Is Prevalence of Diarrhea a Better Predictor of Subsequent Mortality and Weight Gain Than Diarrhea Incidence?

Saul S. Morris, 1 Simon N. Cousens, 1 Betty R. Kirkwood, 1 Paul Arthur, 1, 2 and David A. Ross 1

A number of different outcome measures have been proposed for use in prospective studies of morbidity associated with childhood diarrhea. These include the number of episodes experienced by each child over a defined period (a measure of incidence) and the number of days of diarrhea divided by the total number of days of observation for each child (a measure denoted "longitudinal prevalence"). The authors examined data from Ghana to determine which of these measures is more strongly associated with weight gain over a 4-month period and subsequent mortality. Both diarrhea incidence and longitudinal prevalence were associated with weight gain in children aged 6-23 months, but a statistically stronger association was observed with longitudinal prevalence (likelihood ratio statistic 28.95 on 1 degree of freedom against 19.70 for incidence). Neither measure was associated with weight gain in younger or older children. Longitudinal prevalence, but not incidence, was strongly associated with subsequent mortality (p = 0.002 for longitudinal prevalence; p = 0.557 for incidence). Although many epidemiologic studies of diarrhea focus on incidence, these data suggest that longitudinal prevalence is more strongly predictive of long-term health outcome. The authors conclude that longitudinal prevalence merits greater attention as a measure of outcome in diarrhea studies. Am J Epidemiol 1996;144:582-8.

Received for publication September 13, 1995, and accepted for publication March 8, 1996.

1 Maternal and Child Epidemiology Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom.
2 Health Research Unit, Ministry of Health, Accra, Ghana.

Diarrheal diseases continue to be the second most common cause of death of children younger than 5 years in demographically developing countries (1). Considerable efforts have been devoted to establishing the magnitude of the problem of diarrhea morbidity at a global level (2) and to identifying potential interventions for disease control (3, 4). In contrast, little consideration has been given to the question of how the burden of diarrheal disease on young children should most adequately be described. The following four major categories of outcome measures have been suggested for use in studies of diarrheal disease (5): 1) the number of episodes experienced by each child; 2) the average duration of each episode experienced by the child; 3) the days of illness suffered by each child, as a percentage of all days of observation; and 4) the severity of each episode. Measures 1-3 all attempt to quantify the amount of illness experienced by each child. Although the first is a measure of disease incidence, it differs from the classic measure of incidence used in chronic disease epidemiology (6) because the same children may experience the disease outcome more than once. The second measure will effectively distinguish between children with a high and low burden of disease only when, for each child, duration tends to be consistently long or short from one episode to the next. However, when the same children are observed to have both short and long episodes, their average duration of illness will be difficult to interpret. The third measure is familiar from studies of the impact of diarrheal illness on growth (7-12) and is often referred to simply as "prevalence." However, it differs from the usual measure of prevalence in that it measures the burden of disease suffered by particular individuals over time, whereas "point prevalence" is a cross-sectional description of the burden of disease on a population. Kleinbaum et al. (13) have named the proportion-of-time-ill measure "episodic prevalence." We consider this term misleading, inasmuch as it suggests that it incorporates information about numbers of episodes, which is not the case. We therefore prefer the term longitudinal prevalence.

In this paper, we investigate whether the marked preference for incidence as a measure of disease burden may in some circumstances be unwarranted in the case of childhood diarrhea. Because diarrhea is known to provoke decreased food intake, malabsorption, loss of endogenous nutrients, and increased metabolic rate (4), it is plausible that the long-term impact on chil-
Children’s health may be determined to a greater degree by the total number of days spent ill than by the number of new episodes. We therefore set out to compare the relative explanatory power of a measure of diarrheal incidence and a measure of longitudinal prevalence in predicting subsequent health outcomes in a cohort of Ghanaian children less than 5 years of age. Specifically, we seek to examine the associations both with weight gain over a 4-month interval and with mortality in the 4-month period after morbidity was measured.

**MATERIALS AND METHODS**

This paper uses data collected during the course of a trial of the impact of periodic, prophylactic supplementation with large doses of vitamin A on the health and growth of young children in northern Ghana (Ghana VAST Child Health Study). Both children randomized to receive vitamin A and those randomized to receive placebo, as well as those too young to have been dosed, are included in this analysis. The background of the study and a description of the methods used in the data collection have been described elsewhere (14). In brief, morbidity surveillance involving weekly home visits over a 1-year period beginning in June 1990 was set up for a dynamic cohort of 1,872 children aged 0–5 years. In addition to the home visits, clinical examinations with anthropometry were conducted at the beginning of the study and every 4 months thereafter. The intervals between these clinical examinations will be referred to as rounds 1, 2, and 3, respectively. In the area where the study was conducted, acute gastroenteritis and chronic diarrhea/malnutrition together accounted for between one quarter and one third of all deaths of infants and young children (14).

**Incidence and longitudinal prevalence of diarrhea** were calculated from a day-by-day morbidity record obtained at weekly visits to the children’s homes. At these visits, inquiries were made about seven different locally defined terms for diarrheal illnesses, based on the findings of a focused ethnographic study of local concepts relating to illness in young children conducted before the beginning of the study. Using a pictorial illness diary kept by the children’s mothers or guardians to aid memory, morbidity status was established for each day of the week preceding the interview. Subsequently, a single variable was created combining all seven different diarrhea terms. Epidemiologic measures of diarrhea burden were calculated separately for each child in each round. Longitudinal prevalence was defined as the number of days with diarrhea divided by the total number of days of observation. Episodes of diarrheal illness were defined using a 2-day symptom-free gap to mark the beginning of a new episode (15). Diarrhea incidence was standardized to a 4-month (119-day) time period by dividing each child’s observed number of incident episodes over the round by the total number of days of observation in that round and then multiplying the result by 119. Rounds in which the child was observed for a total of less than 7 weeks were excluded from the analysis because a 50-day period of observation was considered adequate to reliably estimate measures of illness burden. This criterion resulted in the exclusion of data from 400 of 4,842 (8.2 percent) child-rounds, corresponding to just 1.9 percent of all days of observation in the study.

The analysis of weight gain is based on measures made at the beginning of the study and at the three subsequent clinical examinations. Because enrollment into the study was continuous, each child was weighed between one and four times. On each occasion, weight was measured twice to the nearest 100 g by two independent teams of trained field workers using 25-kg Salter hanging scales (CMS Weighing Equipment, London, United Kingdom). An independent arbitrator determined whether the measurements coincided to within 300 g; and if they did, the second measurement was recorded. If they did not, the child was returned to the first anthropometry team for reweighing, and so on, until agreement was reached. Only those child-rounds with anthropometric measurements at both the beginning and end of the interval \( n = 3,641 \) were included in this analysis.

**Mortality** was recorded by the morbidity surveillance system. A total of 77 deaths were identified. However, for the half (38) who died before the start of the second 4-month round, no information was available on morbidity over the previous round. Of the remaining 39, the analysis was restricted to the 20 deaths in children who had at least 7 weeks of morbidity information over the previous round. Four “controls” were matched to each death, selected from those children alive at the time the “cases” died. The controls were chosen so that they were all the same age (in months) as the dead children, had at least 7 weeks of morbidity information in the previous round, and were, of all such eligible children, those living closest to their matched cases as determined by the sequential numbering of the residential compounds in the field. Their morbidity over the round preceding the deaths of their respective cases was compared with that of the cases.

The association between weight gain and diarrheal morbidity was assessed using “multilevel modeling” (16) to account for possible within-subject correlation in weight gain from one round to the next. Indicator
variables were included to model the effect of round of observation. The model thus took the following form:

\[
\text{weight gain} = \alpha_1 + \beta \text{diarrhea} + \gamma \text{round 2} + \delta \text{round 3}
\]

The intercept terms, \(\alpha_i\), were allowed to vary randomly at the level of the individual \((i)\). Because the effect of diarrheal morbidity on weight gain has previously been shown to be dependent on the age of the child \((11)\), the model was fitted separately for three different age groups: 0–5 months, 6–23 months, and 24–59 months. After examining the associations between weight gain and each of diarrheal incidence and longitudinal prevalence separately, secondary analyses were performed to determine whether adding a second measure of diarrhea frequency to a model already containing the other measure would result in a significant improvement in goodness of fit. Such an improvement would indicate that the second measure contributed additional information significantly associated with weight gain even after taking into account the value of the first measure. The analyses were conducted using the computer package MLn (Institute of Education, London, United Kingdom).

The association between diarrheal morbidity and subsequent mortality was assessed using conditional logistic regression \((17)\) and the computer package EGRET (SERC, Seattle, Washington). For both the weight gain outcome and the mortality outcome, the statistical significance of the observed associations was determined by examining the likelihood ratio statistic \((17)\).

**RESULTS**

**Association between diarrhea incidence and longitudinal prevalence**

In figure 1, the distribution of the number of incident episodes of diarrhea per child per round (top panel) and the distribution of longitudinal prevalences of diarrhea for the same children and rounds (bottom panel) are shown. No new episodes of diarrhea were observed in 991 child-rounds (nearly one fourth of the total), whereas a small number of child-rounds were characterized by almost uninterrupted diarrheal illness. Examination of the association between the two measures of diarrhea burden revealed that, broadly speaking, longitudinal prevalence increased as the number of episodes increased (figure 2). However, the two variables are clearly not measuring exactly the same phenomenon inasmuch as some children experienced only one or two incident episodes of diarrhea over a round but were almost constantly sick, whereas others experienced eight or more incident episodes and yet were sick for only a small proportion of the round. Furthermore, a few children did not experience any incident episode of diarrhea over the round but nonetheless experienced moderate to high longitudinal prevalences of diarrhea as a result of the continuation of prolonged episodes beginning in the previous round. In total, just 34 percent of the variation in longitudinal prevalence was explained by the number of incident episodes \(r^2\).

**Association between diarrhea burden and failure to thrive**

The associations between weight gain over a 4-month period and each of diarrhea incidence and longitudinal prevalence of diarrhea were evaluated separately for children aged 0–5 months, 6–23 months, and ≥24 months at the start of the round (table 1). Neither longitudinal prevalence nor incidence was associated with weight gain in the first 6 months of life, nor in children aged 24 months or older. In contrast, both measures of burden of diarrhea were strongly associated with weight gain in children aged 6–23 months \((p < 0.001\text{ in each case})\). In this age group, the negative association between diarrheal illness and weight gain appeared stronger for longitudinal prevalence than for incidence, both in terms of the estimated difference in weight gain between a child at the 75th percentile of the illness distribution and a child at the 25th percentile of the distribution (which was −100 g for longitudinal prevalence vs. −78 g for incidence) and in terms of the likelihood ratio statistic. After adjusting for the number of incident episodes, children’s longitudinal prevalence of diarrhea remained strongly negatively associated with weight gain (adjusted effect of being at the 75th percentile of the illness distribution compared with those at the 25th percentile = −78 g; likelihood ratio statistic = 13.33 on 1 degree of freedom; \(p = 0.001\)). After adjusting for longitudinal prevalence, however, the strength of the negative association between incidence and weight gain was considerably reduced (adjusted effect = −41 g; likelihood ratio statistic = 4.08 on 1 degree of freedom; \(p = 0.043\)). There was no suggestion of any nonlinearity in the association between weight gain and longitudinal prevalence of diarrhea, or in the association between weight gain and diarrhea incidence.

**Association between diarrhea burden and subsequent mortality**

Longitudinal prevalence of diarrhea was strongly associated with the risk of mortality in the subsequent
round (table 2). On average, a 5 percent absolute increase in longitudinal prevalence of diarrhea was associated with a 17 percent relative increase in the risk of mortality in the following round (95 percent confidence interval 5–30 percent). This association was highly statistically significant ($p = 0.002$). Because of the matched design and analysis of this nested case-control study, the association should be free from confounding by age, season, or geographic area. In contrast, there was no evidence of an association between diarrhea incidence and subsequent mortality ($p = 0.557$).

We investigated whether supplementation with vitamin A either confounded the association between diarrhea (longitudinal prevalence and incidence) and outcome (weight gain, mortality) or modified the effect of diarrhea. Inclusion of the variable indicating supplementation with vitamin A or placebo in the

**FIGURE 1.** Distribution of two alternative measures of the burden of diarrhea in children with at least 7 weeks of observation time: Ghana, 1990-1991.
models did not lead to any appreciable changes in the estimated impacts of diarrhea, and the interaction terms in all cases were nonsignificant ($p > 0.2$).

**DISCUSSION**

Our study has shown that although longitudinal prevalence of diarrhea and incidence of diarrhea were, as expected, correlated at the individual level in these Ghanaian children, the correlation between the two measures was far from perfect. Some children with high numbers of incident episodes over a 4-month period had relatively low longitudinal prevalences, whereas other children with just one or two incident episodes experienced prolonged periods of time with diarrhea.

We found that diarrhea burden, however measured, was strongly associated with weight gain in children 6–23 months of age. This finding is consistent with the results of a previous study conducted in The Gambia (11), where it was found that each extra day of diarrhea was, on average, associated with a 3.7-g shortfall in weight gain in children 0–23 months of age after adjusting for age and season. Reanalyzing our data with a combined age group of 0–23 months suggests that each extra day of diarrhea was associated with a smaller 1.7-g shortfall in weight in the Ghanaian children aged 0–23 months. In the Ghana study, the youngest age groups were continually replenished with new births. This difference from the Gambia study, which was a birth cohort, might have led to a lower estimate of the effect of diarrheal morbidity on growth inasmuch as no association was detected in the 0- to 5-month age group in our study.

Although both of the two measures of diarrhea burden that we studied showed an association with weight gain in the 6–23 months age group, comparison of the likelihood ratio statistics from the two models showed that longitudinal prevalence of diarrhea appeared to have more explanatory power with regard to weight gain than did incidence. In particular, the strength of the association between the number of incident episodes of diarrhea and weight gain was markedly reduced after adjusting for the percentage of days with diarrhea (longitudinal prevalence). However, longitudinal prevalence remained very strongly associated with weight gain even after adjusting for incidence.

Longitudinal prevalence of diarrhea was also highly predictive of mortality in the subsequent 4-month period, with a 5 percent absolute increase in longitudinal prevalence of diarrhea associated with a 17 percent relative increase in the risk of mortality. There was no evidence from this study that diarrhea incidence was associated with subsequent mortality.

*Am J Epidemiol* Vol. 144, No. 6, 1996
TABLE 1. Association of longitudinal prevalence of diarrhea and diarrhea incidence with weight gain, Ghana, 1990–1991

<table>
<thead>
<tr>
<th>Age group (months)</th>
<th>No. of child-rounds</th>
<th>Mean weight gain (g)</th>
<th>Median (%</th>
<th>Interquartile range</th>
<th>Effect of diarrhea on weight gain (g)</th>
<th>95% CI</th>
<th>Likelihood ratio statistic</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–5</td>
<td>423</td>
<td>1,412</td>
<td>21.0</td>
<td>9.2–37.3</td>
<td>−24</td>
<td>−119 to +71</td>
<td>0.25</td>
<td>1</td>
<td>0.616</td>
</tr>
<tr>
<td>6–23</td>
<td>1,343</td>
<td>781</td>
<td>17.5</td>
<td>7.1–33.6</td>
<td>−100</td>
<td>−136 to −64</td>
<td>28.95</td>
<td>1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥24</td>
<td>1,875</td>
<td>574</td>
<td>4.2</td>
<td>0.0–12.6</td>
<td>−20</td>
<td>−44 to +4</td>
<td>2.59</td>
<td>1</td>
<td>0.108</td>
</tr>
</tbody>
</table>

* Expected difference in weight gain between children at the 75th percentile of the distribution of diarrhea prevalence/incidence and children at the 25th percentile.
† CI, confidence interval.


<table>
<thead>
<tr>
<th>Mean burden of diarrhea</th>
<th>Longitudinal prevalence (%)</th>
<th>OR*</th>
<th>95% Cl</th>
<th>Likelihood ratio statistic</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases (n = 20)</td>
<td>Control (n = 80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal prevalence</td>
<td>35.8</td>
<td>16.5</td>
<td>1.17†</td>
<td>1.05–1.30</td>
<td>9.86</td>
<td>1</td>
</tr>
<tr>
<td>Incidence</td>
<td>2.47</td>
<td>2.83</td>
<td>0.94</td>
<td>0.76–1.16</td>
<td>0.345</td>
<td>1</td>
</tr>
</tbody>
</table>

* Incident episodes of diarrhea per child-round.
† OR, odds ratio; CI, confidence interval.
‡ Estimated impact of a 5% absolute increase in longitudinal prevalence.

The mortality analysis was limited to only 20 of the 77 deaths that occurred during the study period, in part because for the 38 children who died before the start of the second round, no information was available on prior morbidity. These children died at a different time of year from those analyzed; and should there have been seasonal variation in the relative importance of acute and persistent diarrhea as contributing causes of mortality in this population, it is conceivable that diarrhea incidence could have been more (or less) strongly associated with mortality in these children than it was for those included in the analysis. An additional 19 deaths were not included because they had less than 50 days of morbidity data from the previous round. Although our choice of the 50-day cutoff was to some degree arbitrary, it is important to recognize that neither incidence nor longitudinal prevalence can be reliably estimated in individuals with very short periods of follow-up. The precise length of follow-up required for adequate estimation of these measures requires additional investigation. The inclusion criterion of 50 days of follow-up had little impact on the analysis of weight gain because this analysis was dependent on the availability of anthropometric measurements at the beginning and end of the 4-month interval, and very few children were present at the beginning and end of a round but not present for at least 7 weeks in the interim.

We believe that the differential ability of the two measures of the burden of diarrhea to predict poor health outcomes probably reflects a true difference in their clinical significance. In spite of its relatively low incidence in the community (18, 19), persistent diarrhea is known to account for a large proportion of diarrhea-related deaths (20, 21). Children suffering from persistent diarrhea could easily be classified in longitudinal field studies as "low incidence" if they experience a single, uninterrupted period of illness (in this study, 31 of 60 child-rounds characterized by longitudinal prevalences of 275 percent involved less than three incident episodes). However, five "episodes" of diarrhea each lasting just 1 or 2 days, for example, might have little long-term effect on most children's health. Longitudinal prevalence would more accurately capture these children's true risk status than measures of incidence.
At present, incidence is a much more commonly reported outcome measure than longitudinal prevalence in prospective studies of childhood diarrhea. Although this may be appropriate in studies of interventions expected to reduce transmission of diarrheal pathogens (such as water and sanitation projects), or those aimed specifically at reducing the occurrence of acute diarrhea (such as rotavirus vaccine trials), it is less appropriate in studies of interventions expected to improve host response to diarrhea (such as vitamin A supplementation or treatment of acute diarrhea with zinc) or in those studies seeking to quantify, in a more general way, the burden on children of morbidity from diarrhea. In such cases, longitudinal prevalence may be a more appropriate outcome measure because it better combines the negative effects of acute and persistent diarrhea and is, as we have shown, strongly related to long-term health outcomes such as weight gain and mortality. Furthermore, it is simple to calculate and free of the complexities that surround the definition of an episode (15). Indeed, longitudinal prevalence can even be estimated when the morbidity record is not continuous but is made up of periodic reevaluations of individuals’ health status (such as when the data are recorded for just 1 day each week). In these circumstances, incidence can only be estimated by relying on respondents’ recall of morbidity status on previous days, which may not always be reliable (22). Finally, because longitudinal prevalence is a measure that is child based rather than event based, difficulties of within-subject correlation of multiple outcomes (23) do not arise, and analyses of subgroup differentials are thus relatively straightforward. For all of these reasons, we believe that longitudinal prevalence deserves greater attention as a measure of outcome in longitudinal studies of childhood diarrhea.

ACKNOWLEDGMENTS

This work was supported by the Health and Population Division of the United Kingdom Overseas Development Administration.

REFERENCES