

## Appendix A. Sample Design

---

The major features of the sample design for the Sierra Leone MICS3 survey are described in this appendix. Sample design features described below include target sample size, sample allocation, sample frame and listing, sampling stages, stratification, and the calculation of sample weights.

Sierra Leone is divided into four provinces: the Western Area and the Northern, Southern and Eastern Provinces. The Western Area is divided into Western Urban and Western Rural districts, each of which is divided into wards and then in turn further divided into enumeration areas (EAs). The remaining three provinces are divided into districts of unequal sizes. Each district is divided into chiefdoms, which are in turn are divided into EAs. EAs contain 100 households on average and may contain several villages within their boundaries. EAs are the basic administrative unit that is used by Statistics Sierra Leone (SSL) for the purpose of conducting censuses or surveys. The 2004 Housing and Population Census divided Sierra Leone into 9,673 EAs.

The primary objective of the sample design for the Sierra Leone MICS3 was to produce statistically reliable estimates of most indicators at the national level, for urban and rural areas, and at the province level. The design of the sample allows the estimation of indicators at district level – however, such estimates are likely to be very imprecise, since the sample size was not determined to enable district-level estimates.

### Sample Size and Sample Allocation

The target sample size for the Sierra Leone MICS was calculated as 8000 households. For the calculation of the sample size, the key indicator that was used was the proportion of children aged 12-23 months who are vaccinated with DPT3. The following formula was used to estimate the required sample size for these indicators:

$$n = \frac{[ 4 (r) (1-r) (f) (1.1) ]}{[ (0.1r)^2 (p) (n_h) ]}$$

where

- $n$  is the required sample size, expressed as number of households;
- 4 is a factor to achieve the 95 percent level of confidence;
- $r$  is the predicted or anticipated prevalence (coverage rate) of the key indicator;
- $f$  is the shortened symbol for *deff* (design effect);
- 1.1 is the factor necessary to raise the sample size by 10 per cent for non-response;
- $0.1r^1$  is the margin of error to be tolerated at the 95 per cent level of confidence, defined as 10 per cent of  $r$  (relative sampling error of  $r$ );

---

<sup>1</sup> The margin of error was set at 0.1r, rather than 0.12r as recommended in the in MICS3 Manual.

- $p$  is the proportion of the total population upon which the indicator,  $r$ , is based; and,
- $n_h$  is the average household size in Sierra Leone.

For the calculation,  $r$  (DPT3 coverage rate) was estimated to be 35 percent. The value of  $deff$  (design effect) was taken as 1.75<sup>2</sup> based on estimates from previous surveys,  $p$  (percentage of children aged 12-23 months in the total population) was taken as 3 percent, and  $n_h$  (average household size) was taken as 6.0 households.

The resulting “ $n$ ” or number of required households that was calculated using the formula above was 7944, which was rounded up to 8000 households. It was decided that the cluster size would be 25 households, based on a number of considerations that include the available budget and the estimated time that was required for a team to completely survey one cluster. Dividing the total number of households by the number of households per cluster, it was calculated that a total of 320 clusters was required.

Probability proportion to size ( $pps$ ) method was used to allocate clusters to districts in order to create a self-weighting sample. Clusters and EAs were then selected within districts also according to  $pps$  methods as described in the box below. The table below shows the allocation of clusters to the districts.

**Table 1: Distribution of EAs for Sierra Leone MICS3, by district**

Local Council Area	EAs	EA Prop.	MICS EAs	Households
Kailahun District	704	0.0728	23	575
Kenema District	792	0.0819	26	650
Kenema Town	212	0.0219	7	175
Koidu Town	61	0.0063	2	50
Kono District	544	0.0562	18	450
Bombali District	688	0.0711	23	575
Makeni Town	122	0.0126	4	100
Kambia District	506	0.0523	17	425
Koinadugu District	510	0.0527	17	425
Port Loko District	890	0.0920	29	725
Tonkolili District	825	0.0853	27	675
Bo District	629	0.0650	21	525
Bo Town	209	0.0216	7	175
Bonthe District	346	0.0358	11	275
Bonthe Town	17	0.0018	1	25

---

This was done in order to increase the precision of indicator estimates at both the national as well as at the level of the districts, in support of the government’s national decentralization programme.

<sup>2</sup> The design effect  $f$  was estimated at 1.75 (rather than 1.5, as recommended in the MICS3 Manual) to allow the selection of a larger sample size, and thus to increase the precision of estimates.

Moyamba District	616	0.0637	20	500
Pujehun District	477	0.0493	16	400
Western Rural	176	0.0182	6	150
Western Urban	1,349	0.1395	45	1,125
<b>SIERRA LEONE</b>	<b>9,673</b>	<b>1.0000</b>	<b>320</b>	<b>8,000</b>
% of all EAs included in MICS3	3.3%			

## Sampling Frame and Selection of Clusters

A multi-stage, stratified cluster sampling approach was used to select the survey sample. The 2004 census frame was used for the selection of clusters. Census enumeration areas (EAs) were defined as primary sampling units (PSUs), and were selected in each district using *pps* sampling procedures. The stages of the sampling approach are described below.

### Box 1: Description of sampling approach for Sierra Leone MICS3

#### Stage 1: Selection of EAs

1. The list of all EAs in Sierra Leone was ordered using implicit stratification according to the following variables: province; district; chiefdom; and, population size. 320 EAs were then selected using stratified systematic sampling, thus yielding a self-weighting sample. Selected EAs were then classified as *rural* (population of the settlement where the EA is located is < 2,000) or *urban* (population of the settlement where the EA is located is ≥ 2,000).

#### Stage 2: Selection of households

2. A list of all households in each of the 320 selected EAs as enumerated during the 2004 census was prepared using data contained in the 2004 Population and Housing Census registers.
3. A team of listers/verifiers visited each of the 320 EAs to update the household lists in the EA by verifying each of the households on the list and adding any new households that have been formed in order to control for out-movers, non-existent households, and/or new households. This task produced an updated listing of households in all selected EAs.
4. The newly updated listing of households in each EA was then sequentially numbered from 1 to *n* (the total number of households in the enumeration area of interest) at the Statistics Sierra Leone Office. Sampling experts then selected 25 households in each EA using systematic selection procedures.

## Calculation of Sample Weights

Although the Sierra Leone MICS3 sample was self-weighted, weighting techniques were used to make adjustments to correct for modest inter-PSU differences due to non-response for the household and individual interviews. The adjustment for household non-response is equal to the inverse value of:

$$RR = \text{Number of interviewed households} / \text{Number of occupied households listed}$$

After the completion of fieldwork, response rates were calculated for each sampling domain. These were used to adjust the sample weights calculated for each cluster. Response rates in the Sierra Leone MICS3 are shown in Table HH.1 in this report.

Similarly, the adjustment for non-response at the individual level (women and under-5 children) is equal to the inverse value of:

$$RR = \text{Completed women's (or under-5's) questionnaires} / \text{Eligible women (or under-5s)}$$

Numbers of eligible women and under-5 children were obtained from the household listing in the Household Questionnaire in households where interviews were completed.

The unadjusted weights for the households were calculated by multiplying the above factors for each enumeration area. These weights were then standardized (or normalized). Normalization of weights results in the sum of the interviewed sample units equalling the total sample size at the national level. Normalization is performed by multiplying the aforementioned unadjusted weights by the ratio of the number of completed households to the total unadjusted weighted number of households. A similar standardization procedure was followed in obtaining standardized weights for the women's and under-5's questionnaires. Adjusted (normalized) household weights varied between 0.99 and 1.02 in the 320 EAs. Adjusted woman's weights varied between 0.84 to 1.21 while children's weights varied between 0.88 and 1.16. Details regarding these weights are provided in the tables on the following pages.

Sample weights were appended to all data sets and analyses were performed by weighting each household, woman or under-5 using these weights.

The figure below presents a map of Sierra Leone that shows the locations of the 320 clusters that were selected for the MICS3 survey.

**Figure 1: Location of clusters in Sierra Leone MICS3 survey**

