A CONCURRENT COMPARISON OF A WHO-RECOMMENDED 30-CLUSTER SURVEY AND A MODIFIED VERSION OF IT UNDER INDIAN CONDITIONS IN THE ESTIMATION OF IMMUNIZATION COVERAGES

S. Radhakrishna
B.N. Murthy
N.G.K. Nair
R. Ezhil
S. Venkatasubramanian
N. Ramalingam
V. Periannan
R. Ganesan

ABSTRACT

A concurrent comparison of the WHO 30-cluster sample survey method for estimating immunization coverages (DPT, Polio, BCG, Measles) and an Indian modification of it (GOI) was undertaken in five districts in South India. The essential difference between the two methods is the manner in which the first household is selected in the chosen clusters. With the WHO method, it is chosen at random, whereas with the GOI method it is often close to the village centre. Estimates with the required degree of precision, i.e., 95% confidence limits of ±10 percentage points, were provided in 18 (90%) of 20 instances by the WHO method and in 19 (95%) by the GOI method? findings which are in accordance with expectation. The estimated coverages were, however, higher by the GOI method than by the WHO method in two districts, lower in one district, and in the remaining two districts there was no clear pattern. On the average, there was a suggestion that the GOI method yielded slightly higher coverages, but the differences were not statistically significant.

Key words: *Immunization coverages, 30-cluster survey, Indian modification, Potential for Bias.*

From the Institute for Research in Medical Statistics, Mayor V.R. Ramanathan Road, Chetput, Madras 600 031.

Reprint requests: Dr. S. Radhakrishna, Director, Institute for Research in Medical Statistics, Mayor V.R. Ramanathan Road, Chetput, Madras 600 031.

About 15 years ago, the World Health Organization (WHO) recommended a 30-cluster survey method for the rapid estimation of immunization coverage in children(l). The aim of the survey is to provide a quick estimate of the coverages for a large geographical area, such as an administrative district, so that corrective action could be taken if necessary. Briefly, the first step consists of identifying 30 areas for study *(e.g.,* villages) using a probability proportional to size (PPS) linear systematic sampling technique. Next, within each selected area, a household is selected at random and with this as the starting point a cluster of households is visited in a prespecified manner until 7 children aged 12-23 months are identified and their immunisation status is assessed. It is asserted that the findings in these 210 children (30 clusters x 7 children per

cluster) will provide an estimate of the immunization coverage in the community with 95% confidence limits of ± 10 percentage points. This method has found wide scale application and has performed reasonably satisfactorily according to an evaluation conducted in 1982 by Henderson and Sundaresan, who reported thai 86% of 209 sample survey estimates had 95% confidence limits of ± 10 percentage points(2). More recently, Murthy *et al.* have reported that 95% of 60 such estimates in S. India had the required degree of precision(3).

The above method has been modified at the second stage of sampling by the Government of India (GOI) in its usage under Indian conditions(4). Thus, instead of selecting the first household at random within each chosen area, the field worker goes to the village centre *(e.g.,* market, place of worship, school), selects at random one of the paths leading from the centre, counts or estimates the number of houses from the centre to the boundary along that path, selects a random number between 1 and the total number of households, and commences the survey with the corresponding household. This modification of the WHO recommended method is operationally more convenient, and is apparently necessitated by the fact that up-to-date and reliable household sampling frames are not usually available under rural conditions in India.

The important difference between the WHO method and the GOI method is the manner in which the first household is chosen in the selected cluster. With the latter method, households located at the periphery of the village are not likely to be included in the survey unless the village is small. This could result in some bias, as persons living in the outskirts of villages in India are often of poorer socio-economic status (scheduled castes/tribes, backward classes constituting the labor force) and are likely to have poor immunization coverages. To investigate this possibility, a concurrent comparison of the two methods was undertaken in one of the districts in Tamil Nadu in S. India, and as the findings were not conclusive the comparison was subsequently extended to four more districts.

Material and Methods

Design of the Study

In the event of a 30-cluster survey by either method (WHO or GOI) with 7 children per cluster not providing estimates with the required degree of precision (namely 95% confidence limits of ± 10 percentage points), it is possible that a larger cluster size or a large number of clusters may provide the required degree of precision. To allow for this possibility, in the first district studied (A), 60 clusters were identified by PPS circular systematic sampling technique and 12 children aged 12-23 months were assessed in each cluster. Circular systematic sampling was employed instead of linear systematic sampling, so that the same data set could provide valid estimates for a 30-cluster survey and also for a 60-cluster survey.

As the findings in District A were not conclusive, the comparison between the WHO and GOI methods was subsequently extended to four more districts (B, C, D and E). Although the number of children studied in each cluster was retained at 12, the number of clusters cho-

sen was restricted to 30 (and PPS linear systematic sampling employed), because 60 clusters did not appear to be necessary from the findings in District A, and would have involved a substantially larger work load for the field staff.

Districts A and B were studied in 1991, and District C, D and E in 1993.

Survey Procedures

The procedures adopted for the WHO method were as spelt out in the WHO Manual(l) and those for the GOI method were as set out in the GOI Manual(4). The surveys were undertaken by graduate or post-graduate investigators, after they had received intensive training about interviewing techniques and the 30-cluster survey methodology, for one week. AH the investigators spoke the local language (Tamil) fluently and could therefore elicit information accurately by in-depth interrogation of the mother. To avoid the possibility of any differences in efficiency between the investigators affecting the outcome of the comparison, their postings were rotated on a daily basis so that "balance" was maintained and, on the whole, each of them spent the same amount of time on the two methods (WHO and GOI).

Definition of Coverage

A child was defined as having been "immunized", if he/she satisfied the following conditions:

DPT/Polio : First dose was given at any time 6 weeks after birth. Subsequent 2 doses were given with intervals of at least 4 weeks between successive doses, and all 3 doses were administered before the child had completed one year.

BCG : The vaccination was given at any time before 12 months.

Measles : Immunization was undertaken after completion of 9 months but before the completion of 12 months.

Estimates of Coverage and its Sampling Error

An estimate of the coverage is given by the expression $p = \Sigma fi/\Sigma ni$ where n_i is the number of children assessed in i^{th} cluster and f is the number of children that are immunized. The variance of this estimate is, strictly speaking, not determinable since the 30 clusters were identified by systematic sampling with a random start, and not by simple random sampling; however, in practice, it is usually approximated by the

Expression $\frac{\Sigma i(Pi-P)^2}{m\,(m-1)}$ where $pi = fi/ni$ and

m is the number of clusters(5).

The significance of differences between the estimated coverages by the two methods was determined by using a paired t-test, as the first stage sampling units *(e.g.,* villages) were the same for the two methods. For the same reason, the sampling error of the estimates could not be compared directly by the conventional variance ratio test. Instead, the null hypothesis of equality of variances was tested by computing the correlation coefficient between the sum and the difference of the findings in each first stage sampling unit *(e.g.,* village), and testing the correlation coefficient for equality to zero.

Results

District A

Table I compares the findings with

the WHO method and the GOI method for a 30-cluster survey with 7 children per cluster in District A. Two points of interest emerge from this table:

(i) The estimated coverage with the GOI method appears to be slightly higher than that with the WHO method, by 2.7% for DPT, 4.6% for Polio, 3.6% for BCG and 2.8% for Measles; however, none of these differences was statistically significant (p >0.2), and the 95% confidence intervals were all rather wide.

(ii) The value of 1.96 s.e. varied from 5.8 to 10.8 percentage points for the WHO method, and from 6.7 to 8.1 percentage points for the GOI method; the latter method yielded lower values in the case of DPT, Polio and BCG but slightly higher in the case of Measles; however, the contrast was significant only in thecase of BCG (p = 0.01).

The influence of increased cluster size on the difference between the estimated coverages by the GOI and WHO method is shown in *Table II.* With a cluster size of 12 (the highest studied), the GOI estimate was 5.5% higher than the WHO estimate (p = 0.08; 95% CI = -0.9 to 11.9) in the case of DPT; the corresponding excesses were 6.6% for Polio (p = 0.06; 95% CI = -0.4 to 13.6), 4.0% for

TABLE I- *Findings with WHO 30-Ouster y and GOI Method in District A (7 children per Sun cluster)*

Vaccine	Estimated coverage (%)		Difference		1.96 se	
	WHO (a)	GOI (b)	(b-a)	95% CI	WHO	GOI
DPT	55.8	58.5	2.7	-8.4 to 13.8	9.7	7.5
Polio	50.5	55.1	4.6	-7.5 to 16.7	10.7	7.8
BCG	53.4	57.0	3.6	-3.0 to 10.2	10.8	8.1
Measles	14.6	17.4	2.8	-3.8 to 9.4	5.8	6.7

TABLE II– *Difference in Estimated Coverage* by WHO and GOI Method (30-Cluster) in District A for Different Cluster Sizes*

Vaccine	Number of children in each cluster (cluster size)					
	7	8	9	10	11	12
DPT	2.7	5.5	5.8	5.6	5.5	5.5
Polio	4.6	7.2	7.8	7.7	7.0	6.6
BCG	3.6	3.8	3.9	4.9	4.9	4.0
Measles	2.8	3.8	2.2	2.0	2.0	1.6

* Coverage by GOI method minus coverage by WHO method.

BCG (p = 0.09; 95% CI = -0.5 to 8.5), and 1.6% for Measles (p > 0.2; 95% CI = -3.1 to 6.3).

The sampling error of the estimated coverage by the GOI method was lower than the sampling error of the estimated coverage by the WHO method *(i.e.,* the ratio was less than one) for all cluster sizes (7 to 12) in the case of DPT, Polio and BCG, and slightly higher (*i.e.,* ratio is more than one) in the case of Measles. Again, the contrasts were significant or approached significance only in the case of BCG.

Finally, the effect of investigating a larger number of clusters was examined by considering data from all 60 clusters, with 7 to 12 children per cluster *(Table III).* Once again, the estimated coverages were higher with the GOI method in all cases but one, but none of the differences was statistically significant.

In summary, the findings in this district are not conclusive, although they indicate that the coverages may have been overestimated by the GOI method

Districts B,C,D and E

Of the 32 estimated coverages by the WHO and GOI methods in the 4 districts, all but one had the required degree of precision; the lone exception was with the GOI method in District C, and had 95% confidence limits of ± 10.5 percentage points *(Table IV).*

As regards the comparison between the estimated coverages by the two methods, the GOI method yielded a lower coverage than the WHO method for DPT and Polio immunization and a higher coverage for BCG and Measles immunization, in District B. In District C, all the GOI coverages were higher than the corresponding coverages by the WHO method, a position similar to that in District A. On the other hand, in District D, all the GOI coverages were lower. Finally, in District E, the DPT and Poliocoverages by the GOI method were higher than the coverages by the WHO method, while those for BCG and Measles were lower, which is contrary to the findings in District B. Only one of the contrasts (District C, Measles coverage) was statistically significant (p = 0.01).

TABLE III– *Difference in Estimated Coverage* by WHO and GOI methods (60 clusters) and its 95% Confidence Limits, in District A, for Different Cluster Sizes*

Vaccine	Number of children in each cluster (cluster size)					
	7	8	9	10	11	12
DPT	1.4 ± 7.1	4.5 ± 6.7	4.1 ± 6.2	3.3 ± 5.8	3.0 ± 5.4	2.6 ± 4.9
Polio	2.4 ± 7.4	5.1 ± 6.8	4.6 ± 6.4	4.3 ± 5.9	3.9 ± 5.4	3.1 ± 5.0
BCG	3.5 ± 5.2	3.5 ± 5.2	3.4 ± 4.6	3.7 ± 4.1	2.6 ± 4.0	2.5 ± 3.8
Measles	1.4 ± 5.1	2.1 ± 4.7	1.0 ± 4.2	0.4 ± 3.7	0.4 ± 3.5	–0.2 ± 3.3

* Coverage by GOI method minus coverage by WHO method.

TABLE **IV**-*Findings with WHO 3D-Ouster Survey and* GOI *Method in Districts* B, C, D *and* E *(7 children per duster)*

District	Vaccine	Estimated coverage (%) ± 1.96 se			Difference
		WHO (a)	GOI (b)	(b-a)	95% CI -
	DPT	72.5 ± 7.2	70.7 ± 8.2	-1.8	-9.3 to 5.7
	Polio	72.0 ± 7.6	69.2 ± 8.0	-2.8	-10.3 to 4.7
B	BCG	76.9 ± 8.1	80.8 ± 7.2	3.9	-0.5 to 8.3
	Measles	38.0 ± 8.7	41.8 ± 10.0	3.8	-5.1 to 12.7
	DPT	59.0 ± 8.1	62.0 ± 8.1	3.0	-5.0 to 11.8
	Polio	58.6 ± 8.2	60.6 ± 8.1	2.0	-6.1 to 11.1
e	BCG	66.7 ± 9.1	71.3 ± 10.5	4.6	-1.9 to 11.1
	Measles	22.5 ± 7.0	36.7 ± 9.5	14.2	4.6 to 24.8
	DPT	75.6 ± 6.7	74.5 ± 6.2	-1.1	-9.9 to 7.7
	Polio	73.7 ± 7.0	73.1 ± 5.9	-0.6	-9.1 to 7.9
D	BCG	97.1 ± 3.0	92.9 ± 4.9	-4.2	-8.7 to 0.3
	Measles	54.1 ± 9.2	52.2 ± 9.4	-1.9	-13.1 to 8.3
	DPT	72.4 ± 8.6	74.8 ± 8.7	2.4	-4.4 to 9.2
	Polio	72.4 ± 8.6	76.7 ± 8.4	4.7	-2.1 to 11.7
E	BCG	92.2 ± 5.5	88.7 ± 4.7	-3.5	-11.7 to 3.7
	Measles	48.0 ± 8.1	45.3 ± 8.0	-2.7	-10.8 to 5.4

There was no clear pattern in the matter of precision of the estimated coverages by the two methods. Thus, the sampling error of the GOI estimate was lower than the sampling error of the WHO estimate on 8 occasions, and higher on the remaining 8 occasions. Furthermore, none of the contrasts was statistically significant.

Summary of Findings in 5 Districts

Table V summarizes the main findings, *i.e.,* the difference between the estimated coverage by GOI and WHO methods, in the five districts, for the typical 30 cluster survey with 7 children per cluster and also for the larger survey with 12 children per cluster (the findings with 8, 9, 10 and 11 children per cluster were broadly similar to the above, and are not tabulated here). On average, for the 30-cluster survey with 7 children per cluster, the coverage by the GOI method was higher by 1.0% for DPT, 1.5% for Polio, 0.9% for BCG and 3.2% for Measles. (The corresponding figures were 0.9%, 1.3%, -0.2% and 1.4%, respectively for the survey with 12 children per cluster). Thus, there is a suggestion that, on the average, the GOI

TABLE V—*Summary of Findings in 5 Districts with WHO 30-Cluster Survey and GOI Method*

District	Coverage (%) by GOI method minus coverage (%) by WHO method							
	7 children per cluster				12 children per cluster			
	DPT	Polio	BCG	Measles	DPT	Polio	BCG	Measles
A	2.7	4.6	3.6	2.8	5.5	6.6	4.0	1.6
B	–1.8	–2.8	3.9	3.8	–2.9	–3.7	0.6	0.4
C	3.0	2.0	4.6	14.2	–0.6	–0.8	–1.5	7.3
D	–1.1	–0.6	–4.2	–1.9	–2.1	–1.5	–2.8	–1.7
E	2.4	4.3	–3.5	–2.7	4.8	5.7	–1.5	–0.4
Mean	1.0	1.5	0.9	3.2	0.9	1.3	-0.2	1.4

method yielded slightly higher coverages than the WHO method, though none of the differences in the last row was statistically significant.

Discussion

This concurrent comparison of the WHO recommended 30-cluster survey and the Indian variant of it (GOI method) in five districts in S. India suggests that there may have been a slight overestimation of the estimated coverage by the latter method. The differences were, however, usually small and none of them was statistically significant. While this may be reassuring in the context of Tamil Nadu in S. India where all the 5 districts studied were situated, some caution may have to be exercised in making any sweeping generalization. This is because the composition of Indian villages is usually heterogenous, with backward classes and scheduled castes (SC) and scheduled tribes (ST) tending to live towards the periphery, or in distinct pockets away from the village centre. These are the very subjects who are likely to have poorer immunization coverages because of their lack of health awareness. In this context, it has been reported that in the States of Uttar Pradesh, Madhya Pradesh, Rajasthan .and Bihar, the differences in coverage between SC/ST and the rest of the population are of the order of 10% to 25%(6).

In States with a good record of health services *(e.g.,* Tamil Nadu), the above deficiencies may have been largely overcome with more effective immunization campaigns, with the result that coverages are relatively uniform at the village level in most instances. If so, little difference can be expected between the WHO and GOI methods. Further, the number of occasions a cluster in the periphery or in a pocket containing SC/ST or backward classes was selected by (random sampling) the WHO method (for comparison with a cluster selected close to the village centre by the GOI method) may not have been large enough for the potential bias in the GOI method to

exhibit itself when findings from all 30 clusters were pooled together.

Summing up, in view of die potential for bias with the GOI method, it would be preferable to select the first household at random. Lists of households are usually available, although they may not be up-to-date; this, however, was not a serious constraint in our study in the five districts, as what we required was a sampling frame that would be accurate enough to facilitate a random start, from which the cluster could be built up. Even in areas where no lists are available, households can be enumerated and a random selection made, unless the village is very large and the work load threatens to become too high. Indeed, it is only in such extreme situations that the present GOI method should be regarded as inevitable.

Finally, it must be re-emphasized that both the WHO method and the GOI method are meant only to provide estimates of the coverage for a large geographical area *(e.g.,* health unit district), so that appropriate corrective action can be taken in the event of the performance being unsatisfactory. Lemeshow *et al.* have cautioned against the temptation to look at the findings in individual clusters, or indulge in disaggregation of the survey data(7). If the interest is also in assessing relative coverage levels in different sections of the community (SC/ST vs Others, Rural vs Urban), a design has been proposed that requires 300 children(6). On the other hand, if the interest is in determining whether immunization coverage is below a prespecified target level *(e.g.,* 80%) in a specified block *(e.g.,* a health centre with a population of about 5000), and if accurate sampling frames of eligible subjects are available, the use of lot quality assurance sampling techniques could be explored with advantage(5,8).

REFERENCES

1. Training for mid-level managers. Evaluate vaccination coverage. World Health Organization, Geneva, 1979.
2. Henderson RH, Sundaresan T. Cluster sampling to assess immunization coverage: A review of experience with a simplified sampling method. Bull WHO 1982, 60: 253-260.
3. Murthy BN, Radhakrishna S, Nair NGK, Ezhil R, Venkatasubramanian S. Estimation of immunization coverage in children by WHO-cluster survey. Indian J Med Res (A) 1993, 97: 234-238.
4. Universal Immunization Programme. Evaluate vaccination coverages. Government of India, New Delhi, 1987.
5. Lemeshow S, Robinson D. Surveys to measure programme coverage and impact: A review of the methodology used by expanded programme on immunization. World Health StatQ 1985, 38: 65-75.
6. Methodological Investigations on Surveys on Immunization Coverage. Institute for Research in Medical Statistics, New Delhi, 1991.
7. Lemeshow S, Tserkovnyi AG, Tulloch JL, Dowd JE, Lwanga SK, Keja T. A computer simulation of the EPI survey strategy. Int J Epid 1985,14: 473-481.
8. Lonata CF, Stroh G Jr, Black RE. Lot quality assurance sampling in health monitoring. Lancet 1988 1: 122.