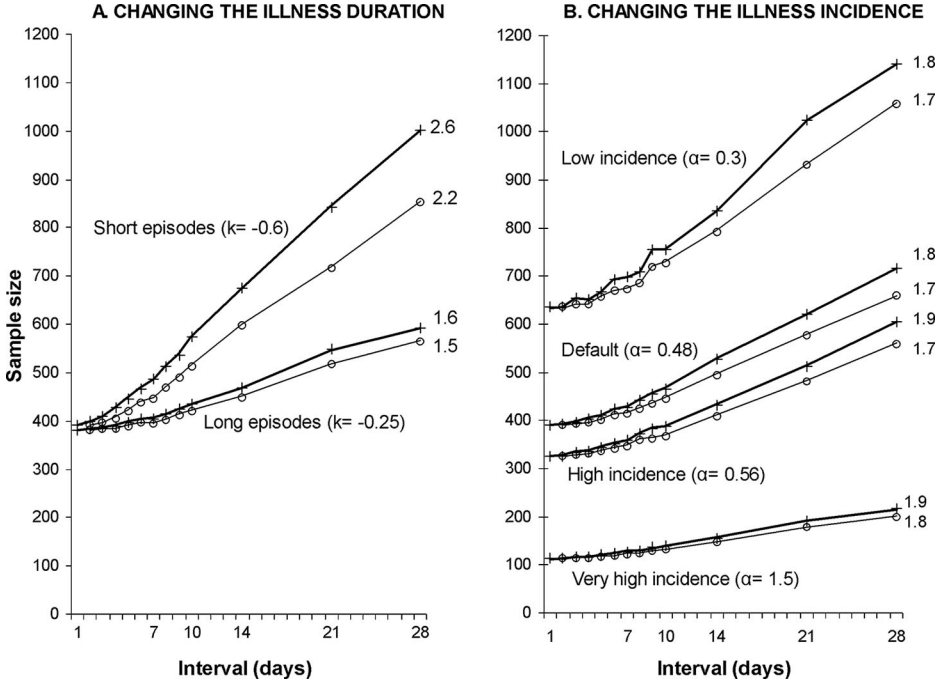

FIGURE 5. Effect of changing the simulated duration and the number of episodes on the sample size. Sample size per arm (y-axis) for the comparison of 2 mean longitudinal prevalence estimates; k is the parameter of the exponential distribution $y = \exp(kx)$ representing episode durations; α is the shape parameter of the gamma distribution representing the number of episodes in an individual (stretch parameter kept constant at $\beta = 6$); numbers on right side of both graphs indicate sample size increase relative to baseline (all days sampled).

days, upper line). This was broadly confirmed by applying this approach to the 2 real datasets, although the marked fluctuations in diarrhea occurrence over the year made it difficult to achieve comparable estimates for different study durations (not shown). The curves for a 48-hour recall leveled off in a very similar way (not shown). Again, varying the number of episodes (as in Fig. 5B) had little impact on the slope, showing that the preferred interval does not depend on the incidence (not shown).

DISCUSSION

Our results suggest, that in many situations, sampling every seventh day yields only slightly less precise estimates of the mean longitudinal prevalence of diarrhea in a population than collecting disease records for every single day. To achieve the same precision, a study based on morbidity data for every seventh day may require a 5%–24% larger sample size than a study with daily records, depending on the average duration of episodes (Table 1). In settings with short epi-

TABLE 1. Examples of Sample Sizes for a Hypothetical Diarrhea Intervention Trial Based on the Simulated Datasets and Different Sampling Strategies

	Surveillance Scheme				
		Visit Every Seventh Day		Visit Every 14th Day	
	Daily Records	24-Hour Recall	48-Hour Recall	24-Hour Recall	48-Hour Recall
Default dataset (as in Fig. 3; mean episode length 3.8 d)	390	428	415	527	494
Changing episode length					
Short episode dataset (k = −0.6*; mean length = 2.4 d)	391	486	446	675	597
Long episode dataset (k = −0.25*; mean length = 4.6 d)	382	404	395	468	448
Changing incidence					
Low incidence dataset ($\alpha = 0.3^{\dagger}$; mean number of episodes = 1.6)	636	698	673	835	793
High incidence dataset ($\alpha = 0.56^{\dagger}$; mean number of episodes = 3.0)	325	358	347	431	408

Sample size per group for the comparison of 2 mean longitudinal prevalence estimates assuming a 30% reduction in the intervention group, $P = 0.05$, 80% power.
*k is the parameter of the exponential distribution y = exp(kx) representing episode durations.
†α is the shape parameter of the gamma distribution representing the number of episodes in an individual (stretch parameter kept constant at $\beta = 6$).

sodes, the increase in the sample size can be reduced (in our simulated example from 24% to 14% for 7-day intervals) by applying a 48-hour recall period.

Recording daily disease occurrence (eg, by frequent visits or by relying on a long recall period) is resource-intensive, could affect the reporting and risk behavior of the study population, and could produce imprecise estimates. Unless the study requires measuring the incidence of diarrhea or close surveillance for other reasons (such as prompt treatment or collection of stool samples), intermittent sampling with a 24- or 48-hour recall period could improve reporting and reduce expenses. Compared with daily records, the increase in required sample size with sampling intervals of up to 14 days appears moderate. However, interval and sample size also depend on the expected average duration of illness, which affects the degree to which consecutive days are correlated. Likewise, depending on the average episode length, investigators may opt for either a 24-hour recall period, which may be simpler and more precise, or a 48-hour recall period, which still seems to yield valid data[10–12] but with little advantage if the average illness duration is long.

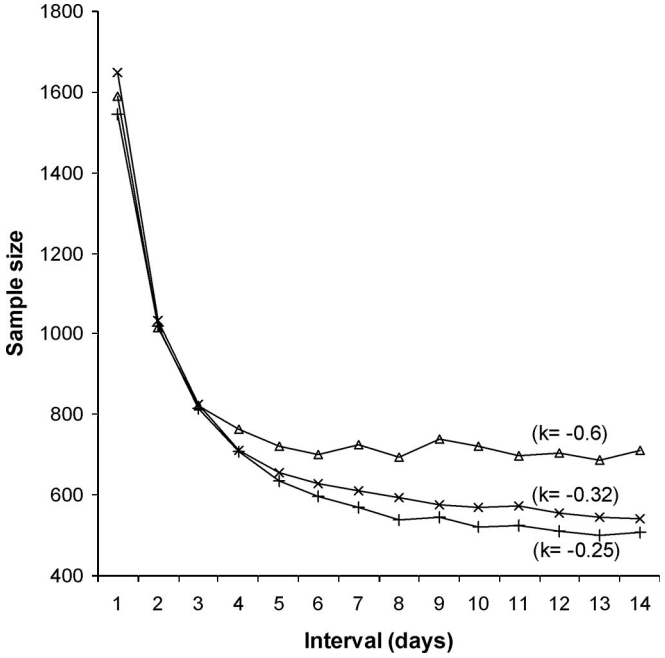

FIGURE 6. Effect of increasing the sampling intervals for a fixed number of visits (n = 25). Sample size per arm (y-axis) for the comparison of 2 mean longitudinal prevalence estimates; upper line (Δ) indicates simulation with short episodes as observed in Brazil; middle line (x) shows default simulation model; lower line (+) shows simulation with long episodes similar to Guatemala; k is the parameter of the exponential distribution y = exp(kx) representing episode durations.

Often the logistical constraints lie not so much in the overall duration of the study as in the total number of visits performed. For a fixed number of visits to each household, spacing disease recordings to at least 7–14 days maximizes efficiency. Investigators may thus choose to employ a small number of well-trained field workers for a longer time, rather than a large group of field workers for a short and intensive period. A longer duration of a study has the additional advantage that it better captures seasonal variations in disease occurrence.

Our findings confirm some aspects of previous work by Morris and colleagues[9] who used real datasets to show that sampling morbidity every 5 days can reliably classify study participants into longitudinal prevalence quintiles. By using simulated datasets and comparing with real data, we have identified illness duration as the key parameter for estimating the longitudinal prevalence on population level. Morris and colleagues had suggested that prevalence of disease is the main determinant of the required number of visits to estimate the longitudinal prevalence on individual level without specifically considering illness duration. Further, Morris and

colleagues proposed as a rule of thumb that visits should be separated by at least the average duration of episodes. Our analysis suggests that intervals between visits should be at least twice the average episode duration to maximize efficiency.

For this analysis, we made a number of simplifications that may affect the interpretation of the findings. We assumed perfect recall, without under- or overreporting of diarrhea occurrence within the chosen recall periods. A number of studies have previously suggested that applying either a 24- or 48-hour period results in a similarly high level of recall.[10–12] There is also evidence that the recall process is complex, with illness more than 48 hours earlier being underreported, or remembered as having occurred more recently, possibly leading to over-reporting of diarrhea.[10] However, although imprecise disease reporting may affect the size of the estimate, it is unlikely to affect the proportional loss of precision with decreasing number of visits, as identified in our analysis.

Further, the 2 real datasets were based on weekly/twice-weekly household visits with recall periods of up to 7 days, whereas we assumed a shorter recall to allow comparison with the simulation. Also, the study populations in Guatemala and Brazil both displayed variations in the diarrhea incidence over time. However, we found that the findings from the real datasets were well reflected by our model (Fig. 5) and that the effect of seasonality was limited.

Finally, one may question the choice of mean longitudinal prevalence and its SD as our outcome measures. Although the mean is influenced by extreme values, it also represents, in the case of the longitudinal prevalence, the population prevalence for all days of observation. Also, as long as extreme values of longitudinal prevalence are plausible and not due to measurement error, they are of public health interest because they are associated with poor nutritional status and higher mortality.[6] In this situation, presenting the mean longitudinal prevalence and comparing 2 groups with the *t* test may be a better approach than relying on nonparametric methods such as the median and the Wilcoxon test if the sample size is large.[21]

Instead of the SD or the standard error of the mean, we could have used the standard error of illness-days as a proportion of all days observed, taking into account the clustering of illness days in individuals. We applied this approach to the data from Brazil and found that the standard errors with increasing intervals were almost exactly proportional to the SD and the standard error of the mean longitudinal prevalence. Likewise, we could have used sample size formulae commonly used for clustered data. Most of these formulae include a measure of the between-cluster variation, which enters the formula as square. Although these formulae might have resulted in slightly different overall sample sizes, the proportional changes would have been very similar, if not identical.

In conclusion, our analysis suggests that the longitudinal prevalence of diarrhea can be efficiently estimated by periodic sampling while minimizing expense and inconvenience to study participants. Our findings have implications for the longitudinal prevalence of other episodic conditions and symptoms, such as respiratory infections, cough or fever.[6] Sampling only a fraction of days during a study period deserves to be tested in the field.

REFERENCES

1. Chiller TM, Mendoza CE, Lopez MB, et al. Reducing diarrhea in Guatemalan children: randomized controlled trial of flocculant-disinfectant for drinking-water. *Bull World Health Organ*. 2006;84:28–35.
2. Luby SP, Agboatwalla M, Feikin DR, et al. Effect of handwashing on child health: a randomised controlled trial. *Lancet*. 2005;366:225–233.
3. Prado MS, Cairncross S, Strina A, et al. Asymptomatic giardiasis and growth in young children; a longitudinal study in Salvador, Brazil. *Parasitology*. 2005;131:51–56.
4. Sharieff W, Bhutta ZA, Schauer C, et al. Micronutrients (including zinc) reduce diarrhea in children: the Pakistan sprinkles diarrhea study. *Arch Dis Child*. 2006;91:573–579.
5. van der Hoek W, Feenstra SG, Konradsen F. Availability of irrigation water for domestic use in Pakistan: its impact on prevalence of diarrhea and nutritional status of children. *J Health Popul Nutr*. 2002;20:77–84.
6. Morris SS, Cousens SN, Kirkwood BR, et al. Is prevalence of diarrhea a better predictor of subsequent mortality and weight gain than diarrhea incidence? *Am J Epidemiol*. 1996;144:582–588.
7. Baqui AH, Black RE, Yunus M, et al. Methodological issues in diarrheal diseases epidemiology: definition of diarrheal episodes. *Int J Epidemiol*. 1991;20:1057–1063.
8. Morris SS, Cousens SN, Lanata CF, et al. Diarrhea–defining the episode. *Int J Epidemiol*. 1994;23:617–623.
9. Morris SS, Santos CA, Barreto ML, et al. Measuring the burden of common morbidities: sampling disease experience versus continuous surveillance. *Am J Epidemiol*. 1998;147:1087–1092.
10. Alam N, Henry FJ, Rahaman MM. Reporting errors in one-week diarrhea recall surveys: experience from a prospective study in rural Bangladesh. *Int J Epidemiol*. 1989;18:697–700.
11. Boerma JT, Black RE, Sommerfelt AE, et al. Accuracy and completeness of mothers' recall of diarrhea occurrence in pre-school children in demographic and health surveys. *Int J Epidemiol*. 1991;20:1073–1080.
12. Ramakrishnan R, Venkatarao T, Koya PK, et al. Influence of recall period on estimates of diarrhea morbidity in infants in rural Tamilnadu. *Indian J Public Health*. 1999;43:136–139.
13. Kirchhoff LV, McClelland KE, Do Carmo PM, et al. Feasibility and efficacy of in-home water chlorination in rural North-eastern Brazil. *J Hyg (Lond)*. 1985;94:173–180.
14. Strina A, Cairncross S, Barreto ML, et al. Childhood diarrhea and observed hygiene behavior in Salvador, Brazil. *Am J Epidemiol*. 2003;157:1032–1038.
15. Genser B, Strina A, Teles C, et al. Risk factors for childhood diarrhea incidence: a dynamic analysis of a large longitudinal study in a Brazilian city. *Epidemiology*. 2006;17:658–667.
16. Quick RE, Kimura A, Thevos A, et al. Diarrhea prevention through household-level water disinfection and safe storage in Zambia. *Am J Trop Med Hyg*. 2002;66:584–589.
17. Semenza JC, Roberts L, Henderson A, et al. Water distribution system and diarrheal disease transmission: a case study in Uzbekistan. *Am J Trop Med Hyg*. 1998;59:941–946.
18. Strina A, Cairncross S, Prado MS, et al. Childhood diarrhea symptoms, management and duration: observations from a longitudinal community study. *Trans R Soc Trop Med Hyg*. 2005;99:407–416.
19. Barreto ML, Strina A, Prado M, et al. Saneamento básico e saúde: fundamentos científicos para avaliação do impacto epidemiólogico do programa de saneamento ambiental da Bahia de Todos os Santos (Bahia Azul). In: Heller ML, Monteiro TC, et al., eds. *Saneamento e Saúde nos Países em Desenvolvimento*. Rio deJaneiro: CC&P Editores; 1997:7–35.
20. Reller ME, Mendoza CE, Lopez MB, et al. A randomized controlled trial of household-based flocculant-disinfectant drinking water treatment for diarrhea prevention in rural Guatemala. *Am J Trop Med Hyg*. 2003;69:411–419.
21. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ*. 2000;320:1197–1200.