

São Tomé and Príncipe - Multiple Indicator Cluster Survey 2014

United Nations Children's Fund, National Institute of Statistics, UNDP/Global Fund project, National Centre for Endemic Diseases

Report generated on: December 1, 2016

Visit our data catalog at: <http://microdata.worldbank.org/index.php>

Sampling

Sampling Procedure

The primary objective of the sample design was to produce statistically reliable estimates of most indicators, at the national level, for urban and rural areas, for the Southern and Northern regions, and each one of the two districts of Agua Grande and Me Zochi in the Central region.

The first stage statistical units, or primary sampling units, are the enumeration areas (EAs) designed during the cartographic operations of the 2012 General Census of Population and Housing (GCPH). The list of EAs constitutes the first stage sampling frame of primary sampling units.

A sample of EAs was drawn at the first stage within each stratum. The second stage sampling units are the households within the EAs drawn at the first stage. They define the secondary stage sampling frame.

A study domain is a portion of the national territory for which valid results are sought, in other words individual estimates of sufficient precision. The districts were not chosen as study domains given their small size. Beside the urban and rural areas, four other study domains were identified. These are the district of Agua Grande, the district of Me Zochi, the Southern region comprising the districts of Cantagalo and Caue, and the Northern region comprising the districts of Lemba and Lobata. The region of Príncipe, which has only 11 EAs, was not regarded as a study domain.

The stratification is defined as the urban/rural area of residence within each district, which led to 13 strata.

The household sample size for the MICS5 survey was calculated as 3930 households. For the calculation of the sample size, the key indicator used was the immunization coverage of children age 12-23 months.

In order to estimate how many clusters or primary units would correspond to 900 households selected in each study domain, three options of cluster size were considered:

- 20 households per primary unit
- 25 households per primary unit
- 30 households per primary unit

It was found that the first two cluster size options of 20 or 25 households did not provide the required number of clusters in certain strata. On the other hand the option of 30 households per primary unit allowed drawing the EAs in all the strata.

In the Region Autonomous of Príncipe, which is not a study domain, all 11 EAs have been surveyed on the basis of 30 households each, giving a national sample of 3930 households. This results in a global sample of 131 primary sampling units or clusters, 30 per study domain and 11 for the region of Príncipe.

The sample selection is implemented independently within each stratum. The primary sampling units or clusters are drawn systematically with probability proportional to size. The probability of selection of a cluster is proportional to the size of the cluster, the size being here defined as the number of households in the cluster from the frame.

At the second sampling stage, i.e. the selection of the households, systematic sampling with equal probability is used. An equal number of households, 30 in this case, is drawn in each cluster selected at the first stage.

The sampling of primary units is done independently one stratum at a time. As previously indicated, the EAs are drawn systematically with probability proportional to size.

The drawing of the primary units is implemented with a computer software called TIRAGE 2.1 specially designed for random sampling. In preparation for the draw, it was first verified if any of the 13 strata included atypical EAs, i.e. EAs with a selection probability greater than 1.

The sampling procedures are more fully described in "Multiple Indicator Cluster Survey 2014 - Final Report" pp.234-241.

Response Rate

Of the 3,930 households selected for the sample, 3,625 were found to be occupied. Of these, 3,492 were successfully interviewed which leads to a household response rate of 96 percent. The women's response rate was 95 percent, the men's

response rate 82 percent, and the children's response rate 98 percent.

Weighting

The following notation is used in the formulae determining the selection probabilities and the sample weights of the sampling units for estimates within a stratum h .

- h represents the stratum in the study domain;
- m_h is the number of PSUs (primary sampling units) drawn in the stratum h ;
- the stratum h is composed of M_h PSUs labeled 1, 2, ..., M_h ; the PSU i of stratum h is noted UP_{hi} ;
- N_{hi} represents the size of PSU UP_{hi} ;
- The size N_{hi} is from the sampling frame being used, the number of households in the PSU UP_{hi} ;
- N_h represents the sum of the sizes N_{hi} of the PSUs UP_{hi} and is defined by the relation
- n is the fixed number of households selected at the 2nd stage of the PSU UP_{hi} in the stratum h .

At the 1st stage, m_h PSUs are drawn from stratum h systematically with probability proportional to size.

At the 2nd stage, a fixed number n of households are drawn in each sampled PSU in stratum h for the three questionnaires of the survey related to the household, the women and the children below the age of 5 years.

Questionnaires

Overview

The questionnaires for the Generic MICS were structured questionnaires based on the MICS5 model questionnaire with some modifications and additions. Household questionnaires were administered in each household, which collected various information on household members including sex, age and relationship. The household questionnaire includes List of Household Members, Education, Child Labour, Child Discipline, Household Characteristics, Insecticide Treated Nets, Indoor Residual Spraying, Water and Sanitation, Handwashing, and Salt Iodization.

In addition to a household questionnaire, questionnaires were administered in each household for women age 15-49, men age 15-49 and children under age five. The questionnaire was administered to the mother or primary caretaker of the child.

The women's questionnaire includes Woman's Background, Access to Mass Media and Use of Information/Communication Technology, Fertility/Birth History, Desire for Last Birth, Maternal and Newborn Health, Post-natal Health Checks, Illness Symptoms, Contraception, Unmet Need, Attitudes Toward Domestic Violence, Marriage/Union, Sexual Behaviour, HIV/AIDS, Maternal Mortality, Tobacco and Alcohol Use, and Life Satisfaction.

The men's questionnaire includes Man's Background, Access to Mass Media and Use of Information/Communication Technology, Fertility, Attitudes Toward Domestic Violence, Marriage/Union, Sexual Behaviour, HIV/AIDS, Circumcision, Tobacco and Alcohol Use, and Life Satisfaction.

The children's questionnaire includes Child's age, Birth Registration, Early Childhood Development, Breastfeeding and Dietary Intake, Immunization, Care of Illness, and Anthropometry.

The blood test questionnaire includes Anaemia and malaria test for children 6-59 months of age, Anaemia and HIV test for women age 15-49 years, and HIV test for men age 15-49 years.

Data Collection

Data Collection Dates

Start	End	Cycle
2014-04-07	2014-06-18	N/A

Data Collection Mode

Face-to-face [f2f]

DATA COLLECTION NOTES

Training for the fieldwork was conducted for 15 days between 3 and 21 March 2014. The training included lectures on interviewing techniques and the contents of the questionnaires, and mock interviews between trainees to gain practice in asking questions. Towards the end of the training period, trainees spent 7 days practicing interviewing in the district of gua Grande and M Zochi. The training also provided guidelines for the field-level collection and analysis of blood samples and biomarkers data, using a specifically designed manual and questionnaire.

The data were collected by eight teams; each was comprised of four interviewers, two health technicians (for anthropometry and blood sample collection), one driver, one editor and a supervisor. Fieldwork began on 7 April 2014 and was concluded on 18 June of the same year.

Data Collectors

Name	Abbreviation	Affiliation
National Institute of Statistics		

SUPERVISION

There is one supervisor for each of the 8 data collection teams in the field.

Data Processing

Data Editing

Data were entered using the CSPro software, Version 5.0. The data were entered on ten desktop computers, procured specifically for the purposes of the 2014 MICS, and carried out by 20 data entry operators and two data entry supervisors working in two shifts (morning and afternoon). For quality assurance purposes, all questionnaires were double-entered and internal consistency checks were performed. Procedures and standard programmes developed under the global MICS programme and adapted to the 2014 Sao Tome and Principe questionnaires were used throughout. Data processing followed rapidly the start of data collection on 14 April and was completed on 28 June 2014. Data were analyzed using the Statistical Package for Social Sciences (SPSS) software, Version 21. Model syntax and tabulation plans developed by UNICEF were customized and used for this purpose.

The processing of the blood samples was conducted from August to September 2014 for the malaria samples and from January to February 2015 for the HIV samples. The processing of the HIV samples was initiated after the scrambling and anonymization of the MICS data collected through the questionnaires. Blood samples were analyzed at the Hospital Ayres de Menezes Laboratory in Sao Tome and Principe. For HIV testing in particular, Elisa (Vironostika VIH Ag/Ab) was used for all samples as a first test. Negative samples from this first testing were classified as negative whereas positive samples were subjected to a second ELISA test (Enzygnost VIH Integral II). Positive samples from this second test were classified as positive. Discordant cases between the first and second ELISA test were reanalyzed using the two tests. Discordant cases were analyzed once again using Western Blot 2.2. Ten percent of negative cases were also subjected to another ELISA test for quality control purposes. At the end of the process, 261 samples, including all positive cases, were sent to the Centre Pasteur in Cameroon for external quality control (EQC). The results of the EQC, communicated in May 2014, coincided with those obtained in Sao Tome and Principe.

Data Appraisal

Estimates of Sampling Error

Sampling errors are a measure of the variability between the estimates from all possible samples. The extent of variability is not known exactly, but can be estimated statistically from the survey data.

The following sampling error measures are presented in this appendix for each of the selected indicators:

- Standard error (se): Standard error is the square root of the variance of the estimate. For survey indicators that are means, proportions or ratios, the Taylor series linearization method is used for the estimation of standard errors. For more complex statistics, such as fertility and mortality rates, the Jackknife repeated replication method is used for standard error estimation.
- Coefficient of variation (se/r) is the ratio of the standard error to the value (r) of the indicator, and is a measure of the relative sampling error.
- Design effect (deff) is the ratio of the actual variance of an indicator, under the sampling method used in the survey, to the variance calculated under the assumption of simple random sampling based on the same sample size. The square root of the design effect (deft) is used to show the efficiency of the sample design in relation to the precision. A deft value of 1.0 indicates that the sample design of the survey is as efficient as a simple random sample for a particular indicator, while a deft value above 1.0 indicates an increase in the standard error due to the use of a more complex sample design.
- Confidence limits are calculated to show the interval within which the true value for the population can be reasonably assumed to fall, with a specified level of confidence. For any given statistic calculated from the survey, the value of that statistic will fall within a range of plus or minus two times the standard error ($r + 2.se$ or $r - 2.se$) of the statistic in 95 percent of all possible samples of identical size and design.

For the calculation of sampling errors from MICS data, programs developed in CSPro Version 5.0, SPSS Version 21 Complex Samples module and CMRJacki have been used.

The results are shown in the tables that follow. In addition to the sampling error measures described above, the tables also include weighted and unweighted counts of denominators for each indicator. Given the use of normalized weights, by comparing the weighted and unweighted counts it is possible to determine whether a particular domain has been under-sampled or over-sampled compared to the average sampling rate. If the weighted count is smaller than the unweighted count, this means that the particular domain had been over-sampled.

Other forms of Data Appraisal

A series of data quality tables are available to review the quality of the data and include the following:

- Age distribution of household population
- Age distribution of eligible and interviewed women
- Age distribution of eligible and interviewed men
- Age distribution of children in household and under-5 questionnaires
- Birth date reporting: Household population
- Birth date and age reporting: Women
- Birth date and age reporting: Men
- Birth date and age reporting: Under-5s
- Birth date reporting: Children, adolescents and young people
- Birth date reporting: First and last births
- Completeness of reporting
- Completeness of information for anthropometric indicators: Underweight
- Completeness of information for anthropometric indicators: Stunting
- Completeness of information for anthropometric indicators: Wasting
- Heaping in anthropometric measurements
- Observation of birth certificates
- Observation of vaccination cards
- Observation of women's health cards
- Observation of bednets and places for handwashing
- Presence of mother in the household and the person interviewed for the under-5 questionnaire
- Selection of children age 1-17 years for the child labour and child discipline modules
- School attendance by single age
- Sex ratio at birth among children ever born and living
- Births by periods preceding the survey
- Reporting of age at death in days
- Reporting of age at death in months
- Completeness of information on siblings
- Sibship size and sex ratio of siblings

The results of each of these data quality tables are shown in appendix E in document "Multiple Indicator Cluster Survey 2014 - Final Report" pp.264-283.

Related Materials

Questionnaires

Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Questionnaire

Title	Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Questionnaire
Date	2014-12-01
Country	Sao Tome and Principe
Language	Portuguese
Table of contents	Household questionnaire modules Household member questionnaire modules Women questionnaire modules Men questionnaire modules Children questionnaire modules Blood test questionnaire modules
Filename	Sao Tome and Principe 2014 MICS_Portuguese_Questionnaire.pdf

Reports

Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Biomarker Report

Title	Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Biomarker Report
Author(s)	National Institute of Statistics UNDP/Global Fund project National Centre for Endemic Diseases United Nations Children's Fund (UNICEF)
Country	Sao Tome and Principe
Language	Portuguese
Filename	https://mics-surveys-prod.s3.amazonaws.com/MICS5/West%20and%20Central%20Africa/Sao%20Tome%20and%20Principe/2014/Biomarker/Sao%20Tome%20and%20Principe%202014%20MICS%20Biomarker%20Report_Portuguese.pdf

Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Report

Title	Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Report
Author(s)	National Institute of Statistics UNDP/Global Fund project National Centre for Endemic Diseases United Nations Children's Fund (UNICEF)
Country	Sao Tome and Principe
Language	English
Filename	https://mics-surveys-prod.s3.amazonaws.com/MICS5/West%20and%20Central%20Africa/Sao%20Tome%20and%20Principe/2014/Final/Sao%20Tome%20and%20Principe%202014%20MICS_English.pdf

Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Key findings report

Title	Sao Tome and Principe Multiple Indicator Cluster Survey 2014 - Key findings report
Author(s)	National Institute of Statistics UNDP/Global Fund project National Centre for Endemic Diseases United Nations Children's Fund (UNICEF)
Country	Sao Tome and Principe
Language	Portuguese
Filename	https://mics-surveys-prod.s3.amazonaws.com/MICS5/West%20and%20Central%20Africa/Sao%20Tome%20and%20Principe/2014/Key%20findings/Sao%20Tome%20and%20Principe%202014%20MICS%20KFR_Portuguese.pdf
