

**OUTCOME MEASURES IN PROSPECTIVE STUDIES OF  
CHILDHOOD DIARRHOEA AND RESPIRATORY INFECTIONS:  
CHOOSING AND USING THEM**

by  
Simon N. Cousens<sup>1</sup>  
Betty R. Kirkwood<sup>1</sup>

<sup>1</sup>MATERNAL AND CHILD EPIDEMIOLOGY UNIT  
DEPARTMENT OF EPIDEMIOLOGY AND POPULATION SCIENCES  
LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE  
KEPPEL STREET, LONDON WC1E 7HT

## CONTENTS

	<u>Page</u>
SUMMARY .....	2
1. INTRODUCTION .....	3
2. CHOICE OF OUTCOME MEASURE: SOME OPTIONS .....	3
3. SOME EXAMPLES OF RESEARCH QUESTIONS .....	4
4. STATISTICAL TECHNIQUES FOR THE ANALYSIS OF QUANTITATIVE OUTCOMES .....	6
4.1 Description of the data .....	6
4.2 Analysis of the number of episodes .....	7
4.2.1 Categorized data .....	7
4.2.2 Untransformed, continuous data .....	11
4.2.3 Transformed, continuous data .....	12
4.2.4 Non-parametric analysis .....	13
4.3 Analysis of the duration of episodes .....	14
4.3.1 Categorized data .....	15
4.3.2 Untransformed, continuous data .....	16
4.4 Analysis of the total number of days of morbidity .....	16
5. DISCUSSION .....	18
5.1 Choice of outcome measure .....	18
5.2 Statistical methods .....	19
5.3 Interpretation of associations .....	20
ACKNOWLEDGEMENTS .....	20
REFERENCES .....	21
ANNEX. Diarrhoea morbidity data for 600 children followed for 12 weeks .....	22

SUMMARY

Children in developing countries suffer repeated episodes of diarrhoea and respiratory morbidity. A number of outcome measures are available for use in prospective studies of these diseases: number of episodes, number of days of morbidity, average duration of each episode, severity of each episode. While the analysis of binary outcomes such as ill/not ill or survived/died is discussed extensively in the epidemiological literature, less attention has been paid to other types of outcome. In this paper we explain how the choice of an appropriate measure of outcome is governed by the research question posed, and we examine some of the basic statistical techniques available for analysing outcomes, such as number of episodes. A worked example is presented using data on diarrhoea morbidity.

## 1. INTRODUCTION

Diarrhoeal and acute respiratory diseases are both major causes of morbidity and mortality among children in developing countries. Each is responsible for several million child deaths annually (UNICEF, 1990). In addition, they are epidemiologically similar in a number of ways. Young children may suffer from repeated episodes of diarrhoea and from repeated respiratory infections. Both are clinical manifestations of infections due to a range of pathogens rather than a single disease entity. Both are diseases of poverty. These epidemiological characteristics have a number of implications for the design, analysis, and interpretation of studies of childhood diarrhoea and respiratory infections. When studying chronic diseases, which are relatively rare in most settings and normally occur only once, a person's health status can usually be recorded as either diseased or not diseased. When conducting follow-up studies of common, recurrent diseases, such as childhood diarrhoea and acute respiratory infections, however, a simple classification such as this is frequently not sufficient since during a follow-up period of two years, almost every child will suffer from at least one episode of the disease. Thus, other measures of the disease burden, referred to hereafter as outcome measures, are needed in the study of these diseases. In this paper we discuss the different choices of outcome measure that are available in follow-up studies of diarrhoeal and respiratory illness and show how the appropriate choice is dependent upon the research question posed. Some of the measures available are binary, qualitative variables (e.g., survived/died). Binary outcomes are widely used in epidemiological studies and have been much discussed in the literature (e.g., Kleinbaum et al., 1982; Rothman, 1986). Others of the measures available are of types less frequently used, and there is consequently less literature on their analysis. In this paper we examine some of the basic statistical techniques available for the analysis of non-binary outcomes.

## 2. CHOICE OF OUTCOME MEASURE: SOME OPTIONS

Four main categories of outcome measure are available for use in prospective studies of diarrhoeal and respiratory diseases:

- (i) the number of episodes of illness experienced by each child (0,1,2,...),
- (ii) the average duration of each episode experienced by the child (e.g., 3.2 days),
- (iii) the total number of days of illness suffered by each child (0,1,2,...),
- (iv) the severity of each episode.

This last category, severity, can be measured in a number of different ways. For example, it could be measured by the binary variable survived/died. Other examples are use of degree of dehydration (none/some/severe) as the outcome measure in a study of diarrhoea morbidity, or use of the presence/absence of chest indrawing as the outcome measure in a study of respiratory morbidity.

It should be noted that, while the first three of these measures (number of episodes, mean duration, number of days) are all quantitative (can be expressed as a number), the fourth (severity) is more qualitative. The nature of the outcome measure, quantitative or qualitative, has implications for the analysis of the data, a point to which we shall return.

Although the measures outlined above are closely related, they are not equivalent. For example, the number of days of illness experienced by a child is clearly closely related to the number of episodes suffered by that child. The two are, in fact, linked through the duration of each episode (number of days - number of episodes x average duration). There may, however, be factors which affect the duration of an episode but which do not influence the number of episodes suffered by a child, and vice versa. Each of the various outcome measures reveals something different about the disease under study and therefore one cannot say that one measure is "better" than the others. In some situations the number of episodes per child will be the most appropriate measure to use, whilst in others it will be the average duration of each episode, and in yet others some measure of the severity of the episode. Suppose, for example, that a child's nutritional (anthropometric) status affects the duration of respiratory infections (children with poor nutritional status suffering longer episodes), but does not affect the number of times the child suffers such an infection. Then, a well-designed study using average duration per episode or total days of illness as the outcome measure of interest would stand a good chance of detecting an association between nutritional status and respiratory morbidity. On the other hand, a study using number of episodes as its outcome measure could not expect to detect any association between nutritional status and respiratory disease.

In summary, there are a number of outcome measures that can be used in prospective studies of childhood diarrhoea and respiratory infections. In any particular situation, the measure that one chooses to use will be determined by the research question that one is trying to answer. In order to choose the most appropriate outcome measure it is therefore essential to understand and define clearly the research question(s) being posed.

### 3. SOME EXAMPLES OF RESEARCH QUESTIONS

To illustrate the way in which the research question to be answered determines the choice of outcome, and to provide examples of different questions that might be posed, consider the general question of whether or not a relationship exists between measles vaccination and respiratory morbidity and mortality. Some specific questions that might then be raised are discussed below.

#### Question 1

"Can measles vaccination reduce the incidence of respiratory infections in young children?"

In order to answer this question, one might follow a group of children, some vaccinated and others not, to see whether the children who were vaccinated suffered fewer episodes of respiratory morbidity than those who were not vaccinated. Clearly, in this case, the outcome measure of interest is the number of episodes suffered by each child. This measure is quantitative (a number: 0, 1, 2, ...) and the unit of observation is the child. All children contribute information to the study, regardless of whether or not they become ill while it is under way.

#### Question 2

"Can measles vaccination reduce the risk of mortality associated with an episode of respiratory morbidity?"

This question is concerned with the course of the illness once the child has become ill. To study it, one might follow a group of children, identifying episodes of respiratory morbidity as they occur, and then monitoring the outcome, recovered or died, of each episode. (For the purposes of this discussion, we ignore any possible ethical problems surrounding such a procedure.) This procedure differs in two ways from that employed to answer Question 1: first, in the nature of the outcome measure used, and

second, in the unit of observation used. In this example the outcome measure is a binary, qualitative one, survived or died, rather than a quantity that can easily be added to or subtracted from other similar quantities. Other examples of qualitative outcomes already mentioned are: no dehydration versus some or severe dehydration, no chest indrawing versus chest indrawing. Some children, when followed over a period of time, will suffer several episodes of diarrhoea/respiratory morbidity. The qualitative nature of these outcomes makes it difficult to produce a summary measure for each child. For example, suppose a child had three episodes of respiratory morbidity during one of which s/he suffered chest indrawing. How does this child compare with another who suffered five episodes two of which were associated with chest indrawing, or a child who suffered one episode without chest indrawing? A more appropriate unit of observation than the child her/himself may be the individual episode. However, since the outcomes of different episodes in the same child may not be independent, the inclusion in the analysis, as separate observations, of two or more episodes occurring in the same child may introduce statistical problems. A further point to note is that whereas for Question 1 all children contribute information regardless of how many episodes of diarrhoea or respiratory morbidity they suffer, for this second question a child who does not suffer an episode of respiratory morbidity during the study provides no information to assist us in judging whether or not measles vaccination affects the risk of mortality from an episode of respiratory morbidity - i.e., in studies of this type of question, only children who become ill contribute useful information.

#### Question 3

"Can measles vaccination reduce the duration of episodes of respiratory morbidity?"

This question is concerned with the course of an episode (how long it lasts) once it has begun. Thus, as with Question 2, only children who suffer at least one episode of morbidity contribute any information to the study. In contrast with Question 2, however, the outcome measure used this time (duration) is quantitative, and thus for each child we may easily add the durations of each episode and divide by the number of episodes to obtain a single measure for each child (the mean duration). In this instance we may use the child as the unit of observation (thus avoiding the problem of non-independence which may arise when each episode is treated as a separate observation) and the quantitative outcome measure of mean duration.

#### Question 4

"Can measles vaccination reduce respiratory mortality in young children?"

At first sight this question appears to be the same as Question 2. Both are concerned with whether or not measles immunization can prevent respiratory mortality. There is, however, a difference between them.

Question 2 is concerned with the risk of mortality associated with an individual episode of morbidity - i.e., once the child has become ill, does measles immunization reduce the risk that the child will die from that episode? Question 4 is more general. Measles immunization may reduce a child's risk of respiratory mortality by reducing the risk of death associated with a particular episode of morbidity (the subject of Question 2) or by reducing the likelihood that the child will become ill in the first place (the subject of Question 1). To answer this question we might follow a group of (initially) healthy children to determine how many die from respiratory infections, and whether there is any difference in the mortality rates among those immunized against measles and those not. Note that all children contribute information regardless of whether or not they suffer any episodes of respiratory morbidity. The unit of observation is the child, and for each child the outcome measure is the binary, qualitative one of survived/died.

An additional point concerning these questions is worth noting. Questions 1 and 2 are subdivisions of Question 4 and it may be tempting to believe that by answering Questions 1 or 2 we will also be able to answer Question 4 - that if the answer to either

of the first two questions is yes, then the answer to Question 4 must also be yes. This is not so. For example, if measles immunization reduced respiratory morbidity by preventing the mildest episodes from occurring, but had no impact on the more severe, life-threatening episodes, the answer to Question 1 would be yes, but the answer to Question 4 no. While from a scientific viewpoint Questions 1 and 2 appear to be more precise, from a public health point of view the more general question, Question 4, may be the most important one to answer.

#### Question 5

"Can measles vaccination reduce the prevalence of respiratory morbidity in young children?"

The prevalence of respiratory morbidity is equivalent to the number of days of illness divided by the number of days of observation and depends upon both the number of episodes experienced and the duration of each episode. Thus, the appropriate (quantitative) outcome measure is the number of days of respiratory morbidity experienced by each child (adjusted if necessary for period of observation). All children contribute information and the unit of observation is the child.

#### 4. STATISTICAL TECHNIQUES FOR THE ANALYSIS OF QUANTITATIVE OUTCOMES

Binary outcomes such as healthy/diseased or survived/died are widely used in epidemiological studies and the statistical techniques for analysing data of this type are well known (Kleinbaum et al., 1982; Rothman, 1986). The same applies to outcomes that are continuous and more or less normally distributed. Some of the outcome measures that we have discussed do not belong in either of these two categories and have received less attention in the epidemiological literature. We now examine some of the statistical techniques that are available for the analysis of three non-binary, non-normal outcomes:

- (i) the number of episodes of illness experienced by each child (0,1,2,...),
- (ii) the average duration of each episode experienced by the child (e.g., 3.2 days),
- (iii) the total number of days of illness suffered by the child (0,1,2,...).

We illustrate approaches to the analysis of these measures with a worked example.

##### 4.1 Description of the data

Throughout the worked example we use data on diarrhoeal morbidity among 600 children (Annex). All the children were aged 6-23 months at the start of the study and were visited weekly for a period of 12 weeks. In addition to the data on diarrhoea morbidity, data are available on the sex of the child, on her/his age at the start of the study, and on whether or not s/he was weaned before the age of 4 months.

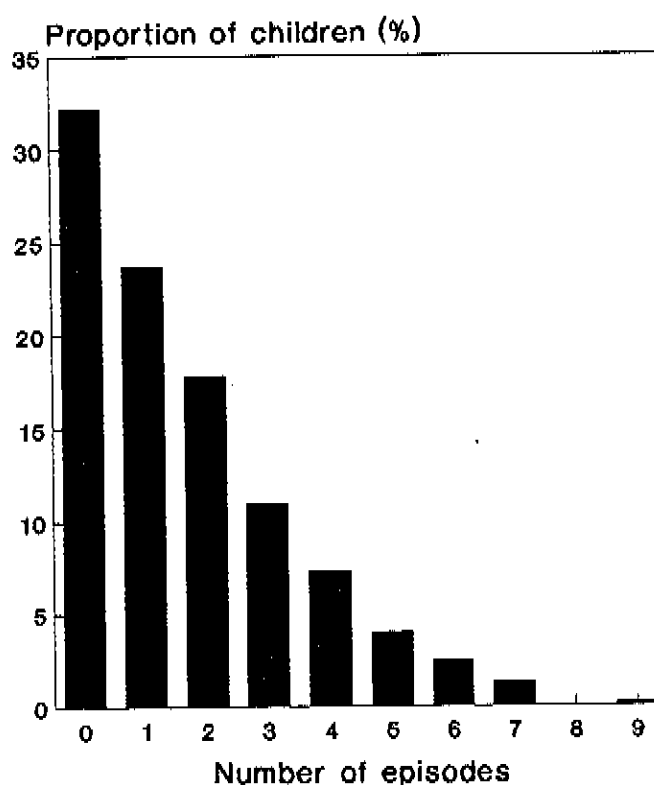
Diarrhoea (and respiratory) morbidity among young children is, typically, strongly associated with age. Usually, diarrhoea morbidity is greatest between the ages of 6 months and 2 years and decreases thereafter (Snyder and Merson, 1982). Our data set includes only children in this high-risk age group. By restricting the children to this relatively narrow range we reduce (but do not remove entirely) the age-dependent variability of diarrhoea rates in our study population and, for the sake of simplicity, we ignore the effect of age in the following analyses.

#### 4.2 Analysis of the number of episodes

Consider the question "Does early weaning lead to an increased incidence of diarrhoea morbidity?". The outcome measure appropriate to answering this question is clearly the number of episodes suffered by each child.

Over a period of 12 weeks, a total of 1005 episodes were recorded among the 600 children included in this analysis, an average of 1.675 episodes per child. The distribution of episodes among these children is presented in Figure 1. It is clear from Figure 1 that the number of episodes suffered by each child does not follow a normal distribution; the distribution is not symmetrical but has a mode at one end (0 episodes) and a long right tail (up to 9 episodes). This has certain statistical consequences to which we shall return later.

FIGURE 1: Distribution of episodes of diarrhoea among 600 children



##### 4.2.1 Categorized data

Analyses of data should always begin with simple techniques before going on to more complex and sophisticated procedures. In our exploration of the association between diarrhoea and early weaning we begin, therefore, by examining the data categorized according to the number of episodes suffered by the child. Before deciding on the categories to use, we examine more closely the distribution of episodes of diarrhoea among the children (Figure 1 and Table 1).

TABLE 1. Distribution of episodes of diarrhoea among 600 children aged 6 to 23 months

	Episodes of diarrhoea										Total
	0	1	2	3	4	5	6	7	8	9	
Children	193	142	107	66	44	24	15	8	0	1	600

In deciding what categories to use, we bear in mind some simple principles:

- Children who suffered no diarrhoea should be kept separate from children who did suffer diarrhoea (or a respiratory infection).
- Each category should include a "reasonable" number of children.
- The number of categories should not be too large.
- The width of each interval/category should, whenever possible, be equal.

In Table 2 we present one way of categorizing the data. Other categorizations are possible and would be reasonable. For example, we might treat children suffering three episodes of illness as one category and have as our last category children suffering four or more episodes.

TABLE 2. Distribution of episodes of diarrhoea according to weaning history among 600 children aged 6 to 23 months

Group	No. of episodes of diarrhoea				Total
	0	1	2	3+	
Weaned early	31 (16%) (26%)	19 (13%) (16%)	23 (21%) (19%)	47 (30%) (39%)	120 (20%)
Not weaned early	162 (84%) (34%)	123 (87%) (26%)	84 (79%) (18%)	111 (70%) (23%)	480 (80%)
	193	142	107	158	600

In Table 2 we present two sets of percentages. The top set are column percentages. Thus the 16% in the top left-hand corner of the table indicates that 16% of children who suffered no episodes of diarrhoea had been weaned early. In contrast, 30% of children who suffered three or more episodes had been weaned early. These percentages thus describe the risk of exposure given the number of episodes of illness. The lower set of figures represent row percentages. Thus, the 26% in the top left-hand corner indicates the probability of suffering no episodes of diarrhoea given that the child was weaned early. The figure of 23% in the bottom right-hand corner indicates the risk of suffering three or more episodes given that the child was not weaned early. These figures, therefore, indicate the proportion of children in each exposure category who suffer a given number of episodes of illness. Looking at the top set of percentages across the

top row of Table 2, there appears to be a trend: the more episodes of diarrhoea a child suffers the more likely it is that s/he was weaned early. The lower set of percentages are not particularly helpful when looking for a trend, but are important in describing the distribution of disease.

There are several possible explanations for the trend we have observed:

- (i) early weaning increases a child's risk of diarrhoea;
- (ii) children who suffer a lot of diarrhoea are more likely to be weaned early;
- (iii) another factor (or factors) associated with early weaning increases a child's risk of diarrhoea (confounding);
- (iv) bias in the selection of children for the study or in the recording of information on exposure (early weaning) and/or disease (diarrhoea) led to the observed association;
- (v) there is no underlying association between early weaning and diarrhoea, and the one we have observed is due to chance.

The last of these explanations is the most straightforward to assess. We may perform a statistical test to determine how likely it is that we would observe such a trend by chance (Kirkwood, 1988). For the data in Table 2, testing for a trend results in a chi-squared statistic

$$\chi^2 = 11.83.$$

In the absence of any underlying trend, this statistic is distributed as a chi-squared random variable with 1 degree of freedom. The probability of obtaining such a result by chance is less than one in a thousand ( $p < 0.001$ ). This is, therefore, strong evidence that an underlying trend towards increased diarrhoea in children weaned early exists. We conclude that it is very unlikely that (v) above explains the association we have observed.

Table 2 also enables us to estimate approximately two important epidemiological measures: the incidence rate ratio of diarrhoea in the two groups and the proportion of all episodes of diarrhoea that are associated with early weaning. First we estimate the total number of episodes that occurred in each group. In order to do this we need to make an assumption regarding the average number of episodes suffered by children in the last category (3+ episodes) (this is where the approximation comes in). We assume an average figure of four episodes. Then we obtain the following estimates:

Number of episodes among children weaned early	= $19 \times 1 + 23 \times 2 + 47 \times 4$
	= 253
Number of episodes among children not weaned early	= $123 \times 1 + 84 \times 2 + 111 \times 4$
	= 735

These estimates produce a total of 988 episodes, slightly below the 1005 actually observed. Using them, we can estimate the incidence rate ratio of diarrhoea in the two groups:

$$\begin{aligned} \text{Incidence rate ratio} &= \frac{\text{incidence rate in exposed}}{\text{incidence rate in unexposed}} \\ &= \frac{253/120}{735/480} = 1.38 \end{aligned}$$

We interpret this as indicating that children who were weaned early suffered 1.38 times as many (38% more) episodes of diarrhoea as (than) other children.

In order to estimate the proportion of all episodes of diarrhoea that might be associated with early weaning, we also estimate the number of episodes we would have expected to observe if all children had suffered episodes of diarrhoea at the same rate as children who were not weaned early.

$$\begin{aligned} \text{Expected number} &= (123 \times 1 + 84 \times 2 + 111 \times 4) \times 600/480 \\ &= 735 \times 600/480 = 918.75 \end{aligned}$$

Comparing this number with the (estimated) number of episodes observed (988) shows that we observed 69.25 more episodes than we would expect to have observed if the children who were weaned early had suffered diarrhoea rates similar to those in other children. This figure constitutes 7% (69.25/988) of the total number of episodes which occurred and suggests that if early weaning is responsible for the extra episodes of diarrhoea (explanation [i] above), then by successfully educating mothers to delay weaning until after 4 months of age it might be possible to reduce the number of episodes of diarrhoea in the population by about 7%. This estimate is of great value to us in assessing the possible public health importance of early weaning.

In Table 2 we presented the data in a very simple fashion, categorizing children according to the number of episodes of diarrhoea they suffered. We used four categories. We could have presented these data in an even simpler form, a 2 x 2 table, by categorizing children according to whether or not they suffered any episodes of diarrhoea (Table 3). In doing so, we have to sacrifice some of the information contained in Table 2, a sacrifice that is not without cost.

TABLE 3. Risk of diarrhoea according to weaning history among children aged 6 to 23 months

Group	Episodes of diarrhoea		Total
	None	One or more	
Weaned early	31(26%)	89(74%)	120(100%)
Not weaned early	162(34%)	318(66%)	480(100%)
	193	407	600

Table 3 reveals that, of 120 children who were weaned early, 89 (74.2%) suffered at least one episode. On the other hand, among children who were not weaned early, 318 (66.3%) suffered one or more episodes of diarrhoea. Thus, in our sample, children who were weaned early were more likely to suffer an episode of diarrhoea than children who were not weaned early.

Before performing any further statistical test we compare the "commonsense" strength of the evidence presented in Tables 2 and 3. In Table 3 there appears to be a difference between the two groups of children but, because the data are presented in such a simple way, we are unable to look for any pattern. In Table 2, on the other hand, we are able to observe a pattern, the trend. At an intuitive level, data that follow a pattern present a more convincing case for an association than data in which no pattern can be observed.

We may also compare the statistical strengths of the two methods of data presentation by performing a chi-squared test of the data presented in Table 3 and determining the statistical significance of the association. For Table 3,

$$\chi^2 = 2.41.$$

Comparing this with a table of values for a chi-squared statistic with 1 degree of freedom, we obtain a p-value of 0.12. The data, as presented in Table 3, therefore do not provide strong evidence of an association between early weaning and diarrhoea. This result is in contrast to that obtained from Table 2, which did provide strong evidence of an association. Thus the analysis of the data as presented in Table 2 is both intuitively and statistically more powerful than the analysis of the data as presented in Table 3. On the basis of an analysis of Table 3, we would conclude that these data do not provide any strong evidence of any underlying association between early weaning and diarrhoea, and that explanation (v) above is a plausible explanation for the results we have observed. The sacrifice we made in the detail of the data in order to present them in a simpler form has removed our ability to look for patterns and reduced the statistical power of our analysis.

The data as presented in Table 3 allow us to estimate one important epidemiological measure. A common way of quantifying the association between an exposure (early vs late weaning) and a disease (diarrhoea) is by a measure known as the risk ratio. This is defined in the following way:

$$\text{Risk ratio} = \frac{\text{risk in exposed}}{\text{risk in unexposed}} = \frac{89/120}{318/480} = 1.12$$

This estimate of the risk ratio indicates that, in our sample, children who were weaned early were 1.12 times as (12% more) likely to suffer at least one episode of diarrhoea as (than) children who were not weaned early.

#### 4.2.2 Untransformed, continuous data

An alternative approach to analysing these data is to treat the number of episodes of diarrhoea as a continuous variable and to compare the mean number of episodes among children weaned early with the mean number of episodes among other children. One way in which we may do this is by means of a t-test (Kirkwood, 1988).

TABLE 4. Results of a t-test comparing the mean number of episodes among 120 children weaned early with the mean number of episodes among 480 children not weaned early

Group	Mean number of episodes	Standard deviation	Number
Weaned early	2.23	2.01	120
Not weaned early	1.54	1.63	480

$$t = 4.00, \quad p < 0.0001$$

This analysis provides strong evidence that children who are weaned early suffer more episodes of diarrhoea than other children ( $p < 0.0001$ ). Over the period of the study, children weaned early suffered, on average, 0.70 more episodes of diarrhoea than other children (95% confidence interval 0.36, 1.04). We may also calculate the (incidence) rate ratio of diarrhoea in the two groups. This is done as follows:

$$\text{Rate ratio} = \frac{\text{rate in the exposed}}{\text{rate in the unexposed}} = \frac{2.23}{1.54} = 1.45$$

The rate ratio indicates that children weaned early suffered 1.45 times as many episodes of diarrhoea as children who were not weaned early or, expressed in another way, children weaned early suffer, on average, 45% more episodes of diarrhoea than other children. This (exact) estimate of the rate ratio is larger than the (approximate) estimate that we obtained from Table 2 (1.38). The approximation (and hence mismeasurement) involved in the calculations used with Table 2 has led to an underestimate of the rate ratio. While the continuous analysis gives a better estimate of the rate ratio, the analysis of the categorized data allows us to look for trends that are more convincing than simple differences.

It is also instructive to contrast the estimate of the rate ratio with the estimate of the risk ratio (- 1.12) which we obtained from Table 3. The rate ratio is larger than the risk ratio and, for common recurrent diseases like diarrhoea and respiratory infections, is the better measure. To understand why rate ratio is to be preferred to risk ratio consider a hypothetical example. Suppose that we are studying a population in which early weaning leads to increased incidence of diarrhoea, and that diarrhoea is very common in our study population. To take an extreme situation, suppose that every child suffers at least one episode of diarrhoea each year. Then, if we follow the population for one year, the risk in both the exposed and unexposed populations will be 1.0 and so the risk ratio will be one; i.e., we will not observe any association between early weaning and risk of diarrhoea. If, instead, we consider the number of episodes (rate) in each group, we shall observe a higher rate in the children weaned early and thus we still detect the association. The relationship between risk ratio and rate ratio may be summarized as follows: over a short follow-up period during which the disease of interest is rare, the risk ratio is approximately equal to the rate ratio, but as the period of follow-up becomes longer the risk ratio will move away from the rate ratio towards 1.0.

From the data in Table 4 we may once again estimate the proportion of all episodes of diarrhoea that are attributable to early weaning. The mean number of episodes among children not weaned early was 1.54. Therefore, in a population of 600 children none of whom were weaned early, we would expect to observe a total of 921.25 (600 x 1.5354) episodes. In practice, we observed 1005 episodes. This suggests that 8.3% of all episodes of diarrhoea may be attributable to early weaning. This (exact) estimate is greater than the (approximate) estimate obtained from Table 2.

#### 4.2.3 Transformed, continuous data

Strictly speaking, a t-test is only valid when comparing the mean values of two samples from normally distributed populations with equal standard deviations. The number of episodes of diarrhoea in each group of children is clearly not normally distributed (Figure 1). Fortunately, when the two samples are reasonably large (each greater than 30, say), the t-test is "robust" in the presence of departures from normality. Thus, this is not likely to be a serious problem in the analysis of this data set. There is, however, some evidence that the standard deviations of the two groups are different (2.01 versus 1.63, F=1.52, p=0.02). Again, for large samples, this does not constitute a serious problem since a modified version of the t-test may be used (Kirkwood, 1988).

One possible approach to dealing with the problems of non-normality and unequal standard deviations when they do arise is to transform the data. By transforming the data it is hoped to make the distribution more normal (in the statistical sense) and/or the standard deviations more equal. However, with a distribution like that in Figure 1, with a mode occurring at one end of the distribution (zero episodes), no sensible transformation will produce a normal-looking distribution. When dealing with data with a long right tail (Figure 3) the log transformation (log[1+number of episodes]) is commonly used. Performing a t-test on the log-transformed data produces the following results (Table 5).

TABLE 5. Results of a t-test comparing the mean of  $\log(1 + \text{number of episodes})$  among 120 children weaned early with the mean among 480 children not weaned early

Group	Mean of $\log(1+\text{episodes})$	Standard deviation	Number
Weaned early	0.42	0.29	120
Not weaned early	0.32	0.27	480

$$t = 3.49, \quad p < 0.0005$$

While the transformation has not resulted in two normally distributed groups of children, it has been effective in rendering the standard deviations more equal. There is now no evidence that they differ between the two groups ( $F=1.2, p=0.20$ ). The results of this analysis confirm those of Table 4, although with a somewhat reduced t-statistic. In fact the t-statistic obtained using the transformed data (3.49) is very similar to that which would have been obtained from the original data using the modified t-test (3.53). It is not, however, easy to derive any simple and meaningful epidemiological measures of the association between diarrhoea and early weaning from this analysis.

#### 4.2.4 Non-parametric analysis

An alternative approach to the problem of comparing the means of samples from non-normal populations or samples with unequal standard deviations is to perform a non-parametric test. Such tests do not make any assumptions about the distributions from which the samples are drawn. An appropriate non-parametric test for comparing two unmatched, independent samples is the Wilcoxon rank sum test (Kirkwood, 1988). The results of performing this test are presented in Table 6.

TABLE 6. Results of a Wilcoxon rank sum test comparing the number of episodes of diarrhoea among 120 children weaned early with the number of episodes among 480 children not weaned early

Group	Number	Sum of ranks	Expected	Standard deviation
Weaned early	120	41 790	36 060	1651.96
Not weaned early	480	138 510	144 240	1651.96

$$Z = 3.47, \quad p = 0.0005$$

This result confirms the statistical significance of the association between diarrhoea and early weaning observed when t-tests were conducted. The value of the Z statistic is very close to that of the t-statistic resulting from the analysis of the log-transformed data.

The Wilcoxon rank sum test makes no assumptions about the distribution of the number of episodes of diarrhoea per child and is almost as powerful as the parametric t-test. It is therefore excellent for hypothesis-testing in this situation. Unfortunately, however, it does not enable us to estimate any useful epidemiological measure(s) with which to assess the biological and/or public health importance of the association. Thus, while it may be useful as a confirmatory test, used alone it is not very revealing.

In Table 2 we presented the data in a simple form with children categorized according to the number of episodes of diarrhoea they experienced. We were able to look for pattern (trend) in the data and to conduct a powerful statistical analysis. The analysis of Table 2 provided strong evidence of an underlying association between early weaning and increased diarrhoea rates ( $p < 0.001$ ). In addition, we were able to estimate two important epidemiological parameters: the rate ratio of diarrhoea and the proportion of episodes of diarrhoea associated with early weaning. In trying to present the data even more simply (Table 3), we sacrificed much of the information in the data. We were no longer able to look for trends in the data and our statistical analysis failed to provide strong evidence of an association between early weaning and diarrhoea. The epidemiological parameter estimated from Table 3, the risk ratio, was less appropriate than the rate ratio estimated from Table 2.

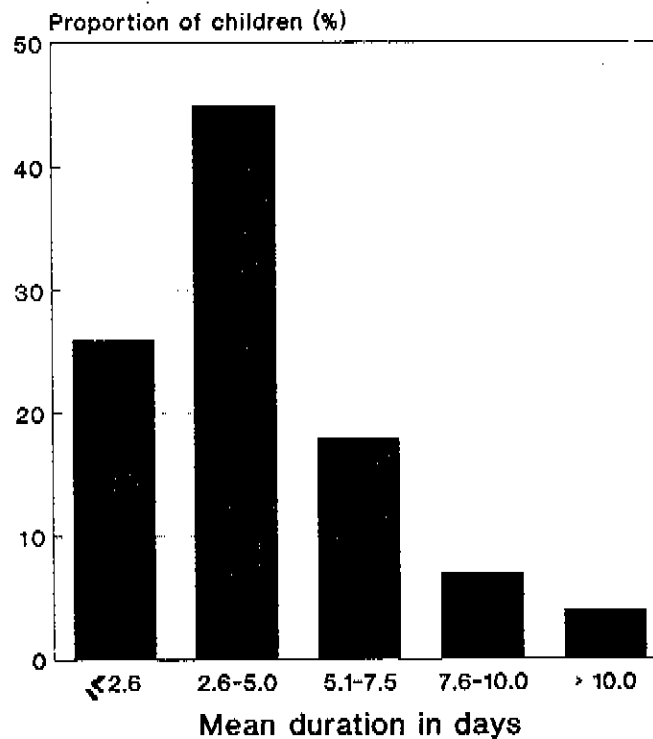
An alternative and complementary approach to the analysis involved the comparison of the mean number of episodes of diarrhoea in each group, treating the number of episodes as a continuous outcome variable (Table 4). This approach also provided strong evidence of an association between early weaning and diarrhoea and we were able to estimate the rate ratio of diarrhoea and the proportion of all episodes associated with early weaning. While the use of the t-test is, strictly speaking, only valid when comparing samples drawn from normally distributed populations with equal standard deviations, the results obtained applying a t-test to the raw data were similar to those obtained from a t-test of the log-transformed data (Table 5) and from a non-parametric test (Table 6).

#### 4.3 Analysis of the duration of episodes

Consider the question "Does early weaning lead to increased duration of diarrhoeal episodes?". This question is concerned with the course of an episode (how long it lasts) once it has begun. Only children who suffer at least one episode of morbidity contribute any information towards answering this question. Among the 600 children studied, 407 suffered at least one episode of diarrhoea. We restrict our attention to these children. Altogether, 1005 episodes of diarrhoea occurred (a mean of 2.47 episodes per child experiencing diarrhoea), lasting a total of 4222 days (a mean of 4.20 days per episode). The mean duration per child is shown in Figure 2.

While still very skewed to the right, this distribution appears somewhat closer to a normal distribution than that of number of episodes of diarrhoea. The minimum mean duration observed was 1.0 days (28 children), and the maximum 17.0 days (2 children), with a median of 3.5 days. Mean duration was slightly correlated with incidence rate (number of episodes divided by days at risk), the correlation co-efficient of this association being 0.14 ( $p < 0.01$ ). Thus, in this population, there is some statistical evidence that children who suffer more episodes of diarrhoea suffer longer episodes. The correlation itself is, however, weak.

FIGURE 2: Mean duration of episodes of diarrhoea among 407 children aged 6-23 months



4.3.1 Categorized data

We begin by examining the data with children grouped into four categories of duration (Table 7).

TABLE 7. Distribution of mean duration per episode of diarrhoea (4 categories) among 89 children weaned early and 318 children not weaned early

Group	Mean duration per episode of diarrhoea				Total
	<2.5 days	<5 days	<7.5 days	7.5+ days	
Weaned early	17 (16%)	35 (19%)	22 (31%)	15 (33%)	89 (22%)
Not weaned early	89 (84%)	148 (81%)	50 (69%)	31 (67%)	318 (78%)
	106	183	72	46	407

$$X^2 = 7.86, p = 0.005$$

A trend is apparent, with children who were weaned early tending to suffer longer episodes of diarrhoea than children who were not weaned early. It is highly unlikely that this trend is due to random variation (p=0.005).

4.3.2 Untransformed, continuous data

TABLE 8. Results of a t-test comparing the mean duration of diarrhoea among 89 children weaned early with the mean duration among 318 children not weaned early

<u>Group</u>	<u>Mean duration</u>	<u>Standard deviation</u>	<u>Number</u>
Weaned early	4.66	2.48	89
Not weaned early	4.00	2.71	318

$$t = 2.04, \quad p = 0.04$$

The results of a t-test comparing the mean durations among children weaned early and children not weaned early (Table 8) suggest that, on average, episodes of diarrhoea in children who were weaned early lasted 0.66 days (14%) longer than those in children who were not weaned early. It is unlikely that this difference is due to sampling variation ( $p = 0.04$ ).

The sample size is large enough to prevent any major problems arising due to non-normality of the distribution, and there is no strong evidence of a difference between the standard deviations of the two groups ( $F=1.19, p=0.32$ ). We shall not, therefore, repeat the additional analyses that we performed when examining number of episodes as an outcome, but simply mention the results in passing. An analysis of the log-transformed data results in a t-statistic of 2.63 ( $p=0.009$ ). A non-parametric analysis results in a Z-statistic of 3.05 ( $p=0.002$ ).

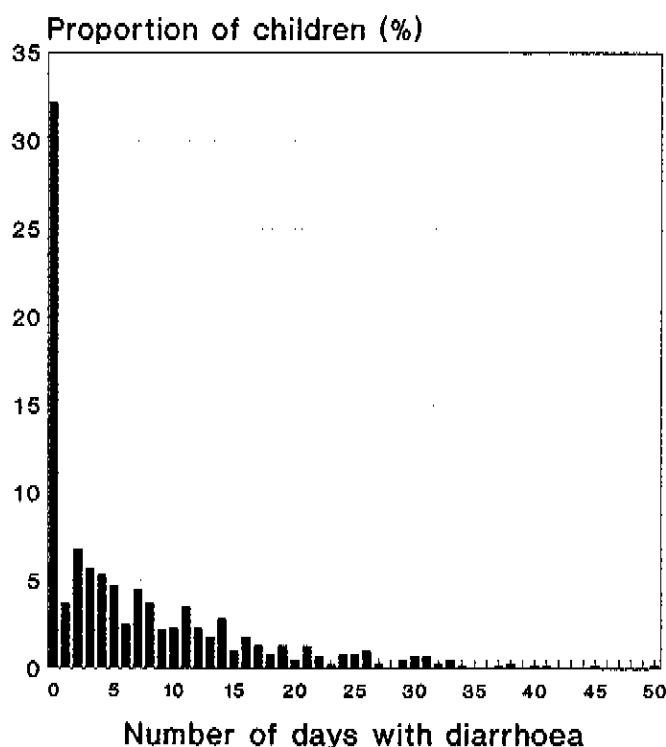
Overall, we have observed some evidence that children who were weaned early suffered longer episodes of diarrhoea than children who were not. On average, episodes in children who were weaned early lasted 0.66 days longer than episodes in other children.

4.4 Analysis of the total number of days of morbidity

Consider the question "Do children who are weaned early suffer a larger total number of days (i.e., higher prevalence) of diarrhoea than children who are not weaned early?".

As with the analysis of number of episodes, all children contribute information to the analysis. The 600 children included in the study suffered a total of 4222 days with diarrhoea, with a mean of 7.04 days per child and a median of 4 days. The distribution between children of the number of days with diarrhoea is presented in Figure 3. It is clearly non-normal with a mode at 0 days (193 children) and a long right tail up to a maximum of 50 days (1 child).

FIGURE 3: Distribution of number of days with diarrhoea among 600 children aged 6-23 months



We shall not go through the analyses of the association between early weaning and number of days with diarrhoea in detail. The appropriate analytical techniques are the same as those used in the analysis of the relationship between early weaning and the number of episodes of diarrhoea. We shall instead present a summary of the results and invite readers to perform their own analyses of these data.

Results of statistical analyses

(i) Analysis of categorized data

Group	Number of days with diarrhoea				Total
	None	1-10	11-20	21-50	
Weaned early	31 (16%)	40 (16%)	25 (24%)	24 (44%)	120 (20%)
Not weaned early	162 (84%)	208 (84%)	79 (76%)	31 (56%)	480 (80%)
Total	193	248	104	55	600

Test for trend,  $\chi^2 = 21.48$ , 1df,  $p < 0.00001$

- (ii) t-test, comparing number of days in each group (untransformed data)

Mean number of days among children weaned early = 10.43

Mean number of days among children not weaned early = 6.19,

$$t = 4.08, \quad p < 0.0001$$

- (iii) Estimation of prevalence rate ratio

$$\text{Prevalence rate ratio} = \frac{10.43}{6.19} = 1.69$$

- (iv) Percentage of all days of diarrhoea attributable to early weaning

$$= 12.1\%$$

- (v) t-test, log-transformed data

$$t = 3.81, \quad p = 0.0002$$

- (vi) Wilcoxon rank sum test

$$Z = 3.87, \quad p = 0.0001$$

These results suggest a strong link between early weaning and an increased prevalence (number of days) of diarrhoea. Children who had been weaned early suffered, on average, an extra 4.24 days of diarrhoea (69% more) compared with children who had not been weaned early. These results suggest that, if early weaning causes the extra days of diarrhoea, the prevention of early weaning in this community could reduce the prevalence of diarrhoea by about 12%.

## 5. DISCUSSION

### 5.1 Choice of outcome measure

We began by describing the outcome measures available for use in prospective studies of childhood diarrhoea and respiratory infections, and by stating that the choice of outcome measure was dependent upon the research question being posed. Examples of research questions were examined with regard to their implications for the choice of outcome measure. From these considerations it is possible to present a broad summary of the relationship between type of research question and both the appropriate unit of observation and the choice of outcome measure (Table 9).

For questions concerned with the transition from health to ill-health (occurrence of disease), the appropriate (quantitative) outcome measure is the number of episodes suffered by the child and the appropriate unit of observation is the individual child. All children contribute information to the analysis.

For questions concerned with the course of the episode after the onset of illness, the outcome measure may be quantitative (duration) or qualitative (severity of the episode) depending on the question itself. The appropriate unit of observation depends on the outcome measure. When the outcome is quantitative (duration), the unit of observation is the child. When the the outcome is qualitative, the appropriate unit of observation is the individual episode rather than the child. Whether the outcome is quantitative or qualitative, only children who suffer an episode of illness contribute information to the analysis.

TABLE 9. Summary of units of observation and outcome measures appropriate to different types of research question

Type of research question	Unit of observation	Type of outcome measure	Appropriate outcome measure(s)
Questions concerning associations between exposures and disease incidence	child	quantitative	number of episodes
Questions concerning associations between exposures and the course of an episode once it has started	child	quantitative	mean duration
	or episode	or qualitative	or severity
Questions concerning associations between exposures and disease prevalence	child	qualitative	severity
Questions concerning associations between exposures and the transition from a state of health to a specified outcome (e.g., death)	child	qualitative	severity

The appropriate outcome measure for questions concerning the prevalence of illness is quantitative (number of days of illness, adjusted if necessary for different periods of follow-up). The unit of observation is the child and all children contribute information to the study.

The final category of question listed in Table 9 includes questions of the form "Can measles immunization prevent respiratory mortality?". The outcome measure for such questions is qualitative (in this example survived/died). The unit of observation is the child and all children contribute information to the study regardless of whether or not they became ill during the course of the study.

## 5.2 Statistical methods

We have performed a series of analyses to examine the association between early weaning and each of three measures of diarrhoea morbidity: number of episodes, mean duration of an episode, and number of days with diarrhoea.

In analysing the association between early weaning and incidence of diarrhoea (number of episodes), we found strong evidence of a link. By grouping the children into several categories we were able to demonstrate a trend, children who were weaned early tending to suffer more episodes of diarrhoea. A comparison of the mean number of episodes in the two groups enabled us to estimate two epidemiological measures of importance, the rate ratio of diarrhoea and the proportion of all episodes associated with early weaning. We explained why the rate ratio is a more appropriate measure than the risk ratio in studies of childhood diarrhoea and respiratory infections. The use of the t-test to compare the mean number of episodes was not entirely appropriate because it assumes that the samples are drawn from normally distributed populations with equal

standard deviations. There was some evidence that the standard deviations were not equal. The use of the log transformation was effective in rendering the standard deviations of the two groups more similar. A comparison of the means of the log-transformed data did not, however, provide us with any easily interpretable epidemiological measure. The same drawback applied to the use of the Wilcoxon rank sum test. The findings obtained from analyses of both mean duration and number of days with diarrhoea were broadly similar to those for number of episodes. We recommend that analyses of prospective studies of childhood diarrhoea and respiratory infections should use categorized data to look for evidence of trends in the data. Epidemiological parameters of interest (rate ratio, proportion of episodes attributable, etc.) can be estimated from the raw, untransformed data. In studies with reasonable sample sizes (say more than 30 children in each group) departures from normality are unlikely to cause any statistical problems. Where there is evidence that standard deviations vary between different groups of children, the statistical significance of any difference in mean values between the different groups should be confirmed. One way of doing this would be to perform a Wilcoxon rank sum test.

### 5.3 Interpretation of associations

Finally, we remind ourselves of the need to exercise caution in the interpretation of the crude associations we have observed between early weaning and the three outcome measures that we have considered here. We outlined above five possible explanations for an observed association between early weaning and increased diarrhoea (incidence, prevalence, or duration). These are:

- (i) early weaning causes increased diarrhoea;
- (ii) increased diarrhoea causes early weaning;
- (iii) early weaning is not itself associated with increased diarrhoea but is related to other confounding factors that are;
- (iv) bias in the selection of the children for inclusion in the study, or in the ascertainment of exposure (early weaning), or in the ascertainment of disease (diarrhoea) led to the observed association;
- (v) there is no underlying association between early weaning and diarrhoea; the observed association is due to sampling variation.

Of these five possible explanations we may with some safety dismiss the last. We have conducted a number of statistical tests which have indicated that it is extremely unlikely that the associations which we have observed arose by chance. Any one of the other four possible explanations could be the correct one. In order to determine which is the most likely explanation, substantial amounts of additional work need to be done, in examining the design of and methods used in the study, and in further data analysis.

### ACKNOWLEDGEMENTS

The Diarrhoeal Disease Control Programme of the World Health Organization provided financial support for the preparation of this document. The authors would like to thank J. Martines and T. Mertens for their constructive comments on earlier drafts of this document.

REFERENCES

- Kirkwood, B.R. (1988) Essentials of medical statistics. Oxford, Blackwell Scientific Publications.
- Kleinbaum, D.G., Kupper, L.L. and Morgenstern, H. (1982) Epidemiologic research. New York, Van Nostrand Reinhold.
- Rothman, K.J. (1986) Modern epidemiology. Boston, Little, Brown and Company.
- Snyder, J.D. and Merson, M.H. (1982) The magnitude of the global problem of acute diarrhoeal diseases: a review of active surveillance data. Bulletin of the World Health Organization, 60: 605-613.
- UNICEF (1990) The state of the world's children. Oxford, Oxford University Press.

## ANNEX

## Diarrhoea morbidity data for 600 children followed for 12 weeks

Age in months	Sex	No. of episodes	No. of days	Early weaning	Age in months	Sex	No. of episodes	No. of days	Early weaning
10	1	3	21	2	14	1	0	0	2
20	1	1	17	2	14	2	0	0	2
17	1	5	30	1	10	1	0	0	2
23	1	1	4	1	9	1	0	0	2
21	1	2	12	2	21	2	0	0	2
22	1	2	14	2	10	2	1	2	2
20	1	3	17	1	13	2	1	3	2
20	1	4	30	2	17	2	0	0	2
11	2	2	34	2	13	2	0	0	2
9	1	3	33	1	15	1	0	0	2
11	2	4	26	2	10	1	1	5	2
13	2	3	29	2	10	1	1	14	2
10	1	4	33	2	11	2	0	0	2
14	1	7	41	2	11	2	2	5	2
15	2	0	0	1	20	1	2	5	2
9	2	4	19	2	20	1	0	0	2
20	2	3	29	2	11	2	3	20	2
6	2	4	25	2	10	1	1	2	2
11	1	4	15	2	7	2	1	2	2
8	1	2	21	2	18	1	3	7	2
17	2	4	32	2	7	1	2	5	2
19	1	4	13	1	20	2	0	0	1
21	1	3	26	2	15	1	0	0	2
20	1	0	0	2	7	2	1	2	2
23	1	2	8	2	23	1	2	2	2
14	1	3	16	2	6	2	1	3	2
9	1	1	9	2	14	2	2	7	2
13	1	2	15	1	6	2	1	4	2
7	2	1	2	2	9	1	2	5	2
21	2	3	19	2	9	2	1	2	2
14	1	2	14	2	8	2	1	3	2
18	2	1	10	2	8	1	0	0	2
23	2	1	10	1	15	1	1	10	2
20	1	1	9	1	20	2	1	2	2
21	2	3	24	1	6	1	0	0	2
16	1	4	19	2	22	1	3	12	2
10	1	3	40	2	15	1	0	0	2
17	1	1	1	2	15	2	2	5	2
7	2	1	4	2	15	2	3	15	2
16	2	1	2	2	15	2	2	17	2
15	2	1	3	2	7	2	1	4	2
16	2	1	1	1	8	1	2	9	2
13	2	0	0	2	18	1	1	3	1
11	2	2	5	2	21	1	2	4	1
17	1	1	5	2	16	1	3	3	2
9	2	1	2	2	18	2	1	1	2
6	1	0	0	2	8	1	4	7	2
13	1	2	5	2	19	2	5	14	2
21	1	1	1	2	6	2	1	9	2
9	1	4	11	2	14	2	0	0	2
17	2	2	8	1	11	2	0	0	2
18	2	2	3	1	13	1	2	4	2
23	1	1	2	2	7	1	3	4	2

Annex

Age in months	Sex	No. of episodes	No. of days	Early weaning	Age in months	Sex	No. of episodes	No. of days	Early weaning
11	1	2	11	2	7	1	4	18	2
21	1	0	0	1	9	2	1	3	2
15	1	3	11	2	18	2	0	0	2
22	1	1	2	2	14	1	0	0	2
11	1	3	6	1	9	1	1	2	2
17	2	0	0	2	19	2	2	3	2
10	2	3	26	1	19	2	0	0	2
22	2	0	0	2	12	2	1	3	2
13	1	2	4	1	20	2	4	11	2
19	2	2	2	2	19	1	0	0	1
20	1	2	7	2	23	2	0	0	2
18	2	2	4	2	19	2	1	5	2
6	2	0	0	2	8	2	0	0	2
16	1	2	6	2	7	2	1	3	1
16	1	0	0	2	11	2	0	0	2
16	1	1	3	2	21	2	0	0	2
9	1	3	16	2	10	1	0	0	1
7	1	1	5	2	23	1	0	0	2
21	1	0	0	2	12	1	0	0	1
17	2	2	4	2	20	2	0	0	2
8	1	5	27	1	20	1	0	0	2
12	2	1	1	1	12	2	3	7	2
17	1	2	14	2	10	1	1	3	2
15	1	3	4	1	10	2	1	2	2
22	2	1	2	2	10	1	0	0	2
9	2	1	2	2	10	2	0	0	2
11	1	2	7	1	12	2	0	0	2
6	2	5	22	2	19	2	0	0	2
22	1	0	0	1	11	2	0	0	2
6	1	3	7	2	15	2	3	7	2
16	2	0	0	2	19	1	1	3	2
19	1	0	0	2	6	1	2	7	2
21	1	0	0	2	18	1	0	0	2
22	2	0	0	2	23	2	0	0	1
11	1	3	7	2	20	2	0	0	2
11	1	2	8	2	19	2	1	4	2
14	1	1	4	2	20	2	0	0	2
17	1	1	2	2	21	2	0	0	2
22	1	5	10	2	13	1	1	5	1
20	2	4	11	2	16	1	0	0	2
16	2	0	0	2	8	1	1	4	2
10	2	2	8	2	18	1	1	3	2
11	2	0	0	2	22	2	2	8	1
6	2	2	9	2	10	2	0	0	1
15	2	1	1	2	15	2	1	2	1
9	2	1	4	2	23	2	0	0	2
20	2	1	1	2	18	1	0	0	2
8	2	4	12	2	16	2	0	0	1
18	2	2	4	2	13	1	2	3	2
22	2	2	2	2	21	2	0	0	2
21	1	0	0	2	21	1	1	4	2
18	2	1	1	2	8	1	2	8	2
14	1	0	0	2	22	1	0	0	2
14	2	1	1	2	10	1	1	1	2
14	2	3	6	2	6	2	0	0	2

Annex

Age in months	Sex	No. of episodes	No. of days	Early weaning	Age in months	Sex	No. of episodes	No. of days	Early weaning
7	1	5	22	2	22	1	0	0	2
21	1	2	9	2	23	1	0	0	2
15	1	0	0	2	14	2	3	10	2
19	1	2	8	2	22	1	0	0	2
17	1	1	1	2	9	2	1	9	1
20	1	0	0	2	8	2	1	8	2
6	1	3	6	2	7	1	3	24	2
7	1	1	8	2	19	2	0	0	2
11	2	0	0	2	23	1	0	0	1
6	2	1	1	2	8	1	2	3	1
7	2	2	11	1	19	1	2	14	2
22	2	3	7	2	21	2	0	0	1
19	2	1	2	2	23	1	1	2	1
9	1	0	0	1	8	2	1	12	2
6	2	0	0	2	22	1	0	0	1
16	1	3	11	2	9	2	0	0	1
10	2	1	1	2	8	1	1	2	2
9	1	0	0	2	11	1	2	14	2
18	2	0	0	2	9	1	0	0	1
16	1	0	0	2	17	2	0	0	1
15	1	2	4	2	14	1	5	50	2
20	2	2	3	2	17	1	5	9	2
14	2	0	0	2	21	2	1	6	2
6	2	0	0	2	9	2	0	0	2
16	1	1	9	2	15	2	2	4	2
16	1	0	0	2	22	2	1	2	2
6	2	0	0	2	7	2	2	7	1
15	1	2	12	2	21	1	2	11	2
21	2	1	1	2	9	2	1	3	2
16	2	1	2	2	13	2	0	0	2
9	2	0	0	2	9	1	0	0	2
10	2	1	13	2	6	2	0	0	2
18	1	0	0	2	22	1	0	0	2
13	1	1	7	2	8	2	0	0	2
17	1	1	1	2	20	1	1	3	2
23	2	0	0	2	7	1	0	0	2
7	2	0	0	2	23	1	0	0	2
23	2	0	0	2	6	2	0	0	2
22	1	1	7	2	16	1	2	12	2
16	1	1	5	2	7	1	0	0	2
21	2	2	2	2	8	2	0	0	2
18	2	0	0	2	8	1	1	2	2
20	2	0	0	2	11	2	1	3	2
11	2	2	8	2	12	1	0	0	2
15	1	0	0	2	8	2	0	0	1
7	1	1	1	2	8	1	0	0	2
15	2	3	11	2	10	2	0	0	2
21	2	0	0	2	12	2	2	4	2
20	2	0	0	2	16	1	0	0	1
10	1	1	2	2	16	2	0	0	1
19	2	1	1	2	11	1	0	0	2
13	1	2	7	2	9	1	1	3	2
18	1	4	14	2	21	2	1	2	2
16	2	0	0	2	22	1	0	0	2
12	1	0	0	2	11	2	0	0	2

Annex

Age in months	Sex	No. of episodes	No. of days	Early weaning	Age in months	Sex	No. of episodes	No. of days	Early weaning
6	1	1	2	2	10	1	0	0	1
16	2	2	7	2	19	1	0	0	2
18	2	1	5	2	21	2	2	10	2
18	1	2	5	1	6	1	1	1	1
10	2	1	2	2	12	2	0	0	2
18	2	2	7	2	10	1	0	0	2
22	2	2	5	2	16	2	0	0	2
16	2	0	0	1	11	1	1	14	2
12	2	1	4	2	17	1	1	3	2
18	2	1	3	2	18	1	0	0	2
8	2	0	0	2	14	1	1	3	2
11	1	4	23	2	19	1	0	0	2
14	1	1	8	2	7	1	0	0	2
11	2	0	0	2	21	2	2	8	1
21	2	1	5	2	11	2	0	0	1
23	2	0	0	2	15	1	1	2	2
18	2	0	0	2	22	2	0	0	2
8	2	2	4	1	8	2	2	6	1
13	2	1	1	2	9	1	1	3	2
9	1	2	6	2	9	1	1	5	2
14	1	3	10	2	23	1	0	0	2
11	2	4	12	1	13	1	0	0	2
8	2	2	19	1	16	1	2	8	2
22	1	4	10	2	8	2	0	0	1
23	1	1	3	2	9	2	0	0	2
21	1	0	0	2	9	1	0	0	2
10	2	0	0	2	20	1	0	0	2
10	2	6	19	2	18	2	0	0	2
12	1	3	10	2	12	1	0	0	2
17	1	0	0	2	12	2	1	2	2
17	1	1	3	2	17	1	1	4	2
22	1	0	0	2	16	1	0	0	2
22	1	2	4	2	19	1	0	0	2
19	1	0	0	1	14	2	0	0	2
15	1	3	10	2	7	2	0	0	2
17	1	0	0	2	11	2	0	0	2
10	2	0	0	1	19	2	0	0	2
10	2	2	7	2	7	1	1	1	2
8	2	1	3	2	10	1	0	0	2
6	2	2	10	2	10	2	0	0	2
19	2	0	0	1	13	2	0	0	1
12	1	2	11	2	16	1	1	7	2
9	2	2	4	1	12	1	1	7	2
11	2	1	2	2	23	2	3	10	2
9	1	4	14	2	23	1	0	0	2
18	1	0	0	2	21	2	0	0	2
14	1	2	10	1	21	1	4	38	1
13	2	3	8	2	6	2	3	11	2
13	1	0	0	1	6	2	1	11	2
16	2	0	0	2	8	1	2	9	2
17	1	1	1	2	17	2	2	6	2
21	1	4	7	1	15	2	0	0	2
21	1	1	5	1	23	2	1	5	1
15	2	2	7	2	15	2	0	0	2
19	2	0	0	2	20	2	2	33	2

Annex

Age in months	Sex	No. of episodes	No. of days	Early weaning	Age in months	Sex	No. of episodes	No. of days	Early weaning
16	1	3	9	1	21	1	1	3	2
10	1	2	15	1	12	2	4	21	2
20	1	0	0	2	6	1	0	0	2
15	1	1	3	2	14	1	0	0	2
9	1	0	0	2	10	2	3	12	2
23	2	2	2	2	6	1	5	20	2
11	2	0	0	1	21	2	1	4	2
7	1	1	5	2	14	2	1	3	1
9	2	1	5	2	17	2	1	2	2
10	2	2	6	2	9	2	5	26	1
10	2	1	3	2	14	2	6	31	1
8	1	3	11	2	9	2	4	17	2
13	2	0	0	2	8	1	6	14	1
23	2	2	6	2	9	2	3	13	2
8	2	0	0	2	14	2	3	25	1
7	2	2	4	2	23	1	2	5	1
8	2	2	8	2	9	2	3	11	1
22	1	7	25	2	6	2	4	17	1
10	2	1	5	2	18	2	4	13	1
16	1	6	16	2	12	1	6	25	1
17	2	1	5	2	10	1	2	26	1
9	2	2	4	2	17	2	4	12	2
18	1	4	12	2	22	2	3	25	1
6	1	3	8	2	14	2	7	32	1
23	2	1	2	2	11	2	4	13	2
22	2	3	11	2	11	2	3	13	2
19	1	1	2	2	21	2	6	18	2
15	2	5	14	2	10	1	0	0	2
17	2	2	12	2	14	2	0	0	2
19	1	0	0	2	14	2	1	7	1
6	1	2	4	2	14	2	1	7	1
6	2	6	18	2	11	2	0	0	1
18	1	7	18	2	10	2	3	11	2
20	2	0	0	2	10	2	0	0	2
15	1	2	4	2	11	2	0	0	2
19	2	4	14	2	22	2	1	2	1
18	2	3	9	2	6	1	2	8	2
7	2	6	37	2	7	2	3	8	2
18	2	2	7	2	18	1	0	0	2
17	1	4	17	2	11	1	1	4	2
8	2	0	0	2	9	1	2	12	2
11	2	2	5	2	16	1	4	11	2
15	2	4	14	2	22	1	1	2	2
9	2	3	13	2	9	2	6	16	2
22	2	3	12	1	13	1	2	6	2
9	1	5	27	2	12	2	5	16	2
14	1	5	21	1	12	2	2	6	2
16	1	2	8	2	20	2	3	22	1
21	2	1	3	1	23	2	2	8	2
11	1	3	30	2	21	2	3	6	2
23	2	3	13	2	14	1	4	16	2
17	1	0	0	2	19	1	3	9	2
22	2	0	0	2	18	1	2	6	2
7	1	0	0	2	22	2	1	4	2
11	1	0	0	2	6	1	2	5	2

Annex

Age in months	Sex	No. of episodes	No. of days	Early weaning	Age in months	Sex	No. of episodes	No. of days	Early weaning
7	2	0	0	2	19	2	0	0	2
11	2	0	0	2	8	1	5	15	2
22	2	1	5	2	20	2	6	19	2
22	2	1	2	2	7	2	3	7	2
8	2	1	1	2	18	1	0	0	2
16	2	1	12	2	17	1	5	24	2
6	2	4	21	1	10	1	2	8	2
23	2	7	30	2	14	1	6	14	2
6	2	4	15	1	8	2	4	11	1
22	2	3	14	1	10	1	4	21	2
15	2	5	31	2	10	1	1	8	2
15	1	3	11	2	20	1	5	11	2
17	1	5	18	2	21	2	2	6	2
10	1	6	24	1	10	1	7	24	2
17	2	2	10	1	18	1	7	19	2
13	2	7	31	1	22	2	3	7	1
9	2	4	21	1	13	2	2	16	1
7	2	3	17	1	12	2	6	22	1
9	1	5	19	1	17	2	6	16	1
6	2	4	20	1	21	1	3	13	1
19	1	9	38	1	13	2	2	16	2
19	2	4	21	2	12	2	2	13	1
18	1	5	29	1	17	1	5	31	2
18	2	4	11	2	11	2	6	46	1
11	1	5	13	2	19	1	5	26	1
12	2	4	17	1	17	2	3	16	1
15	2	3	14	2	16	1	4	16	1

= = =