Sample-size implications for population-based cluster surveys of nutritional status¹⁻⁴

Joanne Katz

ABSTRACT  We studied the design effects for population-based surveys that estimated the prevalence of wasting and stunting malnutrition in Malawi, Zambia, Indonesia, and Nepal, and studied the magnitude of different types of malnutrition clustering within villages. Weight, height, and midupper-arm circumference were measured on all children or on systematic samples of children in randomly selected villages. Design effects ranged from 0.53 for low height-for-age in Zambia to 6.12 for low weight-for-age in Nepal. If all sampled clusters were of size 30, as is often the case for nutrition surveys, design effects would have ranged from 0.44 for low height-for-age in Zambia to 2.59 for low midupper-arm circumference in Zambia. Malnutrition did cluster within villages. Stunting malnutrition clustered less than did wasting malnutrition. Nutrition surveys using clusters of 30 can sample fewer clusters than currently recommended if basic prevalence and cluster information are available prior to sample selection.

AM J Clin Nutr 1995;61:155–60

KEY WORDS  Malnutrition, survey sampling, design effect

Introduction

Many surveys that assess the prevalence of malnutrition use a cluster sampling design in which villages or neighborhoods rather than individuals are selected for nutritional assessment at random or proportionate to population size. This is often done for reasons of costs and logistics. For example, sampling frames listing all villages in a region are usually easier to obtain than are those listing all individuals. Similarly, it is not time- or cost-effective to visit widely dispersed or inaccessible villages only to survey one or two selected individuals within the village. The design effect is the amount by which the variance of the prevalence of malnutrition is increased when villages rather than children are sampled. The variance is increased with cluster sampling when the prevalence estimate from each village differs from one another by more than chance alone.

One way in which to estimate the prevalence of malnutrition with a specified degree of precision in a cluster sample design is to increase the sample size from that required for a simple random sample. Cluster sampling has been used in several nutrition surveys, including ones that use the Centers for Disease Control and Prevention's recommended cluster sampling protocol for such surveys (1–4). Studies in which design effects have been reported are often not helpful in planning future surveys if the number of individuals in each cluster is to be different from the previously reported surveys. A measure of malnutrition clustering that does not depend on cluster size would be more useful for sample size estimation. The clustering of malnutrition in villages may also be different for different types of malnutrition. For example, wasting may be differently distributed across villages than is stunting since each indicates a different mechanism by which such nutritional deficits are acquired.

We used data from four nutritional-assessment surveys conducted in Asia and Africa to estimate the amount of malnutrition clustering within villages and the resulting design effects for each survey. The magnitude of clustering and the design effects were assessed separately for different types of malnutrition. These data provide guidelines for sample size estimation when nutritional-assessment surveys that use a cluster sampling design are being planned.

Subjects and methods

Data from four nutritional-assessment surveys conducted in Malawi, Zambia, Indonesia, and Nepal were used to estimate the design effects for different types of malnutrition. These assessments were part of larger surveys of ocular disorders and studies of vitamin A supplementation. All studies were re-

¹ From the Department of International Health, School of Hygiene and Public Health, Johns Hopkins University, Baltimore.
² Prepared under Cooperative Agreement DAN-0045 between the Office of Nutrition of the United States Agency for International Development (USAID) and the International Center for Epidemiology and Preventive Ophthalmology (ICEPO), and National Institutes of Health grants S10RR04060 and AI25520.
³ The surveys from which the data in this report were obtained were collaborative projects between ICEPO and the national partners in each country, supported by the Office of Nutrition, USAID (except where noted). Malawi: The Ministry of Health, Helen Keller International (HKI), and the International Eye Foundation. Zambia: The National Food and Nutrition Commission, Tropical Disease Research Centre, the Flying Doctor Service, and the Ministry of Health (supported by the International Development Research Center/Canada). Indonesia: The Directorate of Nutrition, Department of Health, and HKI. Nepal: The National Society for the Prevention of Blindness (Nepal Netra Jyoti Sangh).
⁴ Address reprint requests to J Katz, Johns Hopkins School of Hygiene and Public Health, Room 5515, 615 North Wolfe Street, Baltimore, MD 21205-2103.

Received July 20, 1993.
Accepted for publication July 18, 1994.

viewed and approved by the internal review board of the Johns Hopkins University and the within-country institutions participating in the studies. A random stratified sample of villages was selected for inclusion in each study. Children within an eligible age range were identified through a house-to-house census. In Malawi and Zambia, a systematic 10% sample of children in each village was selected for anthropometric assessment. In Indonesia, a 15% systematic sample of children was selected from each village. In Nepal, all children within selected wards were eligible for anthropometry. Eligible children were invited to a central site in each village, where anthropometric measurements were taken. Free and informed consent was obtained from all parents or guardians of children for participation in the surveys. Details of each study are provided below.

Malawi

An ocular disease survey was conducted between October and November 1983 among children aged <6 y in the Lower Shire River Valley of southern Malawi (5-7). This is a poor area where subsistence farming is the primary activity. The survey was conducted during the dry season and took place prior to the large influx of refugees from Mozambique that occurred in the latter half of the decade. A systematic cluster sample of 50 villages was selected from a geographically (north to south) ordered list of villages in Chikwawa and Nsanje districts, excluding urban centers and villages with altitudes >76 m. A total of 5441 children aged <6 y were enumerated in a house-to-house census. In a systematic 10% subsample of each village, weight and height (or length for children aged <24 m) measurements were obtained for 552 children.

Zambia

Between August and December 1985, an ocular disease survey was conducted among children aged <6 y in the Luapula River Valley of Zambia (8, 9). Subsistence farming and fishing are the main occupations in this rural area. The study was conducted during the dry season. The sampling frame consisted of 635 villages in three of the five districts of the valley (Nchelenge, Kawambwa, and Mwense), from which a random sample of 110 villages was selected. All children aged <6 y in each village were eligible to participate in the survey. A total of 4316 children were identified and enrolled by a house-to-house census. In a systematic 10% subsample of 532 children in 99 villages, weight, height (or length if aged <24 m), and midupper-arm circumference were obtained. There were 99 rather than 110 villages in the nutritional-assessment survey because some villages with <10 children did not have enough children for a 10% subsample.

Indonesia

A randomized community trial to examine the impact of a bimonthly program of vitamin A capsule distribution on mortality and xerophthalmia was conducted between September 1982 and August 1984 in the Aceh province of northern Sumatra, Indonesia (10, 11). This is a rural, mostly Moslem area identified as having high rates of xerophthalmia in a nationwide survey undertaken during the late 1970s (12). A total of 2048 villages in two districts of Aceh province formed the sampling frame from which 460 villages were systematically selected for inclusion in the study. A total of 28586 children aged <6 y were identified as eligible to participate. A 15% systematic subsample of 4186 children from 454 villages had anthropometric measurements taken at a central site in the village. These measurements included weight, height (or length if the child was aged <24 m), and midupper-arm circumference (13). Only data from the baseline survey were used in this analysis. Anthropometry was done in 454 rather than 460 villages because 6 villages had missing records or too few children for a 15% subsample.

Nepal

A study to assess the impact of vitamin A supplementation on mortality and xerophthalmia was undertaken in the Sagale district of Nepal from September 1989 to December 1991 (14). This is a low-lying rural region of Nepal bordering Bihar, India. Subsistence farming is the main occupation and childhood mortality rates are high. Forty wards (administrative collection of houses) were randomly selected from a list of 261 wards stratified by geographic location and population size. A total of 4764 children in the selected wards had weight, height or length, and midupper-arm circumference measurements taken.

Statistical analysis

All weight and height data were converted to a percentage of the median weight-for-height, height-for-age, and weight-for-age by using the National Center for Health Statistics (NCHS) reference population (15). The Z scores associated with each of these measures were used to establish cutoffs for malnutrition (16). Wasting or acute malnutrition was defined as a weight-for-age <2 SDs of the NCHS reference distribution for children of the same age and sex. Children whose height-for-age was <2 SDs of the NCHS reference distribution were considered to be stunted or chronically malnourished. Weight-for-age reflects a combination of wasting and/or stunting of malnutrition. It is often used to assess nutritional status since it requires only a weight measurement and knowledge of the child’s age. In addition, decreasing weight-for-age as a percent of the median reference has been associated with a marked increase in the risk of mortality (17). A weight-for-age <60% of the median for the NCHS reference population was also used as a definition of malnutrition. This cutoff was chosen because it has been used to identify severely malnourished children for enrollment in feeding programs.

An additional measure of wasting malnutrition was defined as a midupper-arm circumference <11.5 cm in infants and 12.5 cm in children aged ≥1 y. Midupper-arm circumference has also been shown to be a strong predictor of mortality and relatively independent of age beyond infancy in developing countries (18, 19). It has the additional benefit of requiring only a measuring tape and knowledge of whether the child is an infant or not.

Alternating logistic regressions were used to estimate a pairwise odds ratio, which reflects the degree of association of malnutrition between children within villages (20, 21). This statistic gives the odds that a child is malnourished given that a randomly chosen child from the same village is also malnourished, relative to the odds if the randomly chosen child i
not found to be malnourished. When there are no covariates, this odds ratio can be constructed by taking all possible pairs of children from the same villages and counting the number of concordant and discordant pairs. A two-by-two table of these counts can be constructed with the number of discordant pairs split evenly between the upper right and lower left boxes since the ordering of the discordant pair is not relevant in this situation. The pair-wise odds ratio is then calculated in the usual manner for an unmatched analysis. The 95% CI for the pair-wise odds ratio was calculated from the SE obtained by using alternating logistic regression, and accounts for the fact that the same child may be represented in more than one pair. The 95% CI for the prevalence estimates were also obtained from the alternating logistic-regression program. These CIs are appropriately adjusted for the amount of extra variation in prevalence between villages. The design effect is a known function of the pair-wise odds ratio, the prevalence of malnutrition, and the size of the village or cluster (21). The 95% CI for the design effect was calculated by using the SE for the pair-wise odds ratio and the relationship between the pair-wise odds ratio and the design effect as described previously (21).

Alternating logistic regression was fit by using programs written in the statistical package S (22).

Results

The number of children in the nutritional studies ranged from 532 in Zambia to 4764 in Nepal (Table 1). The number of villages sampled for anthropometry was 50 in Malawi, 99 in Zambia, and 454 in Indonesia, and 40 wards in Nepal. The mean number of selected children per village ranged from 5.4 in Zambia to 119.1 in Nepal. The median number of selected children per village ranged from 6 to 94. The mean and median numbers of selected children per village in Nepal was much larger because all children were eligible for anthropometry within selected wards, whereas the other three studies selected a systematic subsample with a random start from each village.

All the nutrition studies except the one in Nepal had an underrepresentation of infants (Table 2). In the Malawi, Zambia, and Indonesia studies, the list from which children were sampled for anthropometry was constructed from children presenting to a central site for measurement. The underrepresentation of infants may have been due to the reluctance of mothers to bring infants to the central site for weight and length measurements. Hence, they would not have been on the central site roster and would not have been available for selection to the subsample. In Nepal, the list from which a systematic sample was drawn was based on the house-to-house census lists that included all children residing in the selected wards regardless of their presence at a central site.

The prevalence of malnutrition varied widely across studies (Table 3). Wasting malnutrition was greatest in Nepal and Indonesia. It was lowest in Zambia, where only 1.9% of children were ≥2 SDs below the NCHS standard. Stunting was much more common than wasting in all studies. Children were most stunted in Zambia and least stunted in Malawi. The prevalence of low weight-for-age ranged from 1.1% in Zambia to 5.0% in Nepal. The prevalence of a midupper-arm circumference <11.5 cm in infants and 12.5 cm in older children was 1.8% in Indonesia and as high as 12.8% in Nepal. Arm circumference was not measured in Malawi.

村村 odds ratios for clustering of malnutrition could be estimated only if there were at least two cases of malnutrition within any of the villages (21). For this reason, only odds ratios

<table>
<thead>
<tr>
<th>Variable</th>
<th>Malawi</th>
<th>Zambia</th>
<th>Indonesia</th>
<th>Nepal</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHZ &lt; -2</td>
<td>4.2 (2.7, 6.4)</td>
<td>1.9 (1.3, 2.7)</td>
<td>6.0 (5.3, 6.8)</td>
<td>8.0 (6.3, 10.1)</td>
</tr>
<tr>
<td>HAZ &lt; -2</td>
<td>36.6 (23.7, 40.7)</td>
<td>63.2 (59.6, 66.6)</td>
<td>46.7 (45.9, 48.6)</td>
<td>50.8 (48.1, 53.4)</td>
</tr>
<tr>
<td>WA &lt; 60%</td>
<td>1.8 (1.0, 3.3)</td>
<td>1.1 (0.2, 5.2)</td>
<td>1.9 (1.5, 2.4)</td>
<td>5.0 (3.9, 6.5)</td>
</tr>
<tr>
<td>MUAC</td>
<td>8.6 (6.2, 11.9)</td>
<td>1.8 (1.4, 2.4)</td>
<td>12.8 (12.1, 13.5)</td>
<td></td>
</tr>
</tbody>
</table>

1 95% CIs in parentheses. NCHS, National Center for Health Statistics (15).
2 Weight-for-height <2 SDs of the NCHS reference.
3 Height-for-age <2 SDs of the NCHS reference.
4 Weight-for-age <60% of the NCHS median.
5 Midupper-arm circumference: <11.5 cm for infants and <12.5 cm for 1–5-y olds.

TABLE 1
Number of villages and children surveyed

<table>
<thead>
<tr>
<th></th>
<th>Malawi</th>
<th>Zambia</th>
<th>Indonesia</th>
<th>Nepal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of children</td>
<td>552</td>
<td>532</td>
<td>4186</td>
<td>4764</td>
</tr>
<tr>
<td>Number of villages</td>
<td>50</td>
<td>99</td>
<td>454</td>
<td>40</td>
</tr>
<tr>
<td>Mean number of children per village</td>
<td>11.0</td>
<td>5.4</td>
<td>9.2</td>
<td>119.1</td>
</tr>
<tr>
<td>Median</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>94</td>
</tr>
<tr>
<td>Range</td>
<td>1–38</td>
<td>1–64</td>
<td>1–44</td>
<td>20–302</td>
</tr>
</tbody>
</table>

TABLE 2
Age distribution of children surveyed

<table>
<thead>
<tr>
<th>Age (y)</th>
<th>Malawi</th>
<th>Zambia</th>
<th>Indonesia</th>
<th>Nepal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Range</td>
<td>0–5</td>
<td>0–5</td>
<td>0–5</td>
<td>0–4</td>
</tr>
<tr>
<td>Total number of children</td>
<td>552</td>
<td>532</td>
<td>4157</td>
<td>4764</td>
</tr>
</tbody>
</table>

1 Age data missing for 29 subjects (0.7%).
for stunting and low midupper-arm circumference could be estimated in Zambia (Table 4). Pair-wise odds ratios for wasting ranged from 1.01 in Indonesia to 1.45 in Nepal. Stunting showed the least association within villages, with odds ratios ranging from 0.92 in Zambia to 1.17 in Indonesia. The largest amount of malnutrition clustering within villages was seen in Indonesia, where the pair-wise odds ratio for midupper-arm circumference was 2.54. This means that the odds of having a low arm circumference for a child in a village where a randomly chosen child from the same village also had a low arm circumference were 2.54 times greater than if the randomly chosen child did not have a low arm circumference. Except for stunting in Zambia, all odds ratios were ≥1.

The design effect for any particular outcome and study can be written as a function of the odds ratio, the prevalence, and the cluster sizes for that outcome and study (23). The design effects and 95% CIs for malnutrition in each study are presented in Table 5. Design effects for malnutrition ranged from 0.53 to 1.80, except in Nepal where design effects ranged from 2.89 to 6.12. The design effect for stunting in Zambia was <1, indicating a negative association within villages. The 95% CI for this design effect did not include one.

Cluster sizes were much larger in Nepal because all children in the selected villages underwent anthropometric evaluation. This resulted in design effects for malnutrition in Nepal that were much larger than the design effects for the other three studies. The strong dependence of design effect on cluster size is most evident in this example. Table 6 provides the design effects and 95% CIs for hypothetical malnutrition surveys in which clusters of exactly 30 children are selected in each village but the prevalence and odds ratios are the same as those found in the four studies. The design effects would have ranged from 0.44 for stunting to 2.59 for low arm circumference in Zambia. Hence the design effects in Nepal would have been lower or comparable to those of the other studies after adjustment for cluster size.

Design effects for a range of odds ratios and malnutrition prevalence are provided for surveys with clusters of size 30 (Table 7). Design effects are generally <2 when the prevalence of malnutrition is <10% and odds ratios are ≥1.5. The prevalence of wasting malnutrition is usually <10%. However, stunting is often much more prevalent, resulting in a larger design effect than does wasting if both types of malnutrition have the same odds ratio. The largest odds ratio for stunting observed in the four studies was 1.17, resulting in a design effect of 2.13 for a survey with clusters of size 30.

Discussion

The degree to which malnutrition clustered within villages varied by study and by type of malnutrition. Low height-for-age clustered less than low weight-for-height. Low arm circumference clustered more than the other indicators of malnutrition examined in this analysis. One explanation for these differences might be the acute nature of wasting malnutrition, which is more likely to track with infection than stunting, which represents a longer more chronic deficiency. Stunting clustered less than infectious outcomes such as diarrhea, respiratory infections, and measles (21, 24). Another explanation might be that wasting malnutrition varies more with season than does stunting. If villages were examined at different seasons, village clustering of wasting might be greater than that of stunting. This might apply to the Indonesian data, since they were collected over a 1-year period, but not to the other three studies for which data were collected during the same season.
TABLE 6
Village-level design effects for malnutrition if all villages had exactly 30 children

<table>
<thead>
<tr>
<th>Variable</th>
<th>Malawi</th>
<th>Zambia</th>
<th>Indonesia</th>
<th>Nepal</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHZ &lt; -2 (^2)</td>
<td>1.23 (0.74, 2.83)</td>
<td>—</td>
<td>1.02 (0.79, 1.36)</td>
<td>1.90 (1.27, 2.89)</td>
</tr>
<tr>
<td>HAZ &lt; -2 (^3)</td>
<td>1.00 (0.19, 1.86)</td>
<td>0.44 (0.01, 0.94)</td>
<td>2.13 (1.35, 2.73)</td>
<td>1.42 (1.21, 1.69)</td>
</tr>
<tr>
<td>WA &lt; 60(^4%)</td>
<td>1.00 (0.78, 3.55)</td>
<td>—</td>
<td>1.28 (1.01, 1.83)</td>
<td>1.62 (1.20, 2.30)</td>
</tr>
<tr>
<td>Low MUAC(^5)</td>
<td>2.59 (0.90, 5.36)</td>
<td>—</td>
<td>1.75 (1.28, 2.66)</td>
<td>1.38 (1.21, 1.49)</td>
</tr>
</tbody>
</table>

\(^1\) 95% CIs in parentheses. NCHS, National Center for Health Statistics (15).
\(^2\) Weight-for-height <2 SDs of the NCHS reference.
\(^3\) Height-for-age <2 SDs of the NCHS reference.
\(^4\) Weight-for-age <60% of the NCHS median.
\(^5\) Midupper-arm circumference: <11.5 cm for infants and <12.5 cm for 1-5-y olds.

The Z scores were applied over a wide age range (0–5 y of age). These Z scores reflect a standardization of the raw data by using age- and sex-specific norms. Variation in Z scores with age is likely to increase the amount of clustering of malnutrition if the survey focuses on children within a narrower age range. The odds ratios for the surveys presented in this paper are constructed from pairs of children, some of whom are the same age and some of whom are of different ages. If the age range of the survey is narrowed, all possible pairs of children will be of similar ages. These children will be more alike in their nutritional status than will children of different ages if the Z scores do differ by age. Hence, the odds ratios may be larger when the age range surveyed is narrowed. If age is added as a continuous covariate in the logistic-regression model, the household but not the village odds ratios are increased (21, 24).

The design effects obtained from pair-wise odds ratios are comparable to those obtained by using “classical” methods (25). As demonstrated in these studies, the design effect is dependent on the clusters sizes. In surveys that select villages and assess all children within selected villages, the cluster sizes will vary by village and from one survey to another. Hence the odds ratio, which is not dependent on cluster size, is a more useful statistic for assessing the magnitude of the clustering of malnutrition than is the design effect, especially when comparing clustering across surveys with different cluster sizes. In sample-size estimation, once a sample has been selected the approximate cluster sizes are usually known. A guess at the prevalence of malnutrition used to calculate sample size in the simple random sampling case can then be used along with the cluster sizes to obtain an estimate of the design effect.

Given the prevalence of malnutrition and the extent to which malnutrition clustered within villages in these studies, the design effects for nutritional-assessment surveys would rarely exceed two if exactly 30 children were sampled per cluster. The upper limit of the 95% CIs for these design effects would rarely exceed three. The Centers for Disease Control and Prevention’s sampling strategy suggests sampling 30 children from each of 30 clusters. This sample size is based on an assumed design effect of two and would seem to be a reasonable assumption given our data. However, the Centers for Disease Control and Prevention’s strategy is usually used to conduct surveys over short time periods, often for nutritional assessment during famine, whereas our surveys were conducted over longer time periods. Another difference is the potential underestimation of the odds ratios in our studies relative to what might be observed in the Centers for Disease Control and Prevention’s rapid nutrition surveys. Our systematic sampling of a fraction of all village children resulted in the selection of very few children from the same household. This is similar to what occurs in famine situations (one eligible child per household) but different from the rapid nutrition surveys in which there is often more than one eligible child. Furthermore, selected households tended to be geographically spread across each village in all four studies. The Centers for Disease Control and Prevention’s sampling strategy results in the selection of households within close proximity of each other. In addition, all children within the eligible age range who reside in the same household are included in the survey. If malnutrition clusters more within households and more within neighborhoods than within villages, the odds ratios reported in our studies will be lower than those obtained by using the Centers for Disease Control and Prevention’s sampling design. This would result in larger design effects than reported here for clusters of size 30.

A reduction in the design effect is possible if certain village-specific characteristics are strongly associated with malnutrition. If these are known in advance, the sample can be stratified

TABLE 7
Design effects for varying odds ratios and prevalence of malnutrition for clusters of size 30

<table>
<thead>
<tr>
<th>Odds Ratio</th>
<th>2%</th>
<th>4%</th>
<th>6%</th>
<th>8%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1.25</td>
<td>1.1</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
<td>2.1</td>
<td>2.4</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>1.50</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.0</td>
<td>2.2</td>
<td>2.7</td>
<td>3.0</td>
<td>3.5</td>
<td>3.9</td>
</tr>
<tr>
<td>1.75</td>
<td>1.4</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>2.7</td>
<td>3.3</td>
<td>3.8</td>
<td>4.5</td>
<td>4.9</td>
</tr>
<tr>
<td>2.00</td>
<td>1.6</td>
<td>2.0</td>
<td>2.5</td>
<td>2.9</td>
<td>3.2</td>
<td>4.0</td>
<td>4.6</td>
<td>5.4</td>
<td>5.8</td>
</tr>
<tr>
<td>2.25</td>
<td>1.7</td>
<td>2.3</td>
<td>2.8</td>
<td>3.3</td>
<td>3.7</td>
<td>5.2</td>
<td>6.2</td>
<td>6.6</td>
<td>6.8</td>
</tr>
<tr>
<td>2.50</td>
<td>1.8</td>
<td>2.5</td>
<td>3.1</td>
<td>3.6</td>
<td>4.1</td>
<td>5.8</td>
<td>6.8</td>
<td>7.4</td>
<td>7.5</td>
</tr>
</tbody>
</table>
y one or two important characteristics and this would reduce the design effect relative to an unstratified cluster sample. However, such characteristics are rarely identified or available in advance of a survey. Cluster sampling designs such as those described here and the ones based on the Expanded Program on Immunizations coverage surveys do not advocate stratification. Hence, the design effects for clusters of 30 reported here are likely to represent the sampling plans of cluster nutrition surveys that follow the Centers for Disease Control and Prevention's strategy.

Malnutrition did cluster within villages for most of the different types of malnutrition and across most of the surveys in which this phenomenon was examined. Hence an increased sample size is likely to be required for cluster surveys of malnutrition. Design effects with clusters of size 30 were generally below two given the magnitude of clustering and prevalence of malnutrition present in these studies. Ultimately, the purpose of the survey and the logistics of its conduct dictate the sampling strategy. The odds ratios reported here can be used to calculate sample sizes for future nutrition surveys in which different cluster sizes are used (21). Design effects for future cluster surveys can be estimated by using these odds ratios, the projected cluster sizes, and expected prevalence of malnutrition, and will produce more accurate sample size estimates than those that assume a design effect of two.

I gratefully acknowledge Alfred Sommer, Keith West, and James Tielsch for their generosity in making these data sets available. I also thank Scott Zeger for his statistical advice.

References


