SSC Annual Meeting, June 1996
Proceedings of the Survey Methods Section

ESTIMATING THE POPULATION AND CHARACTERISTICS OF HEALTH FACILITIES AND CLIENT POPULATIONS USING A LINKED MULTI-STAGE SAMPLE SURVEY DESIGN

K.K. Singh[1,2] A.O. Tsui[1,3], C.M. Suchindran[1,4], G. Narayana[5]

ABSTRACT

This paper demonstrates the utility of a multi-stage sample survey design that obtains both a total count of health facilities and of the potential client population in an area. The design has been used for a state-level survey conducted in mid-1995 in Uttar Pradesh, India. The design involves a multi-stage, areal cluster sample, wherein the ultimate sampling unit is either an urban block or rural village. All health service delivery points, either self-standing facilities or distribution agents, in or formally assigned to the ultimate sampling unit are mapped, listed and selected. A systematic sample of households is selected, and all resident females meeting predetermined eligibility criteria are interviewed. Sample weights for both facilities and individuals are applied. For facilities, weights are adjusted for multiplicity of ultimate sample units served by selected facilities. For individuals, the weights are adjusted for survey response levels. The survey estimate of the total number of government facilities compares well against the total published counts. Similarly the female client population estimated in the survey compares well with the total enumerated in the 1991 census.

KEY WORDS: Sample survey; Program evaluation; Health services; Developing country.

RÉSUMÉ

Cet article présente un plan de sondage à plusieurs degrés qui permet d'obtenir un compte total des établissements de santé et des clients potentiels dans une région. Ce plan de sondage a été utilisé pour un sondage réalisé au milieu de 1995 pour l'ensemble de l'état de l'Uttar Pradesh en Inde. C'est un plan utilisant un échantillon à plusieurs degrés, aréolaire et en grappes, dans lequel l'unité de sondage fondamentale (USF) est soit un pâté de maisons en milieu urbain ou un village en milieu rural. Tous les points de service de santé, autonomes ou agents représentants situés dans un USF désignés pour desservir l'USF sont repérés, listés et sélectionnés. Un échantillon systématique des ménages est alors choisi, et tous les résidents de sexe féminin qui satisfont les critères d'éligibilité prédéterminés pour le sondage sont interviewées. Deux poids d'échantillonnage sont utilisés soit celui de l'établissement de soins de santé et celui de l'individu échantillonné. Pour les établissements, les coefficients de pondération sont ajustées pour tenir compte de la municipalité des USF desservies. Pour les individus, les poids sont ajustés en fonction du taux de réponse au sondage. L'estimé obtenu à partir du sondage du nombre total d'établissements gouvernementaux se compare très bien avec les totaux officiels. De même, le nombre estimé de clientes se compare bien avec le nombre total recensé en 1991.

MOTS CLÉS: Plan de sondage; programme d'évaluation; services de santé; pays en voie de développement.

1. INTRODUCTION

The evaluation of the impact of health programs on population-level health outcomes often requires knowledge of the number and characteristics of facilities and potential clients. Such information is frequently lacking in developing countries where program record-keeping and vital registration systems

[1] Carolina Population Center, University of North Carolina at Chapel Hill, CB #8120 University Square, Chapel Hill, NC 27516-3997.
[2] Department of Statistics, Faculty of Science, Banaras Hindu University, Varanasi, 221005 India.
[3] Department of Maternal and Child Health, School of Public Health, University of North Carolina at Chapel Hill, CB # 7400 Rosenau Hall, Chapel Hill, NC 27599-7400.
[4] Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill, CB # 7400 Rosenau Hall, Chapel Hill, NC 27599-7400.
[5] The Futures Group International, 1050 17th Street, N.W., Suite 1000, Washington, DC 20036.

tend to be incomplete and poorly maintained. Given the importance of monitoring program performance on a regular basis, the absence of reliable information on service productivity and client characteristics inhibits program improvements. Because financial, personnel and material resources are scarce in developing countries, health information systems are often accorded low priority among initiatives to strengthen the program infrastructure. These constraints have prompted health researchers and evaluators to seek the needed information with sporadic use of other means of data collection.

In order to obtain current information on health status, health service use, service performance, and client needs, programs have resorted to occasional sample surveys, often designed and conducted independently and subareally (Aday, 1991; Ross and McNamara, 1983). Some demographic and health surveys (Macro International, 1996), however, do provide a national profile of population-level health outcomes, such as fertility, child mortality, and nutritional wellbeing. The distinct advantage of a national population sample, insofar as planning health programs, is its ability to measure the attitudes and behaviors of clients as well as non-clients. Program service statistics are limited to actual clients and may not yield the most current or accurate picture of service utilization.

In addition to client behaviors, it is useful to monitor the accessibility and quality of services, but this requires a separate review of service provision at health facilities or related outlets. Efforts in developing countries, like the situation analysis studies (Miller *et al.*, 1991), involve probability surveys of health facilities and can provide a national overview of program performance. However, often they are restricted to reviewing public health programs due to incomplete registration of private health providers, such as private clinics or pharmacies. The lack of complete and accurate registration of private sector service providers prevents probability sample surveys from being used to monitor health care patterns through this sector. Very few surveys in developing countries, therefore, involve public and private health facility samples as well as national samples of client populations.

Constraints on available resources to expand and improve the delivery of health care in developing, as well as developed, countries are increasing. This suggests that a more efficient use of resources available for monitoring and evaluation, particularly through surveys, is a consideration for all concerned. Innovative approaches to sample surveys should be developed to provide health planners and managers with a maximum of information at a minimum of precision loss.

This paper presents results from a multi-stage, cluster sample survey designed to estimate both the population and characteristics of health facilities and target client populations. The cluster sample for the survey, conducted in the large northern Indian state of Uttar Pradesh, is used as a basis for selecting health facilities and households, with subsequent selection of service staff from the facilities and of married women of child-bearing age from the households. The survey was designed to generate independent samples of health facilities, staff, households and client populations for the health services.

The next section of this paper will describe the survey design, its contents and fieldwork procedures as applied in Uttar Pradesh. The following section presents the comparative results on health facilities and population, and the last section will discuss lessons learned for survey design from the Uttar Pradesh application. These lessons will be important specifically for this survey's planned replication in two years but generally informative for other countries that may adopt the linked design.

2. THE PERFORM SURVEY IN UTTAR PRADESH

The PERFORM (Project Evaluation Review For Organizational Resource Management) Survey was designed to measure benchmark indicators for a large family planning project called the Innovations in Family Planning Services (IFPS) project sited in Uttar Pradesh and co-funded by the Government of India and the U.S. Agency for International Development. Uttar Pradesh has a population of over 140 million and by itself would rank as the fifth largest developing country.

Content. Indicator estimates for IFPS are needed at three levels: (1) public and private service delivery points (SDPs), (2) service providers staffing the SDPs or facilities, and (3) client population, represented by women of reproductive age. As IFPS seeks to improve the family planning service environment, it is imperative to obtain measures of indicators at this level but in such a way as to be relatable to the women resident in those environments.

As a result, the PERFORM survey developed seven questionnaires:

(1-2) An urban block and village questionnaire to inventory all potential and actual providers of health services in the sampled village or urban block;

(3) A fixed service delivery point (FSDP) questionnaire to gather information on the staff, services, equipment, supplies, and education and motivation activities at sampled public and private facilities;

(4) A staff questionnaire administered to all FSDP staff involved in family planning services (identified from the FSDP questionnaire) to assess their capabilities and service experiences;

(5) An individual service agent (ISA) questionnaire to all individuals working outside of self-standing facilities (FSDPs) who currently or potentially can provide health planning services, such as private doctors, pharmacists, midwives, lay health workers and retailers;

(6) A household questionnaire to be administered to heads of the sampled households to enumerate household members and selected demographic and social characteristics;

(7) An individual questionnaire for currently married women between the ages of 13 to 49 (identified from the household questionnaire) to collect information on knowledge of and past, current and intended utilization of health services, recent pregnancy and contraceptive behaviors, and additional background characteristics.

Sample design. PERFORM was designed to provide estimates of facility and population characteristics at the state, regional, divisional and district levels. The district was important since it was the focal point for introducing innovative approaches and additional IFPS inputs. At the time of survey design, Uttar Pradesh had 14 administrative divisions; two districts were selected from each using probability proportional to size (PPS) procedures. The districts were also aggregated into five regional groupings.

In each district, the total number of households to be sampled was fixed at 1,500. A sample of 1,500 households per district was determined to be sufficient to provide estimates for the main population-level indicators. Urban blocks and rural villages served as the ultimate sampling units, hereafter referred to as the primary sampling units (PSUs). These areal units have administrative-political boundaries and thus public administration utility. The urban household sample was selected in proportion to the urban population in the district. A minimum of 20 urban blocks in each sample district was selected if the proportion urban was less than 20 percent. This minimum level was prescribed mainly to ensure coverage of a sufficient number of service delivery points. The households from both the rural and urban areas in each sample district were selected using stratified, multi-stage sampling procedures.

Selection of rural villages and households. At first, villages were stratified into four strata according to population size as follows:

Stratum	Population size of the village
I	100 - 499
II	500 - 1,999
III	2,000 - 4,999
IV	5,000 and above

In the second step, villages within each stratum were arranged according to female literacy and the number of villages to be selected from each district was proportionately allocated across the strata. The selection of villages from each stratum was carried out using systematic random sampling.

At the third step, all households were listed and mapped, and a target of 20 households was selected from each village systematically. Villages with more than 500 households or 2,500 population size (*i.e.*, some in stratum III & all in stratum IV) were segmented into four equal parts, and two segments were systematically selected for house listing. The twenty households were then selected from each segment using systematic random sampling. Villages with less than 100 population or 20 households were excluded from the list.

The sampling formula can be given as follows:

Let n_{ij} denote the number of households in the i-th village and j-th stratum. Then the probability of selecting village i from the j-th stratum within a district, p_{ij}, is defined as,

$$p_{ij} = a_j \times \frac{n_{ij}}{N_j}$$

where a_j and N_j are, respectively, the number of villages selected and the total number of households in the j-th stratum.

Let q_{ij} be the probability of selecting a household from the rural areas of a selected district. Then q_{ij} may be defined as

$$q_{ij} = p_{ij} \times \frac{20}{n_{ij}}$$

where 20 is the number of households drawn from the selected village.

The weights for villages and households are then the inverse of their selection probabilities, *i.e.*, $1/p_{ij}$

and $1/q_{ij}$, and are denoted as VW_{1ij} and HW_{1ij} respectively.

Selection of urban blocks and households. At the first stage, all the towns in the district were divided into two strata according to population size as follows:

Stratum	Population size of the town
I	100,000 and more
II	Less than 100,000

At the next stage, towns in the first stratum were self selected. In the stratum II, all the towns were arranged according to population size and selected systemat-ically. At least two blocks were selected from each town using PPS methods. At the third stage, all households in the selected blocks are listed and mapped, and 15 households are selected from each urban block using systematic random sampling.

The sampling formulae are as follows.

Let u_{ij} denote the probability of selecting the i-th urban block from the j-th town. Then u_{ij} may be defined as

$$u_{ij} = b_j \times \frac{x_{ij}}{Y_j}$$

where b_j is the number of urban blocks selected in the j-th town, Y_j is the total number of households in the j-th town, and x_{ij} is the number of households in the i-th block and j-th town.

The probability of selecting a household from an urban block within a district, denoted as v_{ij}, is defined as,

$$v_{ij} = u_{ij} \times \frac{15}{x_{ij}}$$

where 15 is the number of households drawn from the selected urban block.

The weights for urban blocks and households are then the inverse of their selection probabilities, *i.e.*, $1/u_{ij}$ and $1/v_{ij}$, and are denoted as UW_{1ij} and HW_{1ij} respectively. Since the population-level estimates are based on individuals, all individuals in a selected household received the household weight. No selection procedure was used for eligible respondents within a household. The final household weight also included an adjustment to reconstitute the proportion urban in the district, where an oversampling of urban blocks had occurred.

Adjustment for household questionnaire nonresponse. A subsequent step to computing the weights is to adjust for nonresponse to the household questionnaire, which we have done as follows.

Let n_1 be the number of households to be selected and n_2 be the number of households where interviews are completed. Then the adjusted weight for households due to nonresponse is defined as

$$HW_{2ij} = HW_{1ij} \times \frac{n_1}{n_2}$$

Selection of service delivery points in sample districts. To obtain a probability sample of service delivery points, FSDPs and ISAs were selected in relation to the PSUs, *i.e.*, the villages or urban blocks, as follows:

1. All private and public sector health institutions in selected rural and urban PSUs;
2. All subcenters, primary health centers, community health centers, postpartum centers providing services to the population in the selected rural PSUs;
3. All private hospitals with 10 or more beds in the nearest town (with less than 100,000 population) within 30 kms of selected rural PSUs;
4. All municipal hospitals, district hospitals and medical college hospitals;
5. All clinics and hospitals runs by voluntary agencies, the organized sector and cooperatives; and
6. All ISAs in selected villages and urban blocks.

It is probably helpful first to describe the organized delivery of health care through the government sector. Residents of all villages are entitled to obtain health care from a government subcenter (SC), a primary health center (PHC) and a community health center (CHC). Villages with 5,500 population or more often have an SC located within their boundaries. Approximately six SCs will report to one PHC, and PHCs in turn are linked to a CHC. At times the PHC[6] is integrated with the CHC; as a result, our estimation must be of CHCs and PHCs combined, while SCs are estimated separately. All SCs assigned to a sampled village were visited, as were their affiliated PHCs and CHCs.

At the time of listing and mapping households in each urban block and village, the FSDPs and ISAs were also listed and mapped. In addition, key informants in each PSU were interviewed regarding health outlets not visibly obvious. The selection of service delivery points, both FSDPs and ISAs within the PSU boundaries or affiliated with the

[6] Population growth has lead to the establishment of "additional PHCs" and redistricting of the original PHC catchment areas. These additional PHCs have been included in the estimation of the number of PHCs.

government's health subcenter, involved a full census. The one exception to this was for municipal hospitals, district hospitals and medical colleges, which were self-selected and thus had a weight of unity. The selection probabilities of the other FSDPs and ISAs are then a function of the probability of selecting the PSU, and the inverse of the latter serves as the weight of the FSDP or ISA unit. Weights for CHCs, PHCs and SCs were calculated with the procedure below after determining some fieldwork "failure" in selecting these types of facilities correctly. (This failure is discussed later.)

Since CHCs and PHCs are associated with more than one PSU, we have assumed that one PHC exists per 30,000 population (which is approximately the actual average for Uttar Pradesh) and that one SC serves approximately 5,500 (actual district averages range from 4,000 to 6,500). Under this assumption, the CHC/PHC weight for each selected PSU is then

$$W_{CHC/PHC} = \frac{\text{Total population in selected PSU}}{30{,}000} \times VW_{1ij} (\text{or } UW_{1ij})$$

and the SC weight for each selected PSU is

$$W_{sc} = \frac{\text{Total population in selected PSU}}{5{,}500} \times VW_{1ij} (\text{or } UW_{1ij}).$$

All weights for FSDPs that were not self-selected had to be adjusted for multiplicity, *i.e.*, when an FSDP was selected into the sample on the basis of more than one PSU. For example, a CHC/PHC might be selected because of two sampled PSUs. In this case, the weight for the CHC/PHC was the sum of the weights of the two selected PSUs, *i.e.*, $W_{CHC/PHC}$, associated with its selection.

Survey implementation. Fieldwork for the PERFORM Survey was conducted from June to September 1995 in Uttar Pradesh. The survey was executed by four organizations contracted following a competitive procurement process. One organization that had tested the PERFORM survey design in one district a year earlier served as the nodal or coordinating organization. Master training to survey project coordinators and supervisors was provided, including a field pretest. The actual fieldwork for PERFORM was carried out in six-member teams composed of 1 male supervisor, 1 female editor, 1 male interviewer and 4 female interviewers. Each fieldwork organization on average engaged 3 teams to cover one district, or a total of 18 field staff for data collection per district (or 21 teams for a total of 126 field staff to cover 7 districts). Overall field supervision was the responsibility of a specially-appointed four-member team, one assigned to each consulting fieldwork organization. Following field editing, the questionnaires were transported to the home offices of the survey organizations for data entry and cleaning.

3. RESULTS

Table 1 gives the sample coverage for the PERFORM survey, in terms of the number of units selected of each type, the number successfully interviewed, and the completion rate. The completion rates are very high for sample units requiring personal contact, ranging from 94.3 for eligible women to 96.7 percent for households. Interview completion rates were 95 percent for both facilities and agents. Only for fixed facility staff was the rate somewhat lower at 90 percent, a respectable although not as outstanding level[7]. PERFORM's sample of 45,277 individual respondents exceeds the 2,400 required for a simple random sample[8] of the state population to estimate, say, the percentage literate (with an error range of +/-2 points and a 95 percent confidence level). Thus, the precision of most of the population-level estimates should be fairly high.

Population size and characteristics. We compare first population-level measures on selected demographic indicators obtained from other sources with those from the PERFORM survey, as shown in Table 2. The figures indicate that PERFORM results compare favorably with both census measures as well as those from the recent National Family Health Survey (NFHS) conducted in Uttar Pradesh in late 1992 and early 1993, with a sample size of 11,438 ever married women aged 13 to 49. The enumerated population shows a growth of almost 10.5 million persons since the 1991 census, and the percent of households in urban areas is close across all three sources. The ratio of women to men is slightly lower in PERFORM (891) than in the NFHS (917). The percent of the population in the two age groups (0 to 14 and 65 and over) compares well, as does the percent of households belonging to the scheduled castes. The percent of households belonging to scheduled tribes is 3.1, higher than the 1.1 observed in the NFHS. This may reflect an actual growth in such households with increased in-migration to large towns

[7] One type of staffperson, the auxiliary nurse-midwife who is stationed at a subcenter, was difficult to reach, even after the standard three attempts.

[8] Even with a design effect of 1.3 from a cluster design, the required sample would be 3,120.

and cities by scheduled tribe members. The proportions literate show small gains since the NFHS but compare well overall. The total fertility rate and castes. the level of modern contraceptive use also are similar and change in a consistent direction between dates of the two Uttar Pradesh surveys. Table 2's results suggest that PERFORM's sample design, based on traditional multistage cluster sample designs used for demographic surveys, was executed properly to produce state-level results comparable to both the census and earlier NFHS survey.

In Table 3 we compare the age and sex distributions for Uttar Pradesh obtained from the NFHS and PERFORM, as well as from the Sample Registration System, operated by the Office of the Registrar General. The sex ratios for the two surveys are also given. The age-sex distributions are again comparable across the three sources. However, there is a markedly lower sex ratio for the age group 30-49 years (813) in PERFORM and a slightly higher one for ages 50-64 (992) than those in the NFHS (941 and 960 respectively). We suspect some of this difference is due to a "push" of females out of the end of childbearing ages by field investigators of *one* survey organization to avoid completion of the pregnancy calendar and history portions of the questionnaire[9]. As a result, there are somewhat more women in ages 50-64 enumerated in the PERFORM Survey than may actually be the case. This also may mean that births to women who were actually under age 50 were underenumerated. Because this is not a high-fertility age group, the bias is not likely to be large.

Facility size and characteristics. By visiting and interviewing the facilities selected through the PSUs or cluster, we are able to generate an independent sample of health facilities and service providers[10]. The weighted counts of these outlets are shown in Table 4. Our ability to validate the estimates of independent agents is weakened by the fact that many of them are not registered, particularly the "unqualified" (or quack) doctors. Narayana *et al.* (1994: Table 8) report a 1991 total number 112,568 villages in Uttar Pradesh, which would suggest almost one traditional birth attendant per village and 1 anganwandi worker for every 4.5 villages on average. These ratios appear reasonable given known circumstances regarding access to such types of care.

For the FSDPs, however, there are published data on the number of public sector health centers. Using the weighting procedure described earlier, we estimated 3,948 CHCs/PHCs (including additional PHCs) and 20,151 subcentres from the PERFORM sample. The number reported in Narayana *et al.* (Table 19) based on Uttar Pradesh government statistics is 3,985 and 20,153 respectively. The figures are quite close and provide evidence of the utility of the linked cluster sample design.

Table 5 presents the actual and weighted counts of such facilities in each district. The PERFORM sample selected in a total of 633 CHC/PHCs in the 28 districts or 34.8 percent of the total (1,818) and 1,267 or 13.3 percent of the total (9,491) number of subcenters. The actual numbers of CHCs/PHCs and SCs are current for 1995 and have been obtained from the Uttar Pradesh Department of Health and Family Welfare. The estimated figures are based on our original design that assigned PSU weights to the sampled CHCs, PHCs, and SCs.

Overall 1,818 CHCs/PHCs and 9,491 SCs are functioning in the 28 sample districts. Our design estimates 3,472 CHCs/PHCs and 9,495 SCs. The SC figures are nearly identical; and using the villages and urban blocks as PSUs, which coincide with the public administration units used for deciding the location of subcenters, is clearly a reasonable basis for selecting these types of facilities. However, the procedures prove less accurate for the larger health facilities of PHCs and CHCs; and the question is "why". The estimated number of such facilities is proportional to the weight of the PSU that provides the basis for their selection, which in turn is a function of its population size. Hence if CHCs and PHCs associated with residents in (and SCs of) small villages are disproportionately selected, the design will result in a higher-than-actual count of such facilities. This appears to be the case in two districts, Allahabad and Sultanpur. Underestimation of CHCs and PHCs results where the reverse occurs (as in Bareilly district). Because of PPS, larger stratum IV villages have smaller weights and most selected FSDPs in the survey are reported to be in PSUs of this size.

The overall result for the 28 districts is an over-estimation of CHCs/PHCs by 91 percent. If the two most problematic districts of Allahabad and Sultanpur are eliminated, the over-estimation reduces to 22.5 $\pm$ 0.8 percent. Thus, the estimation of these types of facilities is in the acceptable range with the removal of these two districts from the calculations.

The "failure" of the weighting scheme for CHCs/PHCs is a quandary that is investigated further.

---

[9] Upon further investigation, we found the sex ratio for women ages 50-64 to be uniformly higher in the seven districts under one organization's responsibility than those of others.

[10] These include those who currently, as well as potentially, can provide family planning services, *i.e.*, not all the estimated number of retail outlets (general merchant, kirana and pan shops) shown presently dispense contraceptives.

We are able to calculate the expected number of CHCs/PHCs and SCs that should fall within the PSUs based on a priori knowledge that such facilities are located in PSUs of population size at least either 30,000 or 5,500. As the urban block and village questionnaires obtained the total population of the PSU, we use this information to obtain a distribution of the district's population by stratum and divided each stratum by the CHC/PHC or SC catchment size (30,000 and 5,500 respectively). The results are provided in Table 6 and compared against the actual count of facilities reported to be in the PSUs through the survey. The fieldwork organization code (I to IV) is also given. Again, the subcenter figures compare favorably while the CHC/PHC figures deviate quite a bit. If the two districts (Allahabad and Sultanpur) are eliminated, the two counts are more comparable (151 to 137).

It appears that because the selection probability of a CHC/PHC for a given PSU is small, the chances of distortions in their weighted counts are higher than for SCs which have a greater probability of being sampled per PSU. In the case of the two districts, Allahabad and Sultanpur, a higher number of small villages (stratum I) have been selected. These villages' larger weights have disproportionately increased the estimated number of CHCs/PHCs.

Table 1: Coverage of Sample Units of PERFORM Survey: Uttar Pradesh, 1995.

	Sample Units						
	Villages	Urban Block	Households	Eligible Women	Fixed SDPs	FSDP staff	Individual
Sample Coverage							
Number Sampled	1539	738	42,006	48,009	2,549	7,026	23,364
Number Interviewed	1539	738	40,633	45,277	2,428	6,320	22,335
Percent Completed	100.0	100.0	96.7	94.3	95.3	89.9	95.6

Notes: Villages and urban blocks served as the primary sampling units; eligibility criteria for women were currently married and aged 13 to 49 years; SDP=service delivery point.

Table 2: Basic Demographic Indicators for Uttar Pradesh, India.

	Census (1991)	NFHS (1992-93)	PERFORM (1995)
Population	139,112,287	u	149,758,641
Percent Urban	19.8	22.6[a]	21.6[a]
Sex Ratio[x]	879.0	917.0	891.0
Percent 0-14 years old	39.1	41.8	40.6
Percent 65+ years old	3.8	4.8	4.7
Percent Scheduled Caste	21.0	18.0[a]	20.4[a]
Percent Scheduled Tribe	0.2	1.1[a]	3.1[a]
Percent Literate[b]			
Male	55.7	63.5	65.3
Female	25.3	31.4	35.2
Total	41.6	49.9	51.4
Total Fertility Rate	5.1	4.8	4.6
Modern Contraceptive Prevalence	u	18.5[c]	22.1[c]

u=Not available

[a]Based on number of households

[b]Based on population aged 7 and above for the census and population aged 6 and above for NFHS and PERFORM

[c]Percent of currently married women aged 15 to 49 using modern contraceptive methods

[x]Number of females per 1000 males.

Table 3: Percent Distribution of the *De Jure* Population by Age and Sex, Based on SRS, NFHS and PERFORM Sources for 1991-95.

	SRS (1991)		NFHS (1992-93)			PERFORM (1995)		
Age	Male	Female	Male	Female	Sex Ratio	Male	Female	Sex Ratio
0-4	14.4	14.4	14.6	14.6	917	13.3	13.8	908
5-14	24.9	24.4	27.5	26.0	868	27.3	26.8	856
15-29	28.4	26.8	25.1	26.4	967	25.4	26.5	905
30-49	20.7	21.9	19.2	19.7	941	20.0	18.7	813
50-64	8.2	8.5	8.4	8.8	960	8.7	9.9	992
65 +	3.6	4.0	5.2	4.4	718	5.3	4.4	698
Total	100.0	100.0	100.0	100.0		100.0	100.0	

Source for Sample Registration System (SRS): Office of the Registrar General of India (1993a)
Source for NFHS: National Family Health Survey, Uttar Pradesh (1992-93)

Table 4: Total number of estimated public and private sector delivery points by type in Uttar Pradesh, India: 1995.

Fixed Service Delivery Points	Number	Individual Service Agents	Number
Total	31,400	Total	1,099,825
Hospitals		Physicians	
Government Allopathic	968	Private Resident Allopathic	32,182
Government ISM	668	Private Visiting Allopathic	9,011
Municipal Allopathic	57	Private Resident (unqualified)	62,880
Municipal ISM	23	Private Resident ISM	42,343
Private	5,212	Private Visiting ISM	9,138
Private Voluntary	130	Anganwadi Workers	25,994
Private ISM	35	Village Health Workers	65,532
Industrial	61	Traditional Birth Attendants	110,546
Medical Colleges	9	Medical Shops	40,979
CHC/PHC/Add'l PHC	3,948	General Merchants	133,517
Subcenters	20,151	Kirana Shops	376,679
Others	137	Pan Shops	136,353
		Depot Holders	5,818
		Other	48,855

Table 5: Total actual and estimated number of community health centers, primary health centers[a] and subcenters by district in Uttar Pradesh, India: 1995.

	CHC/PHC		Subcenter			CHC/PHC		Subcenter	
Districts	Actual	Estimated	Actual	Estimated	Districts	Actual	Estimated	Actual	Estimated
Aligarh	77	69	399	369	Maharajgang	30	39	195	180
Azamgarh	103	69	475	949	Meerut	76	187	410	119
Almora	44	104	254	468	Mirzapur	64	69	309	302
Allahabad	112	981	594	677	Moradabad	92	81	485	248
Ballia	73	93	357	485	Nainital	53	79	287	344
Banda	89	101	322	302	Rampur	37	19	170	139
Bareilly	71	42	355	162	Saharanpur	60	49	293	388
Dehradun	24	41	139	60	Shahjahanpur	52	59	301	298
Etawah	69	84	323	364	Sultanpur	70	487	394	649
Fatehpur	57	73	309	327	Tehri Garhwal	31	5	159	63
Firozabad	33	34	234	236	Sitapur	87	44	437	450
Gonda	107	183	528	461	Unnao	63	162	344	106
Gorakhpur	59	84	470	460	Varanasi	122	144	616	658
Jhansi	51	77	251	157	Total	1818	3472(±21)	9491	9495(±15)
Kanpur Nagar	12	13	81	74	Total[b]	1636	2004(±13)		

[a]Includes additional primary health centers

[b]Excludes Allahabad and Sultanpur districts

Source for 1995 actual figures from Government of Uttar Pradesh Department of Medical and Family Welfare.

Table 6: Observed and expected number of CHCs/PHCs[a] and subcentres within the PSU by district in Uttar Pradesh, India: 1995.

District	CHC/PHC		Subcenter		Field Work Company
	Observed	Expected	Observed	Expected	
Aligarh	6	5	10	17	II
Azamgarh	3	5	24	15	III
Almora	5	2	14	9	I
Allahabad	19	4	17	18	III
Ballia	9	7	34	27	III
Banda	8	9	19	27	III
Bareilly	5	3	10	16	II
Dehradun	5	7	10	21	I
Etawah	8	7	17	20	II
Fatehpur	9	7	22	25	IV
Firozabad	6	6	28	30	II
Gonda	8	5	15	18	IV
Gorakhpur	5	4	16	20	IV
Jhansi	7	6	16	24	III
Kanpur Nagar	2	2	6	8	II
Maharajgang	4	4	9	13	IV
Meerut	12	8	12	34	II
Mirzapur	7	7	22	22	III
Moradabad	5	5	9	19	I
Nainital	6	4	19	19	I
Rampur	2	5	14	16	I
Saharanpur	6	6	25	21	I
Shahjahanpur	5	3	14	15	II
Sultanpur	16	6	21	15	IV
Tehri Garhwal	1	3	3	10	I
Sitapur	3	6	17	17	IV
Unnao	10	6	9	24	IV
Varanasi	6	5	18	18	III
Total	186	147	450	538	
Total[b]	151	137			

[a]Includes additional primary health centers

[b]Excludes Allahabad and Sultanpur districts

4. DISCUSSION

The cluster-based sample design for generating independent samples of facilities and households, which can be analyzed individually or jointly, does warrant more extensive consideration in data collection efforts for health program research and evaluation in developing countries. Careful design and fieldwork sampling and execution can yield high-quality and acceptably precise survey estimates, as the results in this paper show. The weighted totals, rather than sample totals, themselves are numbers useful to program planners who decide the flow of personnel, material and financial resources to and among various facility sites and area locations. The linkage of facility to individual records offers further important analytic opportunities to assess the relative importance of personal background and service supply factors on health outcomes of interest (*e.g.*, Boyd and Iversen, 1979).

At the same time, our application of this design reveals several lessons. First, there is an obvious need to monitor the survey fieldwork closely with increased on-site data entry so that the apparent "push" of eligible women out of the older age ranges can be prevented. This is difficult to detect through individual questionnaire spotchecks but can be observed in aggregate tabulations produced, say, weekly on completed questionnaires. Second, the excess count of CHCs/PHCs in two districts, where the survey fieldwork involved two *different* organizations suggests that stratum I villages might have been disproportionately selected or that some of the CHCs/PHCs reported to be within the PSU boundaries were in fact not. The former may have occurred as a sampling error since each fieldwork organization was provided with a list of sampled PSUs. Third, the listing and mapping of PSUs for facilities, individual health care providers and households are an important stage of the fieldwork. Careful execution of this task allows the sampled units to be re-located for future follow-up. This is will be an essential measurement effort for evaluating the IFPS project.

Certainly for a survey as complex as PERFORM, scaled to capture the levels of and differentials in the patterns of health service delivery and client utilization in an area as populous as Uttar Pradesh Pradesh, the fact that the quality of the data meets most standards of precision evidences an important fieldwork achievement as well as design innovation.

ACKNOWLEDGMENT

Partial support of this study has been provided by The EVALUATION Project, USAID Contract #DPE-3060-C-00-1054-00. The views contained herein are solely those of the authors and not the sponsoring agency.

REFERENCES

Aday, L. A. (1991). *Designing and Conducting Health Surveys*, San Francisco: Jossey-Bass Publishers.

Boyd, L. H., Jr. and Iverson, G. R. (1979). *Contextual Analysis: Concepts and Statistical Techniques*, Belmont, CA: Wadsworth Publishing Co.

Macro International, Inc. (1996). *Demographic and Health Surveys Newsletter*, 8, 1-12.

Miller, R., Ndhlovu L., Gachara, M., and Fisher, A. (1991). "The situation analysis study of the family planning program in Kenya", *Studies in Family Planning*, 22, 131-143.

Narayana, G., Cross, H.E., and Brown, J.W. (1994). *Family Planning Programs in Uttar Pradesh Issues for Strategy Development: Tables,* Centre for Population and Development Studies, Hyderabad, India.

Ross, J. A. and McNamara, R. (1983). *Survey Analysis for the Guidance of Family Planning Programs,* Liege: International Union for the Scientific Study of Populations.