

The Variability of Childhood Diarrhea in Karachi, Pakistan, 2002–2006

Stephen P. Luby,* Mubina Agboatwalla, and Robert M. Hoekstra

Centers for Disease Control and Prevention, Atlanta, Georgia; Health Oriented Preventive Education, Karachi, Pakistan

Abstract. Diarrhea burden is often estimated using cross-sectional surveys. We measured variability in diarrhea prevalence among children < 5 years of age living in squatter settlements in central Karachi, Pakistan. We pooled data from non-intervention control households from studies conducted from 2002 through 2006. The prevalence of diarrhea varied on average by 29% from one week to the next, by 37% from one month to the next, and during peak diarrhea season by 32% from one year to the next. During 24 months when the same nine neighborhoods were under surveillance, each month the prevalence of diarrhea varied by at least an order of magnitude from the lowest to the highest prevalence neighborhood, and each neighborhood recorded the highest diarrhea prevalence during at least one month. Cross-sectional surveys are unreliable measures of diarrhea prevalence.

INTRODUCTION

Diarrhea is an important global health problem accounting for an estimated 1.9 million child deaths per year.¹ Childhood diarrhea is also associated with impaired growth and cognitive development.^{2,3} Sound measurements of the community burden of childhood diarrhea allow evaluation of trends, study of the relationship between diarrhea and predisposing factors, and assessment of the effectiveness of interventions.

Diarrhea burden in populations is often assessed with a single cross-sectional measurement.^{4,5} Indeed, this is the approach used by demographic and health surveys that draw a nationally representative population sample from an entire country and estimate the proportion of children < 5 years of age with diarrhea in the preceding 2 weeks.⁶ These nationally representative samples become benchmarks for evaluating progress in diarrheal prevention, and are used to compare the diarrhea burden in one country with another.^{6,7}

Longitudinal studies of diarrhea, suggest that diarrhea prevalence in a community varies over time, especially by season.^{8–11} Thus, the appropriateness of single measures of diarrhea incidence has been questioned.¹² However, few population-based long-term longitudinal data are available to assess the magnitude of this variability over time. Without an appreciation of the degree of this variability, we risk misinterpreting differences in diarrhea prevalence.

We conducted a series of studies in squatter settlements in central Karachi, Pakistan over 4 years that included active community-based surveillance for diarrhea in control households that were receiving no intervention.^{13–15} During these studies, the neighborhood water and sanitary infrastructure and food supply remained largely constant. We combined data from these separate studies to quantify the variation in diarrhea prevalence over time.

METHODS

Study population. Karachi is the largest city in Pakistan. Over 4 million Karachi residents live in squatter settlements,¹⁶ where neither people living in the area, nor those constructing dwellings have legal title to the land, and where water and

sanitary infrastructure are limited. Throughout the series of studies combined for this analysis, all of the participating households were located in one of several multiethnic squatter settlements in central Karachi. All of the communities were located within 6 kilometers of each other.

We conducted two controlled intervention studies and one follow-up study from 2002 through 2006. Intervention communities received supplies and a behavior change intervention to encourage regular hand washing with soap and/or treatment of household drinking water. Each study enrolled a non-intervention control group who received a regular supply of children's books, notebooks, pens, and pencils to help with their children's education, but no intervention to improve water or hygiene. Only households that included at least one child < 5 years of age, and had sufficient water supply for children to bathe daily, were eligible for enrollment.

Details on the enrollment and household selection for the studies have been previously reported.^{13–15} Briefly summarizing, in April 2002 we initiated the Soap Health intervention trial in Manzoor Colony and nearby communities.¹³ Field workers identified 42 candidate neighborhoods separated from one another by a street or market area, and enrolled 1,040 eligible households. Eleven of the candidate neighborhoods were randomly assigned to the control group. This control population was followed through March 2003. In April 2003 we initiated a new study, the Floc Health Study.¹⁴ Field workers identified 49 candidate neighborhoods, separated from one another by a street or market area; nine neighborhoods were randomly assigned to the control group. This control group was followed through December 2003. In 2004 no studies including weekly diarrheal surveillance were conducted. In July 2005 field workers revisited each of the households that had participated in the 2003 study and sought consent for re-enrollment and continued follow-up through December 2006 (The Long Term Practices Study).¹⁵

Data collection. Each of these studies was implemented by Health Oriented Preventive Education (HOPE), a local non-governmental organization that operates health services and educational and community development initiatives in the area. For each of the studies trained fieldworkers attempted to visit enrolled non-intervention households twice per week. Field workers asked the child's primary caregiver the same question for each of the studies. The English translation is "During how many days since my last visit has the child had diarrhea (3 or more loose stools in 24 hours)?" If fieldworkers successfully completed two field visits in a week, they added

*Address correspondence to Stephen P. Luby, ICDDR, B: International Centre for Diarrheal Disease Research, Bangladesh, GPO Box 128, Dhaka-1000, Bangladesh. E-mail: sluby@icddr.org

the number of days with diarrhea reported during the two field visits and recorded the number of days in the week with diarrhea. If, because a family was temporarily unavailable, more than 1 week had elapsed since the prior visit, the field worker then inquired only about diarrhea in the prior week.

Data management. Responses were marked on paper forms aggregated by week and double entered into an electronic database. The databases from all weeks of data collection from the separate studies were merged into a single data base for analysis. Because different studies used different strategies to collect data from persons > 5 years of age, this analysis was restricted to all children in the household who were < 5 years of age at the time of the weekly interview. Children born into households after initial enrollment were added into the population under surveillance. The experience of children once they reached 5 years of age was excluded from the analysis.

Pakistan Demographic and Health Survey. To evaluate if the patterns identified in the Karachi longitudinal studies were also seen in the large national cross-sectional Demographic and Health Surveys, we downloaded the data sets for the two completed Pakistan Demographic and Health Surveys, one conducted in 1991/2 and the second in 2006/7 (www.measuredhs.com). Both surveys identified the youngest child in the household and asked “Has (name of the child) had diarrhea in the last 2 weeks?”

Statistics. The primary outcome measure was daily longitudinal prevalence of diarrhea,¹⁷ i.e., the number of days reported with diarrhea divided by the number of days of observation. This was first calculated per person, and then summed for the various analyses. We assessed the relative change in the magnitude of diarrhea longitudinal prevalence using the following formula:

$$LPD_{\Delta} = |(LPD_2 - LPD_1)/LPD_2|$$

where

LPD_{Δ} - Relative change in longitudinal prevalence of diarrhea,
 LPD_1 - Longitudinal prevalence of diarrhea during time period 1,
 LPD_2 - Longitudinal prevalence of diarrhea during time period 2.

We measured variability by calculating the standard deviation and coefficient of variation of the weekly and monthly means of diarrhea prevalence within each of the longitudinal studies in Karachi. For the Pakistan Demographic and Health Survey data we measured the mean, standard deviation, and coefficient of variation of the reported prevalence over the preceding 2 weeks. We calculated the expected standard deviation based on sampling error of a binomial distribution for both weekly and monthly diarrhea prevalence within each longitudinal study and for the preceding 2 weeks for the Pakistan Demographic and Health Survey using the formula¹⁸:

$$WeeklyEsd = \sqrt{MeanWDia * (1 - MeanWDia) / 7},$$

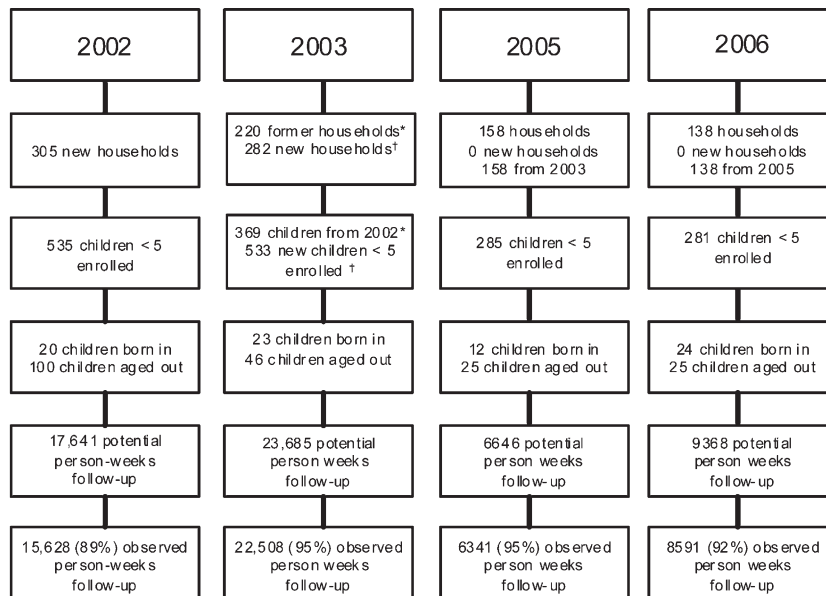
$$MonthlyEsd = \sqrt{MeanMDia * (1 - MeanMDia) / 30.4375},$$

$$BiWeeklyEsd = \sqrt{MeanBiWDia * (1 - MeanBiWDia) / 14},$$

where

WeeklyEsd - weekly expected standard deviation,
MeanWDia - mean weekly diarrhea prevalence,
MonthlyEsd - monthly expected standard deviation,
MeanMDia - mean monthly diarrhea prevalence,
BiWeeklyEsd - Two weekly expected standard deviation,
MeanBiWDia - mean two weekly diarrhea prevalence.

We stratified the data by study because the studies had different numbers of subjects. For the monthly calculation we only included child months with at least four weekly assessments.



*220 Households enrolled in 2002 were followed for the first 14 weeks of 2003.

†282 New households were enrolled in a new study in 2003

FIGURE 1. Enrollment and follow-up by year, Karachi squatter settlements.

The variability of diarrhea prevalence in the population can arise through variability of prevalence among different children and/or through variability that occurs within each child over time. We separately evaluated intra-child variability by identifying the children who had at least 10 weekly measurements of diarrhea. Within each study, for each child we calculated his/her average diarrhea prevalence, standard deviation, expected standard deviation, and the proportion of the observed standard deviation that was caused by sampling error. We excluded children without any reported diarrhea from this analysis, because the zero value for the measured standard deviation of prevalence would render the proportion of observed standard deviation caused by sampling error undefined.

Ethics. Heads of households provided informed consent. All children in each of the studies were assessed by field workers and referred to the appropriate level of care. All participants in the study were eligible for clinical care at HOPE facilities at no charge. Study protocols for each of the studies were approved by local and international human subject protection committees.

RESULTS

A total of 53,068 child-weeks of data were included in the analysis. The number of children < 5 years of age that initially enrolled varied from a low of 281 in 2006 to a high of 535 in 2002 (Figure 1). The percentage of potential child-weeks with completed follow-up ranged from 89% in 2002 to 95% in 2003. The households under surveillance between 2002 and 2006 were of similar size with a similar number of children < 5 years of age and < 2 years of age (Table 1). They had similar levels of parental literacy, a similar occupational profile for the father of the household, and similar household sanitary infrastructure. Households participating in 2005 and 2006 were less likely to report a monthly household income less than 3,000 rupees and more likely to own a refrigerator than in prior years (Table 1).

The mean longitudinal prevalence of diarrhea varied markedly from week to week. Among 145 evaluable weeks with a mean of 360 children evaluated each week, the weekly longitudinal prevalence of diarrhea varied by an average relative magnitude of 29%. Figure 2 illustrates an example of weekly variation from April 2002 through March 2003 when the largest number of children was under surveillance and the weekly mean longitudinal prevalence differed from the preceding week by an average relative magnitude of 24%.

There was a marked seasonal pattern to diarrhea in these Karachi squatter settlements. Overall, diarrhea prevalence was lowest December through March, and peaked in July through October. The average rates in August were 2.6 times higher than rates in December through February (Figure 3). Additionally, the peak month and the relative height of the peak for any given year varied considerably (Figure 4). During the 4 years of observation diarrhea prevalence peaked twice in August, once in September and once in October. Overall, the monthly mean longitudinal prevalence differed from the preceding month by an average relative magnitude of 37%.

Diarrhea prevalence varied markedly year to year (Figure 4). When the analysis was restricted to only those observations during peak diarrhea season, i.e., July through October, the

TABLE 1
Comparison of household characteristics across time

	2002 (N = 305)	2003 (N = 282)*	2005 (N = 158)	2006 (N = 138)
Baseline				
Mean				
Persons per household	9.1	9.0	10.1	10.2
Children < 5 years of age per household	1.8	1.9	1.8	1.8
Children < 2 years of age per household	0.6	0.7	0.7	0.6
Mean age of children < 5 at enrollment	2.7	2.5	2.2	2.6
Percent (%) of				
Mother of the youngest child is literate	98 (32)	105 (37)	60 (38)	52 (38)
Father of the youngest child is literate	196 (64)	175 (62)	96 (63)‡	83 (66)§
Monthly household income < 3,000 rupees†	151 (43)	123 (44)	26 (16)	22 (16)
Father's occupation				
Salaried employee	166 (55)	137 (49)	83 (54)‡	76 (57)§
Works for daily wages	127 (42)	118 (42)	56 (37)‡	46 (35)§
Other	12 (4)	23 (8)	14 (9)‡	11 (8)§
Owens refrigerator	116 (38)	66 (23)	71 (45)	61 (44)
Primary drinking water source				
Municipal supply within the house	29 (10)	91 (33)	48 (30)	41 (30)
Municipal supply at a community tap	45 (15)	104 (37)	15 (9)	14 (10)
Tanker truck	189 (62)	33 (12)	64 (41)	56 (41)
Water bearer	31 (10)	37 (13)	11 (7)	10 (6)
Tube well	9 (3)	15 (5)	4 (3)	3 (2)
Toilet without flush tank in the home	300 (99.7)	271 (96)	150 (98)‡	131 (99)§

* Information from the new households enrolled in 2003.

† 3,000 rupees ranged from US\$58 in April 2000 to \$50 in July 2005.

‡ Information from 153 households.

§ Information from 133 households.

average annual change in the relative magnitude of prevalence was 32% (Table 2).

The mean, the observed standard deviation, and the coefficient of variation of diarrhea prevalence were different for each of the three different studies (Table 3). The coefficient of variation ranged from 158% to 372%. The expected standard deviation based on sampling error of a binomial distribution accounted for 45–46% of the observed standard deviation in the weekly measurements and 33–38% in the monthly measurements (Table 3).

In the child-oriented weekly analysis, which separately calculated the mean diarrhea prevalence and standard deviation

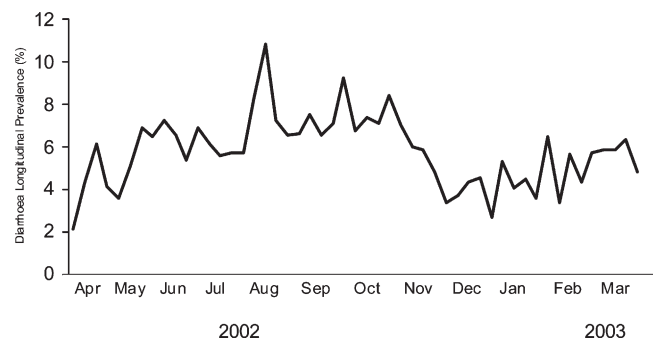


FIGURE 2. Mean diarrhea longitudinal prevalence by week among 489 children < 5 years of age, Karachi squatter settlements Apr 2002–Apr 2003.

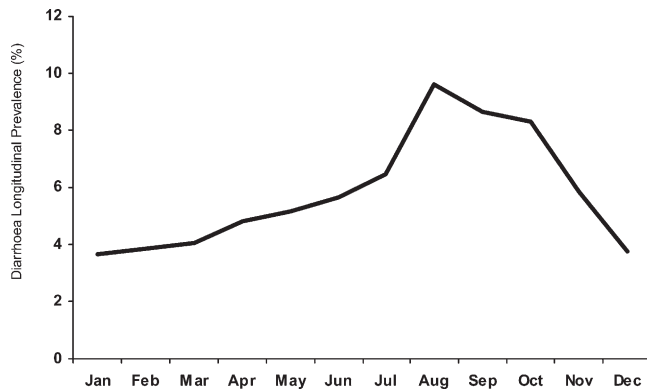


FIGURE 3. Longitudinal prevalence of diarrhea among children < 5 years of age by month, Mansoor, Mujahid, and Bilal Colonies, Karachi, Pakistan 2002–2006.

for each child throughout the course of the study, the expected standard deviation based on sampling error of a binomial distribution accounted for 49% of the observed standard deviation (50% for children followed in the Soap Health Study, 49% for children in the Floc Health Study and 49% for children enrolled in the Long Term Practices Study).

Younger children had more diarrhea, but when we restricted the analysis by age < 2 and age 2–5 years, the magnitude of weekly, monthly, and annual variability was not reduced (Table 4). Further narrowing the age groups also did not reduce variability (data not shown).

Diarrhea prevalence also varied markedly by neighborhood. For 9 months in 2003 and again during 15 months in 2005 and 2006 the same nine neighborhoods were under regular surveillance. The mean longitudinal prevalence for all the children < 5 years of age in these neighborhoods during the 24 months of observation was 6.6%. The neighborhood means ranged from 5.4% to 8.4%. For any given month the prevalence of diarrhea varied by at least an order of magnitude from the neighborhood with the lowest prevalence to the neighborhood with the highest prevalence (Figure 5). The mean number of observations per neighborhood per month was 151 (range 34–332). Specific neighborhoods had neither consistently low nor consistently high prevalence. During the 24 months of observation, each neighborhood recorded the highest diarrhea prevalence during at least 1 month.

Demographic and Health Survey (DHS). Of the 6,428 households surveyed in the 1991/2 Pakistan DHS, 5,956 (93%)

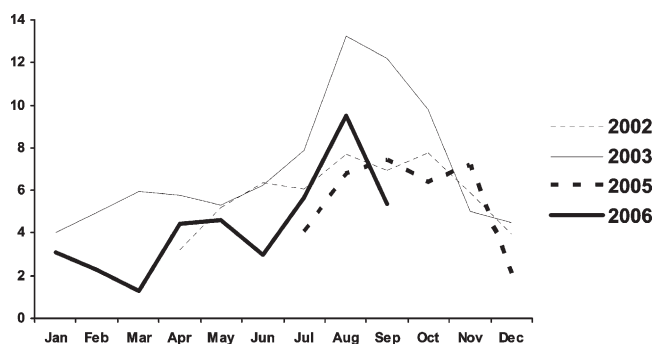


FIGURE 4. Longitudinal prevalence of diarrhea among children < 5 years of age by month and year, Mansoor, Mujahid, and Bilal Colonies, Karachi, Pakistan 2002–2006.

TABLE 2
Mean diarrhea prevalence July through October by year

	Longitudinal prevalence of diarrhea (%)	Relative % change from preceding year*
2002	7.17	—
2003	10.57	+47
2005†	6.30	-40
2006	6.80	+8

* Percent (%) change = (Year 2 – Year 1)/(Year 1).

† 2005 was compared with 2003, because no data from 2004 was available.

were collected from December 1991 through March 1992. The remaining 472 households were interviewed from April through July 1992. We restricted this analysis to the 5,852 households interviewed from December 1991 through March 1992 who had answered the question whether their youngest child who was < 5 years of age and lived in the household had diarrhea in the preceding 2 weeks. In the 2006/7 Pakistan DHS all but four households were interviewed from September 2006 through February 2007. We restricted the analysis to the 8,391 households interviewed from September 2006 through February 2007 who had answered the question whether their youngest child who is < 5 years of age and lived in the household had diarrhea in the preceding 2 weeks.

The differences reported in diarrhea prevalence varied by 25% per month in the 1991/2 survey and by 14% per month in the 2006/7 survey. The changes in the proportion of mothers who were uneducated, the age of the assessed child, and the wealth ranking of the households varied much less month to month (Table 5). The magnitude of the coefficient of variation of these diarrhea measurements was similar in the Demographic and Health Survey and in the longitudinal measurements in Karachi (Table 4). The expected sampling error only accounted for 27% of the observed standard deviation in the Demographic and Health Survey data.

DISCUSSION

In squatter settlements in central Karachi where the same children living in the same households were repeatedly assessed, the prevalence of diarrhea varied remarkably. The prevalence of diarrhea varied on average by 29% from one week to the next, by 37% from one month to the next, and by 32% from one year to the next. The coefficients of variation ranged from 158% to 372% confirming measures as highly variable.

Sampling error accounted for 33–46% of the weekly and monthly variability suggesting that the majority of the variability did not arise from sampling error, but from other factors causing episodic diarrhea. Age, the largest predictor of diarrhea in these data,^{13–15} varied only trivially between weeks, when there were large changes in diarrhea prevalence. When the analysis was stratified by age there was no reduction in the magnitude of variability. Households were followed longitudinally, therefore household level infrastructure and characteristics were quite stable during the repeated assessments, and would not explain the variability. In a child-based analysis, that combined all of the measurements for a single child during the study, the expected variability within each child accounted for just less than half of the observed variability. This suggests that this variability, which exceeded expected sampling variability, did not result from inter-child

TABLE 3
Observed vs. expected standard deviations of diarrhea prevalence by study

Study	Number of observations	Mean diarrhea prevalence	Observed standard deviation	Expected standard deviation	% of observed standard deviation expected as sampling error	Coefficient of variation (%)
Weekly						
Soap health	20372	0.058	0.195	0.089	45	334
Floc health	17764	0.079	0.223	0.102	46	281
Long-term practices	14932	0.049	0.183	0.082	45	372
Monthly*						
Soap health	4217	0.060	0.122	0.043	35	202
Floc health	3752	0.081	0.129	0.049	38	158
Long-term practices	3130	0.050	0.118	0.040	33	237
Pakistan DHS†						
1991/2	5422	0.134	0.341	0.091	27	254
2006/7	8391	0.224	0.417	0.111	27	186

* Child months with fewer than 4 weeks of observation were excluded from this analysis.
† Demographic and Health Survey.

differences, but from high levels of variability within measurements from the same child.

Taken together, these results suggest the high variability of diarrhea was not caused by sampling variation, household characteristics, age, or other interpersonal differences, but rather suggests that this variability resulted from determinants that were highly time dependent. A hypothesis that would explain this pattern of variability is that most episodes of child diarrhea result from ingestion of a sufficient dose of a gastrointestinal pathogen that the child does not have immunity against, and that a child's exposure to the myriad immunologically distinct variants of gastrointestinal pathogens is highly variable over time. Serial measurements of water contamination with fecal organisms suggest highly variable exposure over time.¹⁹ Person-to-person transmission of specific diarrheal pathogens is inherently episodic. We presume that new pathogens are introduced into a new community because of episodic connections with another infected person or community. When a new pathogen is introduced into a previously uninfected community, there are a large number of susceptible individuals and the disease will rapidly infect and be excreted by susceptible persons who in turn will excrete and further contaminate the environment, and so transmit the infection until a high level of population immunity develops to this particular pathogen and transmission is no longer sustained. As new children are born into the population who are immunologically naive to this pathogen and as immunity wanes with the passage of time, eventually the community is again susceptible to re-introduction to this pathogen or a closely related pathogen, but such a re-introduction depends on episodic connections to other infected communities. The weekly variation in diarrhea and the high level of variability in neighboring communities

likely reflects the different phases of the outbreaks of the various circulating pathogens. Weekly, monthly, and yearly differences in diarrhea prevalence and the high levels likely result from the complex interaction of exposure to gastrointestinal pathogens, the effectiveness and consistency of infrastructure, practices that effectively prevent mixing feces with food and water, and the population level of immunity to specific pathogens.

Karachi has a desert climate periodically punctuated by heavy rains. It is possible that the high variability of childhood diarrhea prevalence in the squatter settlements of Karachi is exceptional, and that in most other communities with high child mortality diarrhea prevalence is more stable. However, the 1991/2 and 2006/7 Pakistan Demographic and Health Survey data had similar magnitudes of coefficient of variations compared with the Karachi longitudinal data. The month-to-month variability was somewhat less than in the Karachi data, but an even smaller portion of this variability was explained by sampling error. This suggests that the high level of variability observed in the Karachi longitudinal evaluation is not exceptional, but rather reflects the inherent episodic pattern of diarrhea in a population.

Forsberg and colleagues²⁰ compared diarrhea prevalence measured in cross-sectional surveys at two time points separated by 1 to 3 years in the same country. Applying the same definition of relative change in the magnitude of diarrhea

TABLE 4
Diarrhea variability over time controlling for age

Age group	Mean relative change from preceding time frame number of person-weeks (% change)*		
	Weekly†	Monthly‡	Annually§
< 2 yr	16,550 (38)	16,550 (36)	7221 (40)
2 - < 5 yr	36,518 (38)	36,518 (45)	15,933 (39)
< 5 yr	53,068‡ (29)	53,068 (37)	23,154 (32)

* Percent (%) change = |(Timeframe 2 - Timeframe 1)/(Timeframe 2)|.
† Only weeks with ≥ 30 observations included.
‡ The 53,068 total observations exceed the sum of the age-specific weekly observations (52,266), because 30 person-weeks of observation that were dropped among children < 1 year of age; there were fewer than 30 observations within that age category and were included when the ages were pooled.
§ Restricted to high season July through October.

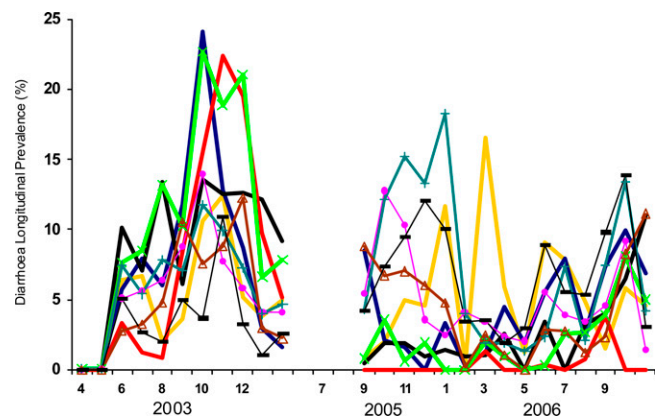


FIGURE 5. Mean monthly longitudinal prevalence of diarrhea among children < 5 years of age in 10 squatter settlement neighborhoods, Karachi, Pakistan.

TABLE 5
Differences in values by month of data collection Pakistan Demographic and Health Survey Data 1991/2 and 2006/7

Month	Households*	Diarrhea prevalence last 2 weeks		Uneducated mothers		Child age in months		Wealth quintile	
		1991-1992	%	% Monthly change†	%	% Monthly change†	Mean	% Monthly change†	Mean rank
Dec	1145	18.4		69		28.6		-‡	-
Jan	1519	12.8	44	75	9	28.1	2	-	-
Feb	1365	10.9	17	74	1	27.5	2	-	-
Mar	1393	12.6	13	80	8	28.3	3	-	-
Mean			25		6		2		
2006-2007									
Sep	1223	29.5		59		28.6		3.2	
Oct	1214	23.9	24	63	5	29.4	3	3.0	5
Nov	1687	24.1	1	68	8	28.6	3	2.8	7
Dec	1786	21.2	14	68	0	29.4	3	2.8	1
Jan	1673	18.5	14	72	5	29.2	1	2.6	8
Feb	808	16.1	15	74	3	28.6	2	2.7	4
Mean			14		4		2		5

*The number of households with mothers education and child months were somewhat less than the reports of diarrhea prevalence in the 1991/2 data because of missing values.

†Percent (%) change = |(Timeframe 2 - Timeframe 1)/[(Timeframe 2 + Timeframe 1)/2]|.

‡Asset information that would permit construction of a wealth index was not collected in the 1991/2 survey.

prevalence used in our Karachi analysis to the data from nine countries reported by Foster and colleagues, the prevalence of diarrhea in the preceding 2 weeks, ranged from a 40% decrease to a 43% increase in prevalence (Table 6). Forsberg and colleagues²⁰ conclude that these "differences in results . . . cannot only be explained by a variation in the true underlying values." The Karachi data suggest, however, that differences of this magnitude could readily be explained by underlying time-dependent episodic variation in diarrhea prevalence. Additionally, Forsberg and colleagues' findings further suggest that the variability in diarrhea prevalence noted in Karachi is not exceptional. Indeed, in other settings that have conducted longitudinal prevalence of diarrhea marked differences in diarrhea prevalence in different years has been reported.^{11,21,22}

Water, sanitation, and hygiene interventions have long been recognized as difficult to evaluate.^{12,23} This study illustrates that part of the difficulty in evaluating these interventions is that diarrhea prevalence is highly variable. Even after implementing a remarkably effective intervention, diarrhea prevalence might increase markedly because of normal variation from year to year. This is particularly difficult when interventions are focused on one geographic area, are relatively small, and do not enroll a simultaneous non-intervention control community for comparison. The common practice of evaluat-

ing an intervention designed to prevent diarrhea by measuring diarrhea before intervening, and after intervening,²⁴⁻²⁷ risks attributing a change in diarrhea prevalence to the intervention when the change actually results from underlying variation.

The high variability of diarrhea also limits the validity of conclusions from limited observations. Although global childhood mortality from diarrheal disease has decreased markedly in the last 3 decades, no change in diarrhea morbidity has been recognized.^{28,29} However, if diarrhea prevalence is highly variable, then meaningful trends are difficult to discern, particularly when few data points are available. To assess global trends in diarrheal morbidity, Kosek collected a mere six data points from the 1990s, only one of which was after 1994.²⁹ A genuine decrease in diarrheal prevalence of 10-20% over a decade would be difficult to separate from the wide year-to-year fluctuations that characterize diarrhea prevalence.

This study has important limitations. An inherent limitation in longitudinal measurements of phenomena that vary by age and time is that the population was different during the different observations. If the same population cohort were followed, diarrhea prevalence would decrease over time because as children age they acquire immunity and their incidence of diarrhea decreases. To evaluate children of similar ages in subsequent years, new children generally from new households need to be enrolled. Therefore, the population for this analysis was different each year. This change in the underlying study population was exacerbated by enrollment of new neighborhoods in 2003. It is possible that an important part of the variability from year-to-year results from differences in the underlying population or differences in vulnerability to diarrhea in the different neighborhoods that were enrolled in the different studies. However, this is unlikely to explain most of the variability in these data for three reasons. First, the mean age of the children was similar in each study and these data come from similar neighborhoods located near each other with similar water, sanitary and solid waste infrastructure, and similar household characteristics. Second, although populations were somewhat different from year to year they were quite similar from month to month and from week to week, where high levels of variability were also noted. Third, in the child-based analysis, half of the observed variability within the

TABLE 6

Change in 2-week diarrhea prevalence in successive cross-sectional surveys (prevalence data from²⁰)

	2-week prevalence of diarrhea (%)		2-week prevalence of diarrhea (%)		Relative % change from preceding survey
	Year	Year	Year	Year	
Botswana	12	1986	11	1988	-9
Kenya-Coastal	10.1	1989	15.3	1990	34
Kenya - Western	17.1	1989	17.9	1990	4
Morocco	29	1987	27	1990	-7
Tunisia	21	1988	15	1988	-40
Northeast Brazil	21.0	1986	15.1	1987	-39
Dominican Republic	24.2	1986	23	1987	-5
Guatemala	20	1986	21.8	1987	8
Mexico	13	1986	23	1987	43

same child was not accounted for by the expected sampling variability.

A second limitation is that the calculation of expected standard deviation assumed a binomial distribution, but the underlying assessments of diarrhea prevalence were repeated measures of the same child. Because a child who has diarrhea on one day is more likely to have diarrhea on the next day, these events are not truly independent and a binomial distribution will underestimate the sampling error, and therefore overestimate the residual unexplained variability. However, because the basic data collection time frame for this analysis was weekly diarrhea prevalence, the week-to-week dependency would not be as strong as the daily measures. There would be some dependence for episodes of diarrhea from the same pathogen that occurred toward the end of one week of assessment and continuing into the next week, but this would occur in a much smaller proportion of episodes comparing diarrhea prevalence month to month or year to year. However, the data suggest high degrees of variability across all time frames with the highest proportion of non-sample variability in the monthly assessments.

A third limitation is that our approach to assessing variability was crude. Variance component models or Fourier decomposition would produce a more sophisticated model of the variability, but the more direct analysis presented is easier to understand, accounts for sampling and other primary causes of variability, and communicates the underlying features of the data.

Taken together, these data suggest that diarrhea prevalence in a community, city, or country should not be conceptualized as a single number that can be reliably measured in a cross-sectional survey. It is better to think of childhood diarrhea prevalence as the result of a complex system, something like rainfall, a phenomenon that is highly variable from place to place, from week to week, from month to month, and from year to year. Valid assessments of the effectiveness of interventions to reduce diarrhea or trends over time need to account for the high time and location dependent variability of diarrhea measurements. Simultaneous longitudinal surveillance in representative intervention communities and comparable non-intervention communities are best suited to assess intervention effects independent of this variability.

Received June 26, 2010. Accepted for publication January 25, 2011.

Acknowledgments: We appreciate the contribution of the field workers and study participants to this analysis. Faisal Sarwari assisted with data management.

Financial support: This work was supported by the Procter & Gamble Company and the U.S. Centers for Disease Control and Prevention.

Authors' addresses: Stephen Luby, Centre for Communicable Diseases, ICDDR,B: International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh, and Global Disease Detection and Emergency Response Division, Centers for Disease Control and Prevention, Atlanta, GA, E-mail: sluby@icddr.org. Mubina Agboatwalla, Health Oriented Preventive Education, Karachi, Pakistan, E-mail: agboat@hope-ngo.com. Robert M. Hoekstra, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, E-mail: rth6@cdc.gov.

REFERENCES

- Boschi-Pinto C, Velebit L, Shibuya K, 2008. Estimating child mortality due to diarrhoea in developing countries. *Bull World Health Organ* 86: 710–717.
- Checkley W, Buckley G, Gilman RH, Assis AM, Guerrant RL, Morris SS, Molbak K, Valentiner-Branth P, Lanata CF, Black RE, 2008. Multi-country analysis of the effects of diarrhoea on childhood stunting. *Int J Epidemiol* 37: 816–830.
- Lorntz B, Soares AM, Moore SR, Pinkerton R, Gansneder B, Bovbjerg VE, Guyatt H, Lima AM, Guerrant RL, 2006. Early childhood diarrhea predicts impaired school performance. *Pediatr Infect Dis J* 25: 513–520.
- Plate DK, Strassmann BI, Wilson ML, 2004. Water sources are associated with childhood diarrhoea prevalence in rural east-central Mali. *Trop Med Int Health* 9: 416–425.
- Pande S, Keyzer MA, Arouna A, Sonneveld BG, 2008. Addressing diarrhoea prevalence in the West African Middle Belt: social and geographic dimensions in a case study for Benin. *Int J Health Geogr* 7: 17.
- Ryland SR, 1998. Childhood morbidity and treatment patterns. *Demographic and Health Surveys Comparative Studies*. Calverton, MD: Macro International Inc.
- Esrey SA, 1996. Water, waste, and well-being: a multicountry study. *Am J Epidemiol* 143: 608–623.
- Robins-Browne RM, 1984. Seasonal and racial incidence of infantile gastroenteritis in South Africa. *Am J Epidemiol* 119: 350–355.
- Mahmud A, Jalil F, Karlberg J, Lindblad BS, 1993. Early child health in Lahore, Pakistan: VII. Diarrhoea. *Acta Paediatr Suppl* 82 (Suppl 390): 79–85.
- Velazquez FR, Garcia-Lozano H, Rodriguez E, Cervantes Y, Gomez A, Melo M, Anaya L, Ovalle JC, Torres J, Diaz De Jesus B, Alvarez-Lucas C, Breuer T, Munoz O, Kuri P, 2004. Diarrhoea morbidity and mortality in Mexican children: impact of rotavirus disease. *Pediatr Infect Dis J* 23: S149–S155.
- Huttly SR, Hoque BA, Aziz KM, Hasan KZ, Patwary MY, Rahaman MM, Feachem RG, 1989. Persistent diarrhoea in a rural area of Bangladesh: a community-based longitudinal study. *Int J Epidemiol* 18: 964–969.
- Blum D, Feachem RG, 1983. Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: problems of methodology. *Int J Epidemiol* 12: 357–365.
- Luby SP, Agboatwalla M, Painter J, Altaf A, Billhimer WL, Hoekstra RM, 2004. Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan: a randomized controlled trial. *JAMA* 291: 2547–2554.
- Luby SP, Agboatwalla M, Painter J, Altaf A, Billhimer W, Keswick B, Hoekstra RM, 2006. Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomized controlled trial. *Trop Med Int Health* 11: 479–489.
- Luby SP, Agboatwalla M, Bowen A, Kenah E, Sharker Y, Hoekstra RM, 2009. Difficulties in maintaining improved handwashing behavior, Karachi, Pakistan. *Am J Trop Med Hyg* 81: 140–145.
- Planning and Development Corporation and Pakistan Environmental Planning and Architecture Consultant, 1990. *Karachi Development Plan 2000*. Karachi, Pakistan: Karachi Developmental Association.
- Morris SS, Cousens SN, Kirkwood BR, Arthur P, Ross DA, 1996. Is prevalence of diarrhea a better predictor of subsequent mortality and weight gain than diarrhea incidence? *Am J Epidemiol* 144: 582–588.
- Fleiss JL, Levin B, Paik MC, 2003. *Statistical Methods for Rates and Proportions*. New York: Wiley-Interscience.
- Levy K, Hubbard AE, Nelson KL, Eisenberg JN, 2009. Drivers of water quality variability in northern coastal Ecuador. *Environ Sci Technol* 43: 1788–1797.
- Forsberg BC, van Ginneken JK, Nagelkerke NJ, 1993. Cross-sectional household surveys of diarrhoeal diseases—a comparison of data from the Control of Diarrhoeal Diseases and Demographic and Health Surveys programmes. *Int J Epidemiol* 22: 1137–1145.
- Pinfold JV, Horan NJ, Mara DD, 1991. Seasonal effects on the reported incidence of acute diarrhoeal disease in northeast Thailand. *Int J Epidemiol* 20: 777–786.
- Hashizume M, Armstrong B, Hajat S, Wagatsuma Y, Faruque AS, Hayashi T, Sack DA, 2007. Association between climate variability and hospital visits for non-cholera diarrhoea in Bangladesh: effects and vulnerable groups. *Int J Epidemiol* 36: 1030–1037.

23. Briscoe J, 1978. The role of water supply in improving health in poor countries (with special reference to Bangla Desh). *Am J Clin Nutr* 31: 2100–2113.
24. Sheth M, Obrah M, 2004. Diarrhea prevention through food safety education. *Indian J Pediatr* 71: 879–882.
25. Elsanousi S, Abdelrahman S, Elshiekh I, Elhadi M, Mohamadani A, Habour A, Elamin SE, Elnoury A, Ahmed EA, Hunter PR, 2009. A study of the use and impacts of LifeStraw in a settlement camp in southern Gezira, Sudan. *J Water Health* 7: 478–483.
26. Kolesar R, Kleinau E, Torres MP, Gill C, Cruz V, Post M, Johnson E, Cordero F, Gavin J, 2004. *Combining Hygiene Behavior Change with Water and Sanitation: Monitoring Progress in Hato Mayor, Dominican Republic*. Washington, DC: Environmental Health Project, 38.
27. Moll DM, McElroy RH, Sabogal R, Corrales LF, Gelting RJ, 2007. Health impact of water and sanitation infrastructure reconstruction programmes in eight Central American communities affected by Hurricane Mitch. *J Water Health* 5: 51–65.
28. Keusch GT, Fontaine O, Bhargava A, Boschi-Pinto C, Bhutta ZA, Gotuzzo E, Rivera JA, Chow J, Shahid-Salles SA, Laxminarayan R, 2006. Diarrheal diseases. *Disease Control Priorities in Developing Countries*. New York: Oxford University Press, 371–388.
29. Kosek M, Bern C, Guerrant RL, 2003. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull World Health Organ* 81: 197–204.