

A Not Quite as Quick but Much Cleaner Alternative to the Expanded Programme on Immunization (EPI) Cluster Survey Design

ANTHONY G TURNER,* ROBERT J MAGNANI** AND MUHAMMAD SHUAIB†

Turner A G (United Nations Statistical Division), Magnani R J and Shuaib M. A not quite as quick but much cleaner alternative to the Expanded Programme on Immunization (EPI) Cluster Survey design. *International Journal of Epidemiology* 1996; 25: 198–203.

Background. Although the Expanded Programme on Immunization (EPI) cluster survey methodology has been successfully used for assessing levels of immunization programme coverage in developing country settings, certain features of the methodology, as it is usually carried out, make it a less-than-optimal choice for large, national surveys and/or surveys with multiple measurement objectives. What is needed is a 'middle ground' between rigorous cluster sampling methods, which are seen as unfeasible for routine use in many developing country settings, and the EPI cluster survey approach. **Methods.** This article suggests some fairly straightforward modifications to the basic EPI cluster survey design that put it on a solid probability footing and render it easily adaptable to differing and/or multiple measurement objectives, without incurring prohibitive costs or adding appreciably to the complexity of survey operations. The proposed modifications concern primarily the manner in which households are chosen at the second stage of sample selection.

Conclusions. Because the modified sampling strategy maintains the scientific rigor of conventional cluster sampling methods while retaining many of the desirable features of the EPI survey methodology, the methodology is likely to be a preferred 'middle ground' survey design, relevant for many applications, particularly surveys designed to monitor multiple health indicators over time. The fieldwork burden in the modified design is only marginally higher than in EPI cluster surveys, and considerably lower than in conventional cluster surveys.

Keywords: cluster sampling, modified EPI cluster survey

The important role played by the World Health Organization's EPI (Expanded Programme on Immunization) Cluster Survey in the success of national immunization programme efforts in many countries is widely recognized. The programme monitoring capability provided through the conduct of periodic cluster surveys has been especially important in developing country settings, where administrative records are often incomplete.

In view of its relative simplicity and low cost, the basic survey methodology has in recent years been

extended to health-related surveys with different and often multiple measurement objectives; for example, the measurement of coverage levels for other public health programmes, mortality rates, disease prevalence, health-related knowledge and practices, and utilization of health services. As the scope of surveys using the EPI Cluster Survey methodology has grown, some of the limitations of the standard design have become increasingly apparent. Many of the difficulties encountered could be resolved through the use of conventional multistage cluster sampling techniques, but such methods are viewed as being logistically unfeasible in many situations in which the EPI sampling methodology is applied.

This article addresses a specific question related to the use of cluster sampling in health interview surveys—can the EPI sampling approach be modified to make it sufficiently rigorous for use in large, national surveys and/or surveys with multiple measurement objectives and at the same time retain its advantages

* United Nations Statistical Division.

** Tulane University Medical Center, School of Public Health and Tropical Medicine, 1440 Canal Street, Suite 2200, New Orleans, LA 70112, USA.

† Dhaka University, Institute of Statistical Research and Training.

Reprint requests to: Robert J Magnani.

The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the institutions with which they are affiliated.

of simplicity and low cost? In other words, is there a 'middle ground' between the EPI design and the more rigorous, but often cumbersome, conventional multi-stage cluster sampling approach?

The Standard EPI Cluster Survey Design

The sample design for the EPI Cluster Survey is a two-stage design involving the selection of 30 primary sampling units or 'clusters' (usually villages or other area units), from which 210 children within a target age range (usually 12–23 months) are chosen, seven children per cluster. The sample size of 210 children (per domain or stratum) is mandated by the desire to estimate the level of immunization coverage to within ± 10 percentage points of the true population proportion with 95% statistical confidence, assuming a design effect (i.e. *deff*) of 2.0. Based upon prior experience with immunization coverage surveys (primarily in the US), 30 clusters are generally thought to be necessary to yield sufficiently reliable estimates.¹

In the standard design, clusters are chosen from a list of primary sampling units (e.g. villages, urban communities, census enumeration areas, etc.) through systematic random sampling with probability proportional to estimated size (*ppes*). The latest estimates of cluster population sizes, which are assumed to be proportional to the number of children in the target age group in each cluster, are typically used as measures of size. The 30 clusters so chosen are then visited by survey field staff who carry out the second stage of sample selection and conduct the household interviews.

The original EPI design called for sample children to be chosen randomly from a list of all eligible children in each sample cluster.^{1,2} However, because the creation of lists of households and children tends to be time consuming, costly, and unfeasible in some settings, this procedure is only infrequently used in actual practice. Instead, one of several simplified second-stage sampling procedures is commonly used. In one variant, children are selected by first choosing a random direction from a central location in a village or community (e.g. by spinning a bottle). The number of households in that direction to the edge of the community is then counted, and one household is randomly chosen to be the first sample household. Subsequent households are chosen by visiting the nearest neighbouring households until information has been gathered on seven children. In a yet simpler variant, a direction from a central starting point is randomly chosen as described above and households are contacted as the interviewer moves in the chosen direction until the required information has been gathered for seven children.

The second-stage sampling methods described above are quota sampling procedures, and some of the problems resulting from the use of this approach have been noted over the years.^{3–6} First, quota sampling does not ensure that every eligible member of the target population has a known, non-zero chance of being selected. Hence, the standard EPI design, as it is usually applied, is not a true probability sample design. Since the theory of statistical inference does not apply to non-probability samples, the use of the EPI design precludes, strictly speaking, the calculation of standard errors based upon the sample data themselves. As a result, it is formally not possible to construct confidence limits around survey estimates or assess the extent to which observed changes in immunization coverage levels may be due to sampling error as opposed to actual changes, although this limitation is usually overlooked.

A second problem concerns sampling weights. Where the measures of size for clusters in the sampling frame used to choose sample clusters are accurate (i.e. are equal to the actual cluster sizes), it can be shown mathematically that a two-stage cluster sample in which clusters are chosen at the first stage of sample selection with probability proportional to size and a fixed, equal number of sampling units chosen in each sample cluster will result in a self-weighting sample; that is, a sample in which all sample children have the identical overall probability of selection.⁷ However, since measures of size in sampling frames are often inaccurate due to census errors and changes in population size since the census was taken, application of the standard EPI Cluster Survey method does not automatically result in a self-weighting sample. The survey data must be weighted in order to yield unbiased estimates. In conventional probability samples, this problem is addressed by applying sampling weights, which are equal to the reciprocals of the overall probabilities of selection, to the survey data. However, since selection probabilities are not known in most EPI Cluster Survey applications, sampling weights cannot be calculated.

Thirdly, a computer simulation study demonstrates that the EPI Cluster Survey based upon quota sampling at the second stage of sample selection is considerably more prone to sampling bias than conventional cluster sampling, particularly where immunized children are 'pocketed' within clusters.⁸

Finally, there is the issue of how second-stage sample selection should proceed in surveys with multiple measurement objectives. Consider, for example, a survey designed to measure the following parameters: (1) the proportion of one year old children receiving measles vaccine (a proxy measure for fully-immunized children), (2) the proportion of <5 year old children receiving

oral rehydration solution (ORS) in connection with a recent episode of diarrhoea (e.g. in the 2 weeks prior to the survey), (3) the proportion of children 7–59 months of age receiving vitamin A capsules, and (4) the proportion of households having access to safe (tube well) drinking water. The issue here is not sample size requirements, as these may be calculated in the same manner as was done for immunization coverage in the EPI design, but rather the fact that each of these measurement objectives has a different reference population. Contacting households until seven children aged 12–23 months have been found will suffice for some, but not all, measurement objectives. Of the objectives in the hypothetical example, finding a sufficient number of children <5 years of age who experienced a diarrhoeal episode in the 2 weeks prior to the survey will usually require contacting many more households than the number required to locate seven children aged 12–23 months.

One solution would be to continue contacting households in each cluster until the required number of diarrhoeal cases have been identified, but this would entail visiting widely varying numbers of households per cluster. The field work would thus be considerably more unwieldy to manage and control; moreover, the 'solution' still results in a quota sample, with the attendant problems noted above.

Thus, while the standard EPI Cluster Survey provides a low-cost means for periodically monitoring the coverage of immunization programmes, methodological limitations make it a less than optimal choice for large-scale and/or multi-purpose surveys.

Proposed Design Modifications

In order to put the standard EPI design on a more solid probability footing, it is proposed that the quota sampling procedure often used in the second stage of sample selection in the EPI Cluster Survey methodology be replaced with a procedure consisting of four steps: (1) sketch-mapping sample clusters, (2) creating subclusters or 'segments' of approximately equal size, (3) selecting one segment at random, and (4) interviewing *all* eligible members of the various survey-defined target groups within *all* households in selected segments.

The purpose of the sketch-mapping and segmenting operations is to permit sample households to be chosen in such a way that the probabilities of selection may be calculated, but without having to conduct a preliminary listing operation. Listing is not required because *all* households in the segment chosen are included in the sample.

The creation of segments is accomplished through a field operation entailing making a rough sketch map

of each first-stage cluster showing external and internal boundaries and a rough indication as to the location of dwellings. The cluster is then divided on the map into segments of approximately equal size, the number of which is equal to the measure of size for the first-stage cluster divided by the desired segment size. Minor variations in segment sizes in terms of numbers of dwellings or households may be tolerated, but segments should be as equal in size as possible.

Sampling probabilities under the proposed design are calculated as follows. If M_i equals the measure of size for the i th cluster (the population figure in the sampling frame), M is the total measure of size (i.e. the sum of M_i across all clusters), m is the number of first-stage clusters to be chosen in the survey, and S_i is the number of segments created in the i th cluster, the probability of selection (P_i) of a household in the i th cluster is:

$$P_i = (m * M_i / M) * (1 / S_i)$$

Note that since $S_i = (M_i / C)$, where C is a constant equal to the predetermined segment size, then

$$P_i = (m * C / M)$$

Therefore, all households in the sample have the same probability of selection and hence the same sampling weight (provided that first-stage units are selected with probability proportionate to size and differential sampling rates by strata are not used).

The strategy described has recently been applied in a large national survey in Bangladesh, the Multiple Indicator Survey, and that experience is used here to illustrate the proposed alternative sampling method. In this survey, first-stage sampling units, defined as mauzas in rural areas and mahallahs in urban areas (mauzas and mahallahs are administrative subdivisions) were selected via systematic *ppes* sampling using counts of households from the 1991 Population Census as measures of size. Each selected mauza or mahallah was then sketch-mapped and segmented in the field. The number of segments created in each selected mauza was equal to the census count of households in the mauza/mahallah divided by 40 (the target segment size), and rounded to the nearest integer. Since each mauza or mahallah contains of the order of 250–300 households (although some are larger), each cluster contained on average six to eight segments, one of which was randomly chosen in the second stage of sample selection.

In order to implement the above strategy, a second departure from the EPI sampling strategy must be made. Whereas in the standard EPI design the number of

households contacted per cluster is determined independently in each cluster by visiting households until the desired number of subjects of specified characteristics are identified (e.g. seven children aged 12–23 months), in the modified sampling method this number has to be determined in advance of survey fieldwork.

The steps involved are relatively straightforward. First, total sample size requirements for each of the variables of priority interest are calculated using standard formulae.⁹ Secondly, the most recent population data available and, where relevant, estimates of the incidence of diseases/conditions of interest (e.g. diarrhoeal incidence among <5 year olds) are used to estimate how many households would have to be contacted in order to locate the required number of subjects for each priority variable. The *largest* of these requirements (in terms of numbers of households) is taken as the minimum sample size for the survey. Finally, the target number of households per segment is obtained by dividing the total number of households required by the number of segments (which is equal to the number of clusters) to be covered in the survey.

In the case of programme coverage indicators such as the immunization coverage rate, only information on the number of subjects in the target age range and the total number of households in the target population for the survey is needed in order to estimate the number of households that would have to be contacted. For indicators that are dependent upon disease incidence (e.g. diarrhoeal disease or respiratory infection case management indicators), additional information is needed in order to provide an estimate of the proportion of the population in a target age range that would be expected to have had an episode of the disease/condition of interest within a given reference period prior to the survey; for example, the proportion of children <5 years of age having had a diarrhoeal episode in the previous 2 weeks. Specific procedures for these computations in the case of diarrhoeal disease (which may also be applied to other diseases/conditions) are described elsewhere.⁸ Note that if variables that are dependent upon disease/condition incidence are to be measured in a survey, these will nearly always require larger sample sizes than other types of variables, and thus it is necessary to calculate sample size requirements only for these variables.

Once a target total sample size has been determined, the number of clusters to be covered and a target segment size needs to be fixed. Note that since the total sample size is the product of the number of clusters covered times the average 'take' per cluster, a decision on either the target number of clusters or target segment size will determine the other parameter.

With regard to number of clusters (or segments) to be chosen, while 30 clusters has become a widely-used standard, there is nothing magical about this number. Instead, the general rule is that as many clusters as can be accommodated within existing resource constraints should be covered in a given survey, as larger numbers of clusters help to minimize cluster sampling design effects, and thus increase the precision of survey estimates.

If a main purpose of the 'expanded' survey is to provide national estimates, then a minimum of 100 clusters should be considered. For example, if the overall sample size is determined to be 6000 households and the segment size is set at 35 households, then the number of clusters (first-stage units) would be $6000/35 = 171$; on the other hand if the sample size is 3500 households and the segment size set at 50 households, this would result in only 70 clusters, in which case it would be better in terms of reliability to reduce the cluster size to 35 and select 100 clusters instead.

The target segment size chosen is thus a compromise among two major considerations. First, in order to simplify the mapping work that has to be done and minimize the need to create artificial boundaries in delineating segments in the field, it is desirable that segments be as large as possible. On the other hand, larger segments result in larger design effects, which reduce survey precision. Experience from large-scale demographic and health surveys over the past 20 years suggests that segment sizes in the range of 20–40 households should be adequate for most health-related measurement objectives.

These procedures are again illustrated using the Bangladesh Multiple Indicator Survey. The variable that has the greatest effect on sample size from among those to be measured in the survey was the proportion of one year old children receiving measles vaccination. This is because the target population for this variable, children 12–23 months of age, is quite small, and thus a comparatively large number of households must be contacted in order to locate the required number of children.

Sample size requirements were calculated to achieve a compromise between national-level and district-level measurement objectives. At the district level, it was determined that 130 children aged 12–23 months would be needed (per district). Data from the 1991 Census of Population indicated that one child in the target age range is found in every four households, on average. Thus, the target sample size was determined to be about 520 households ($4 * 130$ per district). This was taken to be the maximum number of households needed in the sample for Bangladesh; that is, the number necessary to

estimate the most demanding of the main variables in the survey and which is more than adequate to measure the other key variables of the survey.

Since district-level personnel were to carry out the fieldwork in the Bangladesh survey, it was desired to keep the number of clusters to be covered as small as possible (and thus the target segment size as large as possible). As a result of these considerations, it was decided that the average cluster size should be 40 households. This meant that 13 clusters were selected in each district.

DISCUSSION

The EPI Cluster Survey methodology has been widely and successfully used over the past 20 years or so to monitor national immunization programme coverage in numerous settings. Because of its relative simplicity, it has also been used in recent years in surveys with differing, and sometimes multiple, measurement objectives. However, there are several important limitations of the standard EPI design in such applications. Where time and resources permit, these problems may be overcome through the use of conventional multistage, cluster sampling techniques. In this article, we have proposed a 'middle-ground' between more rigorous, but cumbersome, conventional cluster sampling methods and the somewhat 'quick and dirty' EPI Cluster Survey methodology.

By modifying the procedure used to select households/study subjects at the second stage of sample selection, the standard EPI design can be put on a solid probability footing without incurring prohibitive costs, though it does add marginally to the complexity of survey operations. Instead of the quota sampling technique frequently used in EPI cluster surveys, the modified technique entails sketch-mapping of sample clusters (which are chosen in the same manner as in the EPI Cluster Survey), subdivision of the selected clusters into segments of approximately equal size on sketch maps, random selection of one segment per cluster, and interviewing all households found within the segments chosen. Sample size requirements, in terms of numbers of households that need to be contacted, are also determined in advance in the modified procedure.

The relative merits of the modified cluster design vis-à-vis the EPI Cluster Survey design (as it is usually applied) and conventional cluster survey methods are summarized in Table 1. The rationale for preferring the modified cluster survey methodology in many applications is fairly apparent. The modified cluster design is decidedly sounder statistically than the EPI cluster design, and differs from conventional cluster sampling

TABLE 1 Comparison of alternative cluster sampling approaches on selected design features

Feature	Standard EPI design	Modified cluster design	Conventional cluster sampling
Statistical soundness:			
Probability method	No	Yes	Yes
Risk of bias ^a	High	Low	Low
Relative precision ^b	Unknown ^c	Marginally higher	—
Fieldwork burden:			
Avoidance of listing	Yes	Yes	No
Overall simplicity	High	Moderate	Low
Mapping skills required	Low	Moderate	Moderate
Flexibility, appropriateness:			
For immunization coverage	Yes	Yes	Yes
For other variables	Limited	Yes	Yes

^a Assumes that, under each design, sampling frames are unbiased and the conduct of field procedures is faithful to the design.

^b Using a conventional cluster design with the same number of clusters and overall sample size as the basis of comparison.

^c Since the EPI design as it is usually applied is not a probability sample, standard errors cannot be estimated using conventional procedures. But assuming equal sample sizes and design effects as in a conventional cluster sample, the expected level of precision would be approximately the same as in the latter.

methods largely in that it will produce marginally less precise survey estimates when segment sizes vary significantly. With regard to level of fieldwork burden, the modified design lies between the two alternative approaches, sharing with the EPI cluster design the avoidance of household listing operations, but requiring more extensive mapping work. The modified cluster design also shares with conventional cluster sampling methods the design flexibility to contend adequately with different and multiple measurement objectives.

The demand for rapid, low cost data for use in designing and monitoring public health programme efforts in developing country settings has grown dramatically in recent years. In this environment, conventional cluster sampling methods are often viewed as a luxury that often cannot be afforded. Nevertheless, there is a need to maintain sufficient scientific rigor in the conduct of sample surveys such that the resulting data may be confidently used to make important policy and programme decisions. We believe that the modified cluster design described in this article maintains the scientific rigor of conventional cluster survey designs, while at the same time retaining many of the features that have made the EPI Cluster Survey so popular.

While the proposed procedure will entail somewhat higher costs than are incurred in most EPI Cluster Survey applications, the 'payoff' is a bona fide probability sample which provides more robust survey estimates from which important programmatic decisions may be made.

REFERENCES

- ¹ Lemeshow S, Robinson D. Surveys to measure programme coverage and impact: a review of the methodology used by the expanded programme on immunization. *World Health Stat Q* 1985; **38**: 65-75.
- ² Henderson R H, Sundaresan T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. *Bull World Health Organ* 1982; **60**: 253-60.
- ³ Bennett S. Cluster sampling to assess immunization coverage: a critical appraisal. Proceedings of the Meetings of the International Statistical Institute Conference, Cairo, Egypt, 1993.
- ⁴ Kok P W. Cluster sampling for immunization coverage. *Soc Sci Med* 1986; **22**: 781-83.
- ⁵ Kalton G. *Some Proposed Modifications for the WHO Simplified Cluster Sampling Method for Estimating Immunization Coverage*. New York: UNICEF, 1988.
- ⁶ Bennett S, Woods T, Liyanage W M, Smith D L. A Simplified General Method for Cluster-Sample Surveys in Developing Countries. *World Health Stat Q* 1991; **44**: 98-106.
- ⁷ Kish L. *Survey Sampling*. New York: Wiley, 1965.
- ⁸ Lemeshow S, Tserkonvi A G, Tulloch J L, Dowd J E, Lwanga S K, Keja J. A computer simulation of the EPI survey strategy. *Int J Epidemiol* 1985; **14**: 473-81.
- ⁹ Lemeshow S, Hosmer D W, Klar J, Lwanga S K. *Adequacy of Sample Size in Health Studies*. Chichester, UK: John Wiley & Sons, for the World Health Organization, 1990.

(Revised version received July 1995)