Mobile Phone Survey

Sampling Design, Data Management and Analysis

Version 4
October 2025

CDC Foundation
Together our impact is greater

Editor's Note

Financial support was provided by the Bloomberg Philanthropies Data for Health Initiative through the CDC Foundation with a grant from Bloomberg Philanthropies.

The views expressed herein do not necessarily represent the views of Bloomberg Philanthropies or the CDC Foundation.

Table of Contents

List of Tables

List of Figures

Abbreviations

D4H	Data for Health
MPS(s)	Mobile Phone Survey(s)
RDD	Random Digit Dialing
IVR	Interactive Voice Response
SMS	Short Message Service
CATI	Computer-Assisted Telephone Interviewing
MPN(s)	Mobile Phone Number(s)
NCD(s)	Non-Communicable Disease(s)
SRS	Simple random sample
MNO(s)	Mobile Network Operator(s)
INRR(s)	Item-Non-Response-Rate(s)

1. Introduction

1.1 Overview

Efficient monitoring and surveillance are cornerstones to track progress of disease burden, related risk factors and policy interventions. The systematic monitoring of risk factors to generate accurate and timely data is essential for a country's ability to prioritize essential resources and make sound policy decisions.

With increasing access and use of mobile phones globally, opportunities exist to use mobile phone technology as an interim method to collect data and supplement infrequent household surveys. Such technologies have the potential to allow for efficiencies in producing timely, affordable and accurate data to monitor trends, and augment traditional health surveys with new, faster mobile phone surveys.

The Bloomberg Philanthropies Data for Health (D4H) Initiative aims to strengthen the collection and use of critical public health information. One of the components of the initiative aims to explore innovative approaches to risk factor surveillance, including the use of mobile phone surveys. The main objectives of this component are to:

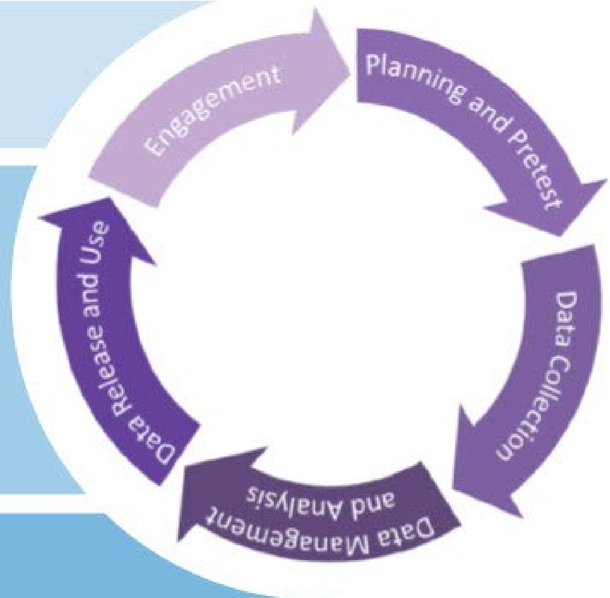
- **Empower Country Governments**: Collaborate with ministries of health to implement mobile phone technology for the systematic collection of representative health-related data, facilitating ongoing risk factor surveillance and informed public health decision-making.
- **Develop a Standardized Mobile Phone Survey Framework**: Create and promote a globally recognized mobile phone survey protocol that ensures uniformity in data collection methods, enhancing the reliability and comparability of health data across diverse populations.
- **Optimize Mobile Phone Survey Effectiveness**: Evaluate and identify best practices in the design and implementation of mobile phone surveys, ensuring they are tailored to effectively capture critical health risk factors.

The Mobile Phone Survey is a representative survey of adults 18 years of age and older. The survey uses standardized instruments and procedures reviewed and approved by international experts. This includes a core questionnaire with optional questions, sample design utilizing random digit dialing (RDD), data management procedures, and data collection using single or mixed modes such as interactive voice response (IVR), short message service (SMS), mobile web, and computer-assisted telephone interviewing (CATI). The implementation process consists of five stages: 1) engagement; 2) planning and pretesting; 3) data collection; 4) data management and analysis; and 5) data release and use. Details on each stage are presented in the Mobile Phone Survey Process Chart (see **Figure 1**).

Figure 1. Mobile Phone Survey Process Chart

Mobile Phone Survey Process

The Mobile Phone Survey is a representative survey of adults aged 18 or older that uses mobile phone technology to collect data on diseases and associated risk factors. Data are collected through computer assisted telephone interviews (CATI), text message (SMS), automated phone call (IVR), mobile web or mixed modes following a standard protocol. The Mobile Phone Survey provides timely data to support monitoring and evaluation of public health programs and policies.



Engagement

- Formalize stakeholder coordination plan
- Determine implementing agency
- Develop workplan and timeline

Planning and Pretest

- Adapt protocol, including questionnaire and sample design
- Obtain ethical and regulatory approvals
- Determine data hosting solution
- Establish connections with mobile network operators
- Set up and test data collection tool
- Perform pretest data collection
- Determine optimal survey parameters and sample size for full-scale data collection
- Perform cognitive testing

Data Collection

- Finalize protocol
- Launch media awareness campaign
- Deploy full-scale data collection
- Monitor data collection progress

Data Management and Analysis

- Aggregate data
- Execute quality assurance checks
- Construct sampling weights
- Analyze data

Data Release and Use

- Finalize fact sheet and executive summary
- Release data
- Disseminate findings

Visit NCDmobile.org for all tools including the following technical manuals:

- Implementation Instructions
- Survey Technology
- Questions and Indicators
- Data Management and Analysis
- Sampling Design

1.2 Purpose

This manual is designed to outline the requirements and guidelines to develop an appropriate sample design for implementation of the Mobile Phone Survey:

- **Survey Design Objectives:** A summary of the survey objectives
- **Target Population and Sampling Frame:** A definition of the target population for the survey and discussion of the sample frame that will be used
- **Basic Survey Design Specification:** The basic features of the sample design that are required to achieve established levels of statistical quality and facilitate intercountry comparability
- **Overview of the Sampling Approach:** An overview of the sampling design
- **First Phase of Sampling:** The recommended approach for selecting mobile phone numbers (MPNs)
- **Second Phase of Sampling:** How the sample will be stratified after selection of MPNs
- **Determining Sample Sizes and Response Rates:** Guidance on determining the selected sample sizes needed to produce a respondent sample size and computing the final eligibility and response rates for the survey
- **Computing Final Sample Weights:** The steps for computing the sampling weights for statistical analysis
- **Data Collection:** Different modes used in data collection, survey contact attempts, time intervals and other features available
- **Data Management and Quality Assurance Processes:** The quality assurance guidelines and procedures used in the post-data collection phase.
- **General Analytic Guidelines.** Data analysis used to generate both the Fact Sheet and the Executive Summary.

2. Survey Design Objectives

Sample design requirements for the Mobile Phone Survey have been developed so that unbiased population estimates of prevalence and health risks can be generated for the target population of each country and for two analysis groups of interest defined separately by age or sex.

The allocation of the sample will be accomplished by explicitly stratifying the second-phase sample by age group (i.e., 18–29, 30–44, and 45 years and older) and sex (i.e., male and female) with disproportionate allocation (to the mobile phone user population) among strata to produce a final respondent sample of mobile phone users that resemble the joint age group by sex distribution of the general population (as defined by the National Statistics Office or other reputable sources, such as the *2024 Revision of World Population Prospects* from United Nations Department of Economic and Social Affairs Population Division).

Designing the survey to generate precise joint estimates by sex and age group is preferred for the D4H Mobile Phone Survey, primarily to allow comparisons of estimates by these domains between different countries participating in this effort.

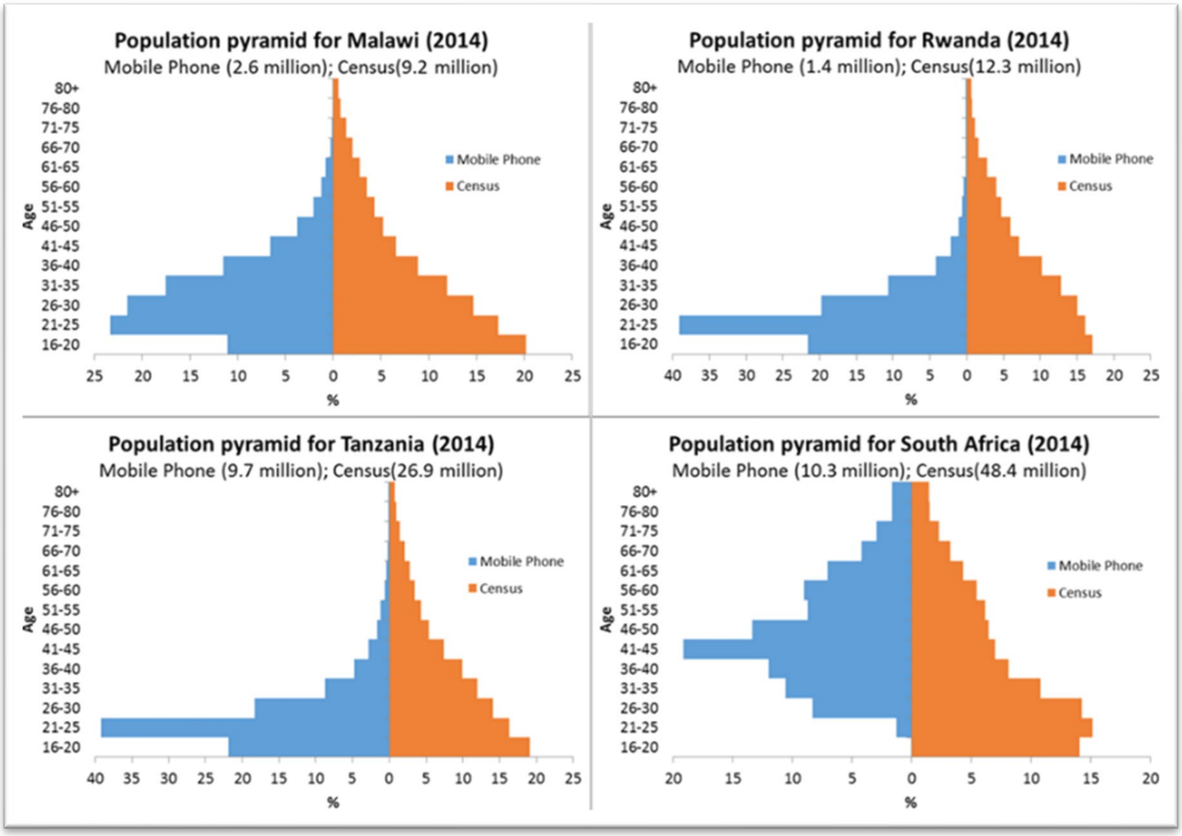
2.1 Sample Design for Countries Implementing the Mobile Phone Survey Protocol

The sample design is a two-phase sampling approach where MPNs are randomly sampled in the first phase and a subsample of respondents from the first phase are stratified by age and sex in the second phase, proportional to the general population.

- **First phase** – Randomly sample MPNs from an implicit MPN frame of all possible MPNs.
- **Second phase** – The allocation of the sample is accomplished by explicitly stratifying the second phase sample jointly by age group (i.e., 18–29, 30–44, 45 years and older) and sex (i.e., male and female) with disproportionate allocation to the mobile phone user population among strata. This allocation attempts to adjust for underrepresentation of certain groups from the general population in the mobile phone user population, namely individuals age 45 years and older.

  This sample design is described in detail in **Section 3**.

Figure 2. Second phase – sample stratified proportional to the general population but disproportionate to the mobile phone user population



3. Target Population and Sample Frame

3.1 Definitions of the Target Population and Study Eligibility

In general, the target population for the Mobile Phone Survey should include all eligible individuals, 18-99 years of age, with MPNs within the country's MPN series (refer to the ***Surveda Technical Manual: IVR, SMS, and Mobile Web MPS Modes*** for definition). To propose an exclusion, a country should explicitly define which MPN ranges are included in and excluded from the target population and provide an estimate of the percentage of the population that resides in those ranges being excluded. This percentage will provide an indication of the potential bias effect of any real exclusions on estimates that are otherwise intended to be national in scope.

The target population for the Mobile Phone Survey should include men and women who are 18-99 years of age. This definition is important when completing the demographic questions, which determines who is eligible to complete the noncommunicable disease (NCD) questions.

3.2 Sampling Frame

In general, a sampling frame is a list of all eligible members of the target population from which the sample could be drawn. For the Mobile Phone Survey, the following parameters have been used:

- A two-phase sampling of mobile phone users from an *implicit* frame of MPNs, with stratified sampling disproportionate to the mobile phone user population in the second phase.
- An *implicit frame* is an implied sampling frame produced from the range of all possible MPNs in the country, some of which will not currently be in service. This frame will be constructed in collaboration with the country's telecommunication regulatory agency.
- For this study design, a comprehensive list of all survey-eligible individuals in a country likely does not exist, making sampling from the preferred frame virtually impossible. Consequently, an indirect list frame should be used to select the sample. This list is implied by the known set of possible MPNs that could be assigned by providers to their customers.
- This option is preferred whenever a reputable national listing of MPNs currently in service is unavailable.

4. Survey and Sample Design

4.1 Basic Survey Design Specifications

- The *sampling unit* during sample selection refers to the entities that are selected for the survey. In this survey, the sampling units are randomly-generated MPNs. For the Mobile Phone Survey, stratification cannot occur until after the sample is selected. In such cases, two-phase sampling can be used to control the demographic composition of the final sample.
- The Mobile Phone Survey sample design uses a two-phase sample of mobile phone users from the sampling frame of all possible MPNs with stratified sampling disproportionate to the mobile phone user population in the second phase. Within each stratum, each mobile phone user linked to a single MPN will have the same probability of being selected.
- In the first phase of sample selection, a random sample of MPNs is generated using the country code numbering plan + operator prefix + random generation of remaining digits from the range of possible MPNs. The sample for MPNs is based on the ratio of provider share in the market. The sample is then screened and filtered for active/non-active MPNs. Disconnected numbers and non-residential numbers are removed from the generated sample.
- In the second phase of sample selection, and to improve bias and representativeness, respondents selected for the Mobile Phone Survey are screened and assigned to one of several strata categorized by age group and sex. Data are collected on respondents within age/sex strata until individual stratum sample sizes are achieved or until the data collection period has expired. Once sample size is met in a specific stratum, data are collected for the demographic questions only, but not for any additional questions, until the targeted respondent sample sizes for all strata are attained.
- As noted, certain requirements and recommendations should be followed when conducting the Mobile Phone Survey to maximize the comparability of the results between countries. However, each country has the option of introducing design enhancements that allow it to increase the usability of the results from this survey (e.g., selecting the sample to ensure precise estimates by region). In the next section, we present some of the basic survey design requirements. Any design enhancement that a country wants to introduce will generally be acceptable, provided it does not interfere with these basic requirements and, thus, contributes to a loss of intercountry comparability of survey estimates.

- For the purposes of clarity in discussion of sampling features, including sampling weights and adjustments, the demographic questions will be referred to as the Demographic Module and the NCD questions will be referred to as the NCD Module.

4.2 Sample Design Features

Requirements related to the sample design and sampling weights include the following:

- Simple random selection (SRS), without replacement, should be used so that every member of the implicit sampling frame of mobile phone users has a computable, non-zero chance of being selected into the sample.
- Sample strata are constructed based on the known distribution of the general population from official population data because the distribution of the mobile phone user population is not known. Once data is collected, the mobile phone user population distribution will be estimated based on age and sex. This estimated mobile phone user population distribution, and the known general population distribution, will be used to adjust the sampling weights.
- Randomly selected mobile phone users should be enrolled into both the Demographic Module and the NCD Module until the appropriate respondent sample size for a given stratum is met, or the data collection period has expired.
- Survey nonparticipation because of ineligibility and nonresponse should be tracked to properly compute response rates in the Mobile Phone Survey. For instance, MPNs may not be active, a mobile phone user may be under 18 years of age and ineligible for the survey, an eligible mobile phone user may refuse to participate, or an active MPN may lose connection to the mobile platform system during an interview. A list of all applicable disposition codes will be assigned to all MPNs selected for the survey to document eligibility and nonresponse. If local definitions of these codes are used, they should be standardized to the codes provided in this manual. Conversion rules of these codes between local and standard definitions should also be provided.

4.3 Sample Sizes and Expected Precision

Requirements and recommendations related to respondent sample size are based on the following indicators of statistical quality that were established for the Mobile Phone Survey:

- The survey should be designed to produce estimates that meet the following precision requirements: estimates computed at the national level by age, by sex, and by the cross of sex and age should have a 95% confidence interval with a margin of error of 5 percentage points or less for NCD risk factor prevalence of 50%.
- The design effect ($Deff_o$) associated with any estimate from a survey is defined as the multiplicative factor increase in the variance of survey estimates because of complex

survey design features, such as unequal sampling probabilities. The multiplicative effect as a result of variable weights, defined here as $Meff_{Wts}$, is multiplied times the comparable effect for cluster sampling ($Meff_{CS}$) to produce the overall value of $Deff_{o.}$ For the Mobile Phone Survey, the $Meff_{CS}$ is 1, due to the assumption that there is no clustering in the MPS sample design, since MPNs were selected using RDD. By definition, $Deff_{o,}$ is the ratio of the variance of an estimate based on the complex survey design relative to the corresponding variance of the same sample size using simple random sampling.

- For the Mobile Phone Survey, sample weights will be variable mostly because of sample variation in second-phase stratum sampling rates. This variability among final sample weights increases the variance of survey estimates by a factor of $Meff_{Wts}$, which may be approximated by

  $Meff_{Wts} \approx 1 + \{CV_{Wts}\}^2 \geq 1.00$, where CV_{Wts} is the coefficient of variation among all sample weights.

- Most quantitative indicators of the statistical quality of estimates from sample surveys are mathematically related to the variance of estimates among all possible outcomes of the survey design. When limiting our attention to health-related estimates of some proportion *(P)* (informed by a country's specific health policy and goals) in the population of those at risk of adverse health outcomes, the variance of the estimator *(* $\hat{P}$ *)* among all possible outcomes of the sample design is approximately,

$$V(\hat{P}) = \left[Deff_o\right]\frac{P(1-P)}{n_r} = \left[Meff_{Wts}Meff_{CS}\right]\frac{P(1-P)}{n_r} \qquad \textbf{\textit{(Eq. 1)}}$$

Where $Deff_o = Meff_{Wts}{*}Meff_{CS}$ is the overall design effect (Gabler et al., 1999) $Meff_{CS}$ is the multiplicative increase in variance due to the use of cluster sampling (in this case $Meff_{CS}$ = 1) and *n* is the number of sample respondents used to produce the estimate of *P*. From Eq. (1) we see that statistical quality is influenced by *n*, the actual size of *P* and the value of $Meff_{Wts}$. Experience indicates that $Meff_{Wts}$ could be quite large (Gabler, 1999), particularly in countries with lower mobile phone penetration. In this instance, highly disproportionate sampling or major calibration adjustments are needed for the weighted respondent sample to match the target population for subgroup comparisons.

The design of the sample should correctly reflect anticipated levels of nonresponse and ineligibility in determining how many MPNs should be selected to yield the recommended number of respondents. For example, a person selected for interview may refuse to participate (nonresponse). Similarly, a selected MPN may prove to not be in service, or selected persons may indicate they are younger than 18 years old and are therefore ineligible. Rates of eligibility and response among eligible mobile phone users are multiplied to determine the overall attrition rate in samples. These components of attrition during sample recruitment should be estimated as accurately as possible from recent relevant survey experience, regardless of which sample design option is proposed.

For example, suppose the survey in a country is designed to produce n_f female respondents and n_m male respondents, and it expects to observe the following parameters shown in **Table 1**.

Table 1. Example parameters for sample size computation

Rate	Comment	Assumption
Eligibility Rate (ER)	Accounts for those cases when respondents are interviewed for the survey and later determined to be ineligible (e.g., they are younger than 18 years old)	80%
Response Rate (RR)	Accounts for those eligible respondents who are selected but do not complete the Demographic Module and at least one question in the NCD Module	4%

The actual values assumed for the active number rate, eligibility rate and response rates should be informed by *apriori* information obtained about the population age and sex distribution from certified sources (e.g., recent census, United Nations projections) as well as by the results of the pretest(s).

The formula for calculating the effective sample size is as follows:

$$n_{eff} = \frac{n}{ER * RR}$$

(Eq. 2)

Additional guidelines for determining an appropriate sample size at each step of the Mobile Phone Survey sample design are provided in **Section 6.**

5. Overview of the Sampling Approach

The sample for the Mobile Phone Survey should be selected using a multi-phase, stratified design to ensure the respondent sample closely matches relevant demographics of the entire target population. In summary, this sample selection process should proceed as follows.

- The first phase of the design involves randomly selecting MPNs. This selection process is highly dependent on the MPN format convention in the country that is fielding the survey and the quantity of mobile network operators (MNO) and mobile telephone prefixes.
- The selected sample for each of the age and sex strata defined for sampling in the second phase of the design should be allocated so that the final respondent sample is proportional to the age-by-sex distribution of the country's target population. The age-by-sex distribution of the country's target population can be constructed from a recent census, administrative records, or a recent large and reputable demographic survey and, in general, should equal the total eligible population, 18 years of age and older. If the Mobile Phone Survey target population counts are not available, corresponding total population counts can be used as stratum-size measures.
- MPNs should continue to be collected until all strata sample sizes are achieved or until the data collection period has expired.

6. First Phase of Sampling: Selecting MPNs

A random sample of MPNs will be selected using simple random sampling according to the following guidelines:

- MPNs should be selected via SRS using RDD without replacement.
- For generation of the MPNs, the country code numbering plan will be used first, followed by the operator prefix. The remaining digits are randomly generated.
- The sample for MPNs should be based on the ratio of provider share in the market to improve bias and representativeness.
- MPNs will be screened and filtered for active/non-active MPNs (disconnected numbers are removed, non-residential numbers are purged from the generated sample.)
- Every MPN on the sampling frame has some nonzero probability of selection, and the probabilities of selection are retained in the final analysis file. These probabilities of selection are one of several pieces of statistical information that are used to compute sample weights for the ultimate respondents to the Mobile Phone Survey.

7. Second Phase of Sampling: Selecting and Stratifying Eligible Mobile Phone Users

After the MPNs are selected for the Mobile Phone Survey, the next phase of the sample design involves stratifying the final sample.

The prevalence of most chronic disease risk factors tends to increase with age and vary by sex. Therefore, survey results should include estimates for specific age groups for each sex in addition to the total survey population estimates to provide a more nuanced picture of the prevalence of NCD risk factors in the target population.

The recommended age groups are 18–29, 30–44 and 45-99 years of age, though any age groups may be used. Estimates may be obtained only for the entire age span of the survey (e.g., 18 years and older) or for age groups (e.g., 18-29, 30 years and older).

A country's final sample design should include a detailed description of how a country plans to form a suitable sampling frame for the selection of MPNs, how the country will address frame incompleteness (such as new mobile phone operators opening or closing for business during data collection) and how randomization will be used to select the sample of MPNs.

The sample mobile phone users will be determined by explicitly stratifying the second-phase selection of users by age group (18-29, 30-44, 45-99 years of age) and sex (male and female) based on information received in the Demographic Module. Final respondent sample sizes in the strata are intended to be proportional to the age group and sex distribution from the general population (as defined by the National Statistics Office or other reputable sources, such as the *World Population Prospects* from United Nations Department of Economic and Social Affairs Population Division).

Randomly selected eligible mobile phone users should be enrolled into **both** the Demographic Module and the NCD Module until the sample size for each second-phase stratum is met.

Once a stratum sample size is met, selected members of that stratum should continue to be recruited to complete a Demographic Module until respondent sample size goals have been met in all strata or until the data collection period has expired. Counts of completed Demographic Modules and of those completing an NCD Module will be needed later for each second-phase stratum, which is defined by age and sex to compute sampling weights.

8. Determining Sample Sizes and Reporting Final Response Rates

In this section, guidance is provided on determining the selected sample sizes needed to produce a respondent sample size in the Mobile Phone Survey that will meet precision and power requirements for analysis. The sample size discussed below is the number of respondents arising out of the selected sample, after nonresponse and ineligibility are considered. Also, guidance is provided on computing the final eligibility and response rates for the survey.

The definition of response rate and eligibility rate are particularly important because having all countries adhere to the same definitions of these rates will allow intercountry comparison of these data quality measures.

8.1 Sample Size Calculation

In the Mobile Phone Survey, *two types of sample sizes* need to be considered.

1. The respondent sample size (n) is the total number of respondents with complete or partially complete interviews. It is determined by the prevalence of the NCD risk factor, precision of the estimate, and the variation due to sample weights for each sex by age stratum.
2. The mobile phone number sample size (n_{MPN}) is the total number of mobile phone numbers required to dial to obtain the respondent sample size. This is determined by the respondent sample size, active number rate, eligibility rate and response rate.

In Section 8.2, guidance is presented on how to determine sample size based on fixed assumptions about active mobile number rates, response rates, eligibility rates, Meff$_{Wts}$, Meff$_{CS}$, prevalence, and Margin of Error ($MOE(\hat{P})$). This approach may be appropriate when there are sufficient resources to attain the possibly large target respondent sample size(s).

However, it is important to realize there is an alternate approach when resources are limited. In such cases, it may be advisable to fix the overall respondent sample size in the same way as the other assumptions and examine the impact on the estimation of $MOE(\hat{P})$. Appendix A offers several scenarios taking this approach, using the same assumptions regarding active number rates, response rates, eligibility rates, Meff$_{Wts}$, Meff$_{CS}$, and prevalence as used in the example in Section 8.2.

Recall that MPNs should be stratified by age and sex to ensure adequate sample size in certain stratum (e.g., women age 45 years and older). In this section, an example is provided that assumes a reporting subgroup sample size of 384 respondents, for a prevalence of 50% and margin of error of 5%. Countries with other respondent sample sizes and features should modify their calculations accordingly.

Prevalence, or proportion (P), estimates discussed below are stated in general terms. However, there are several domains for which P might be estimated.

For instance, there is the overall population, population subgroups defined by individual sampling strata, or population subgroups that cut across strata, such as urban/rural.

In the discussion and tables below, the estimator $\hat{P}$ is used to represent any of these domains.

From Eq. (1), we see that statistical quality will be influenced by *n*, the actual size of *P*, and the values of $Meff_{Wts}$ and $Meff_{CS}$

After solving for *n* in Eq. (2) below, we see that the number of survey respondents needed to achieve a particular 95% margin of error, $MOE(\hat{P}) = 1.96\sqrt{V(\hat{P})}$, for descriptive estimates of IP can be determined as,

$$n = \frac{(1.96)^2 P(1-P)}{\left[MOE(\hat{P})\right]^2} * [Deff_o] = \frac{(1.96)^2 P(1-P)}{\left[MOE(\hat{P})\right]^2} * [Meff_{Wts} * Meff_{CS}]$$

(Eq. 3)

Note that $MOE(\hat{P})$ is a measure of the *reliability* of the estimates of *P*. All surveys that take samples from a population with the intent of generalizing the results from the sample to the whole population have a possibility of error. This is because a sample can never perfectly describe the population. The higher $MOE(\hat{P})$ is, the less likely it is that the results of the survey are true for the whole population. Also note that the margin of error is related to the confidence interval because it is half the interval length.

For the case where $Deff_o$ = 1.00, **Table 2** indicates the respondent sample size requirements for various combinations of P and $MOE(\hat{P})$. For other likely settings where $Deff_o$ > 1.00, respondent sample size requirements are obtained by multiplying the value from **Table 2** times the anticipated value of $Deff_o$. Moreover, to achieve this level of precision for each of K key reporting subgroups (e.g., K = 6 to achieve the level of $MOE(\hat{P})$ for the estimates in each key subgroup), the overall respondent sample required is $K*Deff_o*n$ (the required sample size from **Table 2**).

Table 2. Respondent sample size (n) requirements for different combinations of $MOE(\hat{P})$ and P for age-by-sex reporting subgroups when $Deff_o = 1.00$

P	$MOE(\hat{P})$: 1%	1.5%	3%	5%
0.03	1,118	497	124	45
0.04	1,475	656	164	59
0.05	1,825	811	203	73
0.06	2,167	963	241	87
0.07	2,501	1,111	278	100
0.08	2,827	1,257	314	113
0.09	3,146	1,398	350	126
0.10	3,457	1,537	384	138
0.11	3,761	1,671	418	150
0.12	4,057	1,803	451	162
0.13	4,345	1,931	483	174
0.14	4,625	2,056	514	185
0.15	4,898	2,177	544	196
0.16	5,163	2,295	574	207
0.17	5,420	2,409	602	217
0.18	5,670	2,520	630	227
0.19	5,912	2,628	657	236
0.20	6,146	2,732	683	246
0.21	6,373	2,832	708	255
0.22	6,592	2,930	732	264
0.23	6,803	3,024	756	272
0.24	7,007	3,114	779	280
0.25	7,203	3,201	800	288
0.26	7,391	3,285	821	296
0.27	7,572	3,365	841	303
0.28	7,744	3,442	860	310
0.29	7,910	3,515	879	316
0.30	8,067	3,585	896	323
0.31	8,217	3,652	913	329
0.32	8,359	3,715	929	334
0.33	8,493	3,775	944	340
0.34	8,620	3,831	958	345
0.35	8,739	3,884	971	350
0.36	8,851	3,934	983	354
0.37	8,954	3,980	995	358
0.38	9,050	4,022	1,006	362
0.39	9,139	4,062	1,015	366

P	$MOE(\hat{P})$:			
	1%	**1.5%**	**3%**	**5%**
0.40	9,220	4,098	1,024	369
0.41	9,292	4,130	1,032	372
0.42	9,358	4,159	1,040	374
0.43	9,415	4,185	1,046	377
0.44	9,465	4,207	1,052	379
0.45	9,508	4,226	1,056	380
0.46	9,542	4,241	1,060	382
0.47	9,569	4,253	1,063	383
0.48	9,588	4,261	1,065	384
0.49	9,600	4,267	1,067	384
0.50	9,604	4,268	1,067	384

It can be seen from Eq. (2) that sample size requirements for estimates of P depend on the value of $Deff_o$. Experience sheds some light on the magnitude of $Meff_{Wts}$ in cross-class subgroup estimation (Gabler, 1999), with findings between 3.0 and 8.0. However, $Meff_{Wts}$ within class or stratum estimation is expected to be near 1.0 because our sampling strategy employs a two-phased approach of selecting MPNs with stratified sampling disproportionate to the MP user population. It is unlikely that stratum-level $Meff_{Wts}$ will be of the magnitude of the overall $Meff_{Wts}$. Therefore, the stratum-level is noted $Meff_{Wts}$ as $Meff_S$ with a range of 1.25 to 1.75 for instruction on determining sample size.

Sample sizes should be sufficient to produce behavioral health risk estimates of acceptable precision for the overall population of each participating country and for key population subgroups defined by each of the second-phase strata. Given these requirements and the need to control the variation of respondent sample weights to control losses in the precision of survey estimates, the overall respondent sample size would be determined as follows. Based on the agreed level of $MOE(\hat{P})$ in defining precision requirements, determine *n* for the smallest stratum in the general population. While it is generally recommended to limit $MOE(\hat{P})$ to 5% or less, flexibility is also recommended to facilitate feasibility. For proportional sample distribution sampling, the required overall respondent sample size would be *n* divided by the proportion of the general population in the smallest stratum. This approach to determining sample size targets for the overall sample and each sampling stratum has the advantage of ensuring the stated precision requirements based on the size of $MOE(\hat{P})$ will be met or exceeded for each of the sampling strata and for the overall population estimates. A potential disadvantage of this approach is that the required overall respondent sample size could be quite large when the smallest stratum is extremely small.

8.2 Example Sample Size Computation—Respondent Sample Size and Number of MPNs

To illustrate the sample size computations mentioned in the previous section, suppose a country assumes the following about estimation parameters in **Table 3**:

Table 3. Example sample size calculation with no stratification resulting in MeffWts = 1.00

Rate	Comment	Assumption
$Meff_{CS}$	Will always be 1.00 due to use of RDD	1.00
$Meff_{Wts}$	Increase in variance because of sample variation	1.00
MOE($\hat{P}$)	Margin of error	5%
P	Prevalence of NCD risk factor	50%

To estimate the sample size for a survey with the assumptions listed above we use the following formula from **Eq.2**:

$$n = \frac{(1.96)^2 P(1-P)}{\left[MOE(\hat{P})\right]^2} * [Deff_o] = \frac{(1.96)^2 P(1-P)}{\left[MOE(\hat{P})\right]^2} * [Meff_{Wts} * Meff_{CS}]$$

For this example, the base respondent sample size (*n*) is 384:

$$n = \frac{(1.96)^2 0.5(1-0.5)}{0.05^2} * [Deff_o] = \frac{(1.96)^2 0.5(1-0.5)}{0.05^2} * [Meff_{Wts} * Meff_{CS}] = 384 * 1.0 * 1.0$$

In the Mobile Phone Survey, some sort of stratification will be employed. This affects the respondent sample size in two ways. First, stratum-level estimates mean that each stratum should be at least large enough to estimate *P* based on the assumptions for margin of error and $Meff_{Wts}$. Second, because of stratification, it should be considered that $Meff_s$ (stratum-level $Meff_{Wts}$) will be greater than 1.00, **as discussed in section 8.1**.

Considering a range of $Meff_s$ from 1.25 to 1.75, a range of respondent sample sizes can be seen for prevalences between 10% and 50% (**Table 4**).

Table 4. Respondent sample size (n_i) for smallest stratum based on various prevalences

	Base sample size ($Meff_{Wts}$ & Meffs = 1.00)		
Prevalence	***$Meff_s$ = 1.00***	***$Meff_s$ = 1.25***	***$Meff_s$ = 1.75***
50%	384	480	672
40%	369	462	646
30%	323	404	566
20%	246	308	432
10%	138	174	242

Therefore, the respondent sample size for the *smallest* stratum in this survey is 384 for a prevalence of 50% when $Meff_s$ = 1.00. Consider the distribution of the population age and sex strata shown in **Table 5**.

Table 5. Example population distribution by age and sex (π_{ij})

Age Groups	Males	Females
18-29	π_{11} = 19.7%	π_{12} = 22.4%
30-44	π_{21} = 14.9%	π_{22} = 15.1%
45+	π_{31} = 14.8%	π_{32} = 13.0%

Source: Population of Bangladesh - United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2012 Revision.

The smallest stratum for this population is females 45 years of age and older. The sample size for each stratum is determined as follows, assuming the smallest stratum sample size is 384, for a proportional (to the overall population) stratified overall sample (**Table 6**):

Table 6. Multiplicative increase of minimum respondent sample size (n_i) by age and sex for example population distribution

Meffs	**Prevalence**	**50%**		**40%**		**30%**		**20%**		**10%**	
	Age	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
	18-29	580	664	557	638	488	559	371	426	208	239
	30-44	442	449	424	432	371	378	283	288	159	161
1.00	**45+**	438	384	421	369	368	323	280	246	157	138
	Total Respondent Sample Size	2,957		2,841		2,487		1,894		1,062	
	18-29	725	830	698	799	428	699	465	533	263	301
	30-44	552	562	531	541	430	473	354	360	200	204
1.25	**45+**	547	480	527	462	477	404	351	308	198	174
	Total Respondent Sample Size	3,696		3,558		2,911		2,371		1,340	
	18-29	1,015	1,163	975	1,118	855	979	652	747	365	419
	30-44	773	786	743	756	651	662	497	505	278	283
1.75	**45+**	766	672	736	646	645	566	492	432	276	242
	Total Respondent Sample Size	5,175		4,974		4,358		3,325		1,863	

Based on these multiplicative increases, the final respondent sample sizes (*n*) for this survey are as follows in **Table 7.**

Table 7. Overall sample sizes based on various prevalences and $Meff_s$ for an example population distribution

Age Groups	Males	Females
18-29	$\pi_{11}/\pi_{32} = 151\%$	$\pi_{12}/\pi_{32} = 173\%$
30-44	$\pi_{21}/\pi_{32} = 115\%$	$\pi_{22}/\pi_{32} = 117\%$
45+	$\pi_{31}/\pi_{32} = 114\%$	$\pi_{32}/\pi_{32} = 100\%$

The total respondent sample size (*n*) for this survey would be between 2,957 and 5,175 to estimate a prevalence of 50%, depending on the impact of $Meff_S$.

8.3 Final Disposition Codes

The data file will include **final disposition codes** for all sample cases recorded during data collection. Each mobile phone number interview case will have only **one final result code**, which is derived from disposition states automatically generated depending on **the mode of data collection (*Surveda* or CATI)**. These final codes are essential for calculating and reporting **response rates**, assessing **survey completion trends**, and ensuring **data quality** through standard quality assurance measures.

Final disposition codes may serve multiple functions, including:

- **Classifying cases** as completed, partial, non-response, or ineligible
- **Tracking survey efficiency**, such as identifying common non-response reasons
- **Ensuring standardization** of survey outcomes across different sample groups

In some instances, the **final disposition code** will directly match the **generated code** recorded during the last attempt to contact a respondent or complete recruitment. However, adjustments may be made in cases where manual review or quality control procedures require **reclassification**. A complete list of **Surveda-generated disposition codes** and their corresponding final classifications can be found in the ***Surveda Technical Manual: IVR, SMS, and Mobile Web Modes***.

8.4 Contact and Response Rates

For surveys, in general, contact, response, and eligibility rates are computed for the sample using the formulas noted below. It is frequently recommended to use these formulas so response rates and eligibility rates can be compared among countries. These rates would normally be computed for the entire sample and by stratum. The stratification variables might include age group, sex, and any other stratification variables that a country is using in its design (e.g., region).

$$\mathrm{CON1} = \frac{(\mathrm{I} + \mathrm{P}) + \mathrm{R} + \mathrm{O}}{(\mathrm{I} + \mathrm{P}) + \mathrm{R} + \mathrm{O} + \mathrm{NC} + (\mathrm{UH} + \mathrm{UO})}$$

(Eq. 4)

$$RR1 = \frac{I}{(I + P) + (R + NC + O) + (UH + UO)}$$

(Eq. 5)

$$RR2 = \frac{(I + P)}{(I + P) + (R + NC + O) + (UH + UO)}$$

(Eq. 6)

$$RR6 = \frac{(I + P)}{(I + P) + (R + NC + O)}$$

(Eq. 7)

$$ELR = \frac{I_{ELIG}}{I_{ELIG} + I_{INEL}}$$

(Eq. 8)

Note that the required definition of response rates as defined below is consistent with a standard definition of response rates as suggested by the American Association for Public Opinion Research (AAPOR). For the Mobile Phone Survey, we are using the definition of what AAPOR refers to as contact rate CON1, response rates RR1, RR2, and RR6, as well as eligibility rate ELR in the 2016 version of the AAPOR report (AAPOR, 2016).

These contact and response rates are defined as follows:

- **Contact Rate 1 (CON1)** is the contact rate that assumes all cases of indeterminate eligibility are actually eligible. It is the proportion of all cases in which some respondent was reached by the survey.
- **Response Rate 1 (RR1)** is the minimum response rate. It is the number of complete interviews divided by the number of all interviews plus the number of eligible non-interviews plus cases of unknown eligibility.
- **Response Rate 2 (RR2)** counts partial interviews as respondents, as well.

- **Response Rate 6 (RR6)** is a special case of RR2, which assumes there are no cases of unknown eligibility, which applies to Phase 2 of the D4H Mobile Phone Survey.
- **Eligibility Rate (ELR)** is the number of eligible respondents (I_{ELIG}) divided by the number of respondents with known eligibility, including respondents who are not eligible (I_{INEL}).

For definitions of the *I, P, NC, UH*, and *UO* groupings, refer to **Table 8**.

Table 8. Response Rate Definitions

Definitions	Code
1. Interview	
Complete	I
Partial	P
2. Eligible, Non-Interview	
Refusal and break-off	R
Non-contact (NC)	NC
Other (O)	O
3. Unknown Eligibility, Non-Interview	
Unknown, if housing unit	UH
Other	UO

In the Mobile Phone Survey, we must calculate response rates for each phase separately. Phase 1 response rate is defined as the number of potential respondents screened by interview, both eligible (I_{ELIG}) and ineligible (I_{INEL}) for Phase 2, divided by the number of randomly selected mobile phone numbers that were dialed, including Unknown MPN status (UH) and Unknown Refusal (UO). Note that ineligible respondents include respondents less than 18 years of age or respondents aged 18 or greater who match a stratum where the sample size has been attained and they are deemed ineligible to continue and are, therefore, rejected from participation. This response rate is RR1 for Phase 1.

This response rate is defined as follows:

$$RR1_{P1} = \frac{I_{ELIG} + I_{INEL}}{I_{ELIG} + I_{INEL} + UH + UO}$$

(Eq. 9)

Phase 2 response rate is defined as the number of interviews(complete (I) + partial (P)) divided by the number of respondents from Phase 1 who are eligible for the NCD Module of the interview, including those who provided and complete (I) or partial (P) interview, and those who broke off (NC) after beginning Phase 2. In the D4H Mobile Phone Survey, it is impossible to obtain *R* (eligible refusals/break-offs) and *O* (eligible other non-interview) because eligibility should be obtained **before** those events occur. In the Mobile Phone Survey, eligibility is obtained after these events would occur. As a result, RR6 response rate is reduced to:

$$RR6_{RED} = \frac{(I+P)}{(I+P)+NC}$$

(Eq. 6 and Eq. 7, reduced)

Note that because of the limitation in determining *R* and *O* from the Eligible, Non-Interview subcategories, CON1, RR2 and RR6 are reduced to **identical** definitions.

9. Computing Final Sample Weights

9.1 Overview of the Sample Weights

Sample weights are numerical measurements that are essential to producing and evaluating estimates from sample survey data. They are designed to account for the probability that each respondent came into the sample and the differential effects of nonresponse, imperfect sampling frames and other forces that affect the composition of the sample. Weights for surveys like the Mobile Phone Survey are typically computed in the following basic steps.

1. **Base Weight.** A base weight is computed for each respondent as the inverse of the (unconditional) probability that the respondent was randomly selected in the sample. This respondent selection probability is in turn determined as the product of the probabilities for the sampling phases that led to selecting the respondent. The probability of selecting a respondent depends on the probability of selecting the respondent's MPN and the probability of selection of that respondent in the second-phase strata (i.e., defined by the respondent's age and sex).

2. **Calibration.** The goal of a calibration weight adjustment is to bring weighted sums of the sample data into line with the corresponding counts in the target population. This step involves adjusting the sample to the population distribution of a set of categorical calibration variables through post-stratification (or cell weighting) to the joint or cross-classified population distribution of these variables. The adjusted weight is then multiplied by a factor that calibrates the sample to the demographic distribution defined by characteristics that are likely to correlate with key study outcomes (e.g., age and sex). These calibrated weights become the final adjusted sample weights that should be used for all analyses of the Mobile Phone Survey data in each participating country.

The weighted distribution of the final adjusted weights with respect to the demographic variables used for calibration will thereby match the population counts with respect to these variables.

As noted in **Section 4**, the probability of selection for each stage of the sample design should be retained on the final analytic file for each selected household and individual. These factors should be computed and stored, and the process followed in computing them should be carefully documented at the time that the sample in each stage of selection is chosen. This section describes in detail the steps for calculating the sampling weights for the Mobile Phone Survey.

9.2 Step 1: Base Weight

A base weight was computed for each respondent as the inverse of the overall probability that the respondent was randomly selected in the sample. This respondent selection probability was in turn determined as the result of a sequence of events that led to selecting the respondent. Selection probability is the product of the probabilities for the sampling phases that led to selecting the respondent, including:

- **Phase 1:** probability of selecting the respondent's MPN
- **Phase 2:** probability of selection of the respondent in the stratified second-phase sample, which is estimated as entry into strata based on the respondent's age and sex.

Eligibility for phase 2 is based on age and sex, which is proportionally allocated to strata based on the distribution of the general population.

9.2.1 Phase 1 (p_1) Sampling Probability for Each Potential Respondent

Phase 1 MPNs were selected via simple random sample (SRS) using random digit dialing (RDD) [Event I], therefore every MPN in Phase 1 has the same sampling probability. The overall Phase 1 sample of respondents reflected two effects: (1) the proportion of MPNs selected that have active subscriptions, which is random, and (2) the sample attrition due to non-participation, which is a stochastic process and thus not be considered part of the randomization sequence here.

p_1 = Phase 1 sampling probability = probability MPN is selected and contacted.

MPN_N = number of all possible mobile phone numbers

MPN_A = number of MPNs selected by SRS = total number of MPNs generated and dialed

p_1 = phase 1 sampling rate = MPN_A/MPN_N

9.2.2 Phase 2 (p_{2s}) Sampling Probability for Stratum s

Sampling in Phase 2 is presumed to have been achieved in each age-by-sex cell by applying Bernoulli sampling probabilities to those participants completing Phase 1. Therefore, a Phase 1 respondent is retained for participation in Phase 2 with a probability equal to its Bernoulli probability. The Bernoulli probability for a user falling in the *s*-th age-sex stratum is determined by the relative frequency of the *s*-th category in the frame population of MPN users and the *s*-th category of the distribution for the country's general population aged 18 or older. The frame population of MPN users is all MPN users in the country. Because access to this frame is unlikely, its distribution is estimated from the MPN sample with known age and sex from Phase 1.

p_{2s} = Phase 2 sampling probability = conditional probability of respondent's selection into the stratified subsample, given their selection in Phase 1

P_s = proportion of the frame population in stratum *s*, as estimated from the MPN sample with known age and sex from Phase 1. This can be estimated using the sample distribution from Phase 1 (p_s).

P_{tp_s} = proportion of the target general population in stratum *s*, as provided by the most recent Census data approved by the National Statistics Office.

$Max_s(P_{tp\_s}/P_s)$ = maximum of the ratio P_{tp_s}/P_s among the six age-sex strata.

$p_{2s} = (P_{tp\_s}/P_s)/Max_s(P_{tp\_s}/P_s)$ estimated by $(P_{tp\_s}/p_s)/Max_s(P_{tp\_s}/p_s)$

9.2.3 Combined Sampling Probability (ps) and Base Weight (ws) for Stratum s

The sampling probabilities from Phase 1 and Phase 2 are multiplied to provide the combined sampling probability of being selected for the Zambia Mobile Phone Survey. The inverse of this combined sampling probability was the Base Weight.

p_s = Combined sampling probability for stratum s = $p_1 * p_{2s}$

w_s = Base Weight for stratum s = $1/p_s$

9.3 Step 2: Calibration

The adjusted weight was then multiplied by a factor that calibrates the sample to the demographic distribution defined by characteristics that are likely to correlate with key study outcomes (e.g., age and sex). These calibrated weights are the final adjusted sample weights that were used for all analyses of the Mobile Phone Survey data.

The goal of a calibration weight adjustment is to bring weighted sums of the sample data into line with the corresponding counts in the target population. The final set of weights may be calibrated to the population distribution based on population data from a statistically superior external source (e.g., the most recent census projections). This step essentially involves adjusting the weighted sample to the population distribution of a set of categorical calibration variables through *post-stratification* (or cell weighting) to the joint or cross- classified population distribution of these variables. The necessary information is:

1. Adjustment cells defined by cross-classification of two categorical (or categorized) calibration variables generally known to be correlated with the key measures of NCD risk factors reported from survey samples. For the Mobile Phone Survey this was the age and sex strata *s*.
2. A Mobile Phone Survey respondent *i* assigned to the calibration adjustment cell in the same sex and age strata *s*.
3. The population count of those with these characteristics is $N_{s,}$ as stated in the calculation of base weights.
4. The weighted sum of the respondents assigned to the calibration adjustment cell with these characteristics is

$$w_s^{(nr)} = \sum_{i=1}^{IP_s} w_{s(i)}{}^{(nr)} ,$$

(Eq. 17)

where, $w_{s(i)}{}^{(nr)}$ is the non-response adjusted weight for respondent *I*, and IP_s is the final respondent sample size for adjustment cell/strata *s*.

5. The post-stratification adjustment for respondent *i* is computed as

$$A_s^{(cal)} = \frac{N_s}{w_s^{(nr)}}$$

(Eq. 18)

Note that $A_s^{(cal)}$ is the same for everyone in the same adjustment cell/stratum. Segments of the population that were still underrepresented in the sample after adjustment for nonresponse will have values greater than one and those that were overrepresented will have values less than one.

The final adjusted weight for respondent *i* is

$$w_i = w_{s(i)}{}^{(nr)} * A_s{}^{(cal)}$$

(Eq. 19)

9.4 Assessment of Final Weights

To assess the final sampling weights for quality we provide several summaries. These are designed to illustrate the sampling weights for the Mobile Phone Survey will support estimation of population-level characteristics. The procedures are as outlined and should be examined overall and by strata:

1. The sum of the final calibrated weights should match the population distribution used for calibration.
2. The sum of the final calibrated weights should match the population count from the source used for calibration.
3. The calibration factors should be slightly greater or less than 1.
4. The variation in the final calibrated weights should not be greater than 2, as shown by multiplicative effect, or M_{eff}.

9.4.1 Sum of Final Calibrated Weights

The sum of the final calibrated weights (w_i) should match the population distribution (N_s) used for calibration.

9.4.2 Calibration Factors

The calibration factors should be slightly greater or less than 1. Large deviations from 1 in any age by sex strata indicate under or over representation of that stratum.

9.4.2 Multiplicative Effect (M_{eff}) of Sampling Weights

The variation in the final calibrated weights should not be greater than 2, as shown by multiplicative effect, or M_{eff}.

M_{eff} is calculated as:

$$M_{eff} = 1 + \frac{s_{w_i}^2}{\overline{w}_i^2}$$

(Eq. 20)

Where $s_{w_i}^2$ is the variance of the final calibrated weights and $\overline{w}_i^2$ is the squared mean of the final calibrated weights.

10. DATA COLLECTION

Data are collected using a cloud-based platform called Surveda or via computer-assisted Telephone Interviewing (CATI), depending on the method selected. (More information can be found about Surveda in the ***Surveda Technical Manual: IVR, SMS, and Mobile Web Modes*** and for CATI in the ***MPS Computer-assisted Telephone Interviewing (CATI) Technical Manual: CATI Mode*** (for this mode). Each MPN from the RDD-generated list is assigned to one of two initial contact modes (only one if using CATI), The four possible modes are: IVR, SMS, Mobile Web, or CATI. Determination on the number of contact attempts is dependent on whether a single mode versus dual modes will be implemented. For single mode surveys, a maximum of up to three or four contact attempts is recommended spaced 26 hours apart.

For dual modes, a maximum of four to six contact attempts is made. If no contact is made using the initial mode, Surveda attempts to contact the MPN via the remaining mode, referred to as a fallback mode. The fallback mode also uses a maximum of two or three contact attempts 26 hours apart. These contact attempts include the initial attempt to begin the survey, any follow-up attempts to begin the survey, as well as re-contacts when the survey is cut-off. A summary template of the design and data collection procedures for the MPS can be found in **Table 9.**

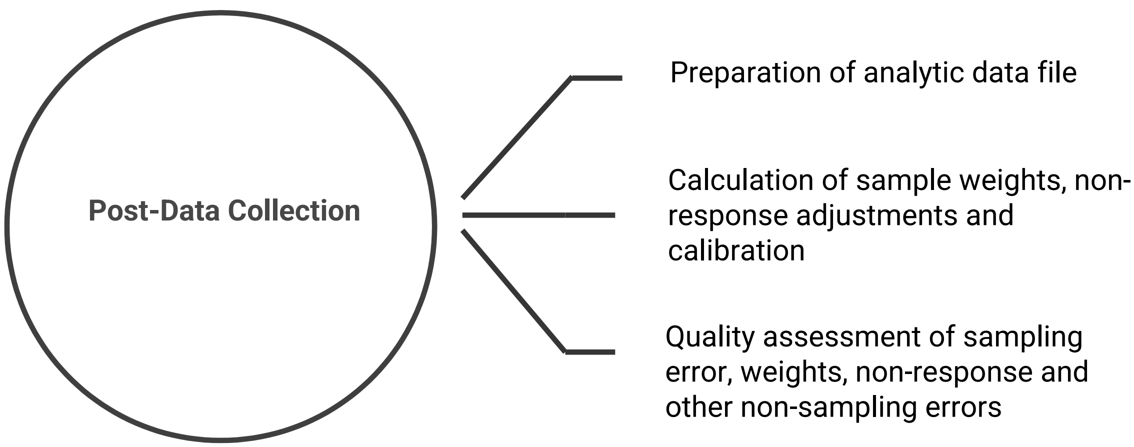
Table 9. Mobile Phone Survey Design

Component	Design
Mode(s)	IVR, SMS, Mobile web or CATI (See CATI Manual for detailed information on this)
Sample	A two-phase sample of mobile phone numbers generated via random digit dialing (RDD), using the appropriate mobile phone prefixes (based on market share) for the country, stratified by age and sex in the second phase.
Number of Interviews	(Final sample size) interviews, allocated proportionally across strata to the general population distribution.
Strata	6 strata, created by crossing sex (male, female) with age categories (18-29, 30-44, 45+)
Questionnaire	The MPS questionnaire consists of <total # of NCD module questions> core questions and administered in <total number> of languages <list languages>.
Contact times	<Number> days of the week, between <start and stop time> each day.
Contact attempts	• Contact #1: <mode used> • Contact #2: <mode used>, 26 hours after Contact #1 • Contact #3: <mode used>, 26 hours after Contact #2
Cost to Respondents	None. Incoming voice calls and SMS messages are free. The channels are configured so respondents do not incur any data charges.
Incentives	If incentives are used, every person who completed the survey was sent <insert>
Tool and Hosting	*Surveda* or CATI vendor, with data hosted <list hosting option>

11. DATA MANAGEMENT AND QUALITY ASSURANCE PROCESSES

The quality assurance process involves systematic activities that ensure and assess the quality of survey data (Biemer, 2003). High-quality data are suitable for their intended purpose and reflect multifaceted characteristics, including accuracy, timeliness, accessibility and comparability. The guidelines described in this document represent standardized procedures for the quality assurance of the Mobile Phone Survey data (see **Figure 3**).

Figure 3. Quality Assurance Diagram



11.1 Post-data Collection: Quality Assurance, Data Preparation and Cleaning Data for Sample Weight Calculations

This section describes the quality assurance guidelines and procedures used in the post-data collection phase. The post-data collection phase refers to the stage after all survey data have been collected and aggregated. It begins with the preparation of the analytic data file for data analysis and encompasses validation of variables and skip patterns, creation of the final disposition codes, preparing the data for sample weight calculations, applying non-response adjustments; assessing the quality of the sampling, sampling error, and weights; and measuring the quality of non-response, and other non-sampling errors.

11.2 Creating the Master Data File

After sample weights are calculated, a master data file is created that will contain the sample information, responses from all attempted interviews, information about mobile network operators and mode, final disposition codes and sample weights.

11.2.1 Clean and Validate the Master Data File

All respondents should have both sex and age reported. Values for age should range from 18-99. Variables are checked for valid values. Skip patterns are also checked to ensure that they worked correctly, and that all missing data are appropriately coded as missing. Data are checked for any fields that are unexpectedly blank, which could indicate an error in the skip pattern. Values are confirmed to be valid for each variable, and a codebook (data dictionary) is created to report the values and number of those values contained within each variable.

11.3 Quality Measures: Margin of Error and Item Non-Response

This section describes evaluation of the quality of estimates from the Mobile Phone Survey samples and shows the effect of unequal weights on these estimates. Guidelines to assess the performance of the calculated weights are also included.

11.3.1 Margin of Error for Key Survey Estimates

An estimate's margin of error (MOE) is one way to report the statistical precision of survey estimates. The Mobile Phone Survey recommends reporting the estimated MOE along with estimates for key survey measures. **Section 4** describes the two main features of Mobile Phone Survey samples that will influence the statistical quality of estimates and findings from the data. These features are the selection of population members with unequal probabilities (hence the need to use sample weights in analysis) and the use of stratification. A general background of these measures is provided in this section.

- **Each estimate has its own MOE.** MOE is the expected half-width of a confidence interval of an estimate of a key survey measure. MOE is interpreted as how close the estimate is likely to be to the actual survey measure in the population.
- **Estimating MOE.** The information necessary to compute an MOE is usually available. Three things are needed to compute and interpret MOE:
    - The estimate of the survey measure.
    - The estimated standard error (or variance).
    - A specified measure associated with the desired statistical confidence in the value of the estimated MOE.

The level of confidence is usually based on a value (Z) of the standard normal distribution. For example, for a 95% level of confidence, we can use $Z = 1.96$. MOE is computed as the product of the desired confidence measure and the standard error of the estimate.

- **Reporting MOE.** Key survey estimates and their associated values of MOE are recommended to be presented. This includes overall national estimates of these measures as well as estimates of these measures for all important reporting subgroups (e.g., by sex and age).
- **Interpretation.** When taken with the value of a survey estimate, MOE indicates how close the estimate is likely to be to the actual value in the population. For example,

when using *Z = 1.96* to compute estimated MOE, the survey estimate and its value of MOE can be interpreted together as follows:

- "We are 95% confident that the estimate, (VALUE OF THE ESTIMATE), is within (VALUE OF ITS MOE) of the corresponding population value."

Example: *Suppose that the reported value of a Mobile Phone Survey estimate is 22.9%, with a standard error of 1.2% that was computed in accordance with the actual sample design in that country:*

We are 95% confident that the estimate, 22.9%, is within 2.4% of the corresponding population value.

11.3.2 Item Non-response Rates for Fact Sheet Indicator Variables

For the Mobile Phone Survey, *item non-response rate* (INRR) is defined as the percentage rate of all respondents who do not answer a specific interview question among all respondents who should have answered the question. INRRs are computed for all indicators included in the country-specific Mobile Phone Survey Questions and Indicator Manual. INRRs are computed as the ratio of the number of respondents for whom an in-scope valid response was not obtained (*M*x for item x) to the total number of unit-level respondents (I) minus the number of respondents with a valid skip for item x (Vx):

$$\boldsymbol{INRR^x = \frac{M^x}{(1-V^x)} * 100}$$

(Eq. 21)

The total number of unit-level non-respondents of x (Mx) will be obtained from an unweighted frequency of respondents with missing data for item x. The total number of unit-level respondents will be obtained from the total unweighted frequency of responding males or females to the sex or age questions because these variables will have no anticipated blank fields. The total number of respondents with a valid skip for item x can be obtained as the frequency of item x with a response of REFUSED/DON'T KNOW (99).

11.3.3 Creation of Analytic Data File

After the sample weighting and all quality assurance checks have been completed, a new file will be created containing only cases with a final disposition code of Completed or Partial. This new file is called the *Analytic Data File* and should be used when conducting any data analyses. This dataset will be provided to countries as a comma delimited (csv) file, that can be opened in a variety of statistical software packages, including Excel and Stata. Technical assistance is available for the use of Excel, Stata and R.

12. GENERAL ANALYTIC GUIDELINE

12.1 Overview

The Mobile Phone Fact Sheet and Executive Summary are important documents that enable countries to present key findings and facilitate cross-country comparisons. The Country Report provides detailed results in the context of each country's unique surveys. The Fact Sheet is intended to provide an overview of the key findings and highlights of the survey for a broad audience. This document provides general data analysis used to generate both the Fact Sheet and the Executive Summary.

12.2 Reporting Point Estimates and Confidence Intervals

The Mobile Phone Survey employs a complex sampling design; therefore, analysis must account for stratification and unequal selection probabilities to obtain valid point estimates, standard errors (SEs), confidence intervals (CIs), and tests of hypotheses. If the sampling design is not accounted for, the variance may either be underestimated, which usually occurs when sampling designs include unequal probabilities of selection, or be overestimated, which can occur with stratification. It is suggested to report the weighted point estimate along with the lower and upper bound of 95% CI. The 95% CIs can be calculated based on the point estimates and their SEs (i.e., lower bound = point estimate − 1.96 × SE; upper bound = point estimate + 1.96 × SE) using appropriate methods for variance estimation of complex survey data.

12.3 Reporting Estimates in Subgroups

The suggested tables in **Appendix C** include the recommended subgroups for reporting NCDs and NCD risk factors. The variables used for classifying subgroups include the following selected demographic characteristics from the core questionnaire:

- **Sex.** Male and female
- **Age.** Three broad age groups (18–29, 30-44 and 45+), However, countries may choose to adjust subgroups based on their specific needs

Statistical tests are used to determine the significance of differences between subgroups. Differences between point estimates should be considered statistically significant if $p<0.05$.

12.4 Evaluating Missing Data

Typically, responses with "REFUSED/DON'T KNOW" are excluded from analysis for each specific indicator. See ***Questions and Indicators Manual*** for specific guidance on addressing missing data for each indicator.

Failing to identify these types of missing data or treating the assigned values for "REFUSED/DON'T KNOW" as real values will distort analysis results. Therefore, the analyst must recode them to missing values. This may be a blank, a specific code such as 99, or NA, depending on the software used for analysis.

Missing data may bias the analysis results, and some adjustments may be considered. As a general rule, if 10% or less of the data for the main outcome variable for a specific indicator are missing for eligible respondents, continuing analysis without further evaluation or adjustment is usually acceptable (Langkamp, 2010). If, however, more than 10% of the data for an indicator are missing, the analyst may need to further examine respondents and nonrespondents with respect to the main outcome variable and decide whether imputation of missing values or use of adjusted weights is necessary. Note that even if the overall item nonresponse rate is less than 10%, a subgroup item nonresponse rate within the indicator may exceed 10% and need to be further examined for statistical bias.

12.5 Reporting Small Sample Size

If an unweighted cell sample size or denominator is less than 25, it is recommended to report only unweighted data. The point estimate and 95% CI should be suppressed and replaced with a dash (—) in the cell and an explanatory footnote at the bottom of the table. For example, "— indicates an estimate based on an unweighted sample size of less than 25 and has been suppressed."

12.6 Using Statistical Analysis Software Packages

To account for the complex survey design, the sample design information should be explicitly used when producing statistical estimates or undertaking statistical analysis of the Mobile Phone Survey data. The sample weights reflect the unequal probabilities of selection, adjustments for nonresponse and adjustments to country-specific population sizes. Thus, the proper sample weight and stratification of the design must be incorporated into an analysis to obtain the correct estimates and standard errors of the estimates.

Currently, most statistical software programs do offer procedures or modules for analyzing survey data with a complex sampling design. Data from this survey can be analyzed using Stata, SAS or R. Stata offers the ***svy*** module. SAS offers ***Survey Procedures***. R offers the ***survey package***. Technical assistance is available for both Stata and R for data analysis. Note that using any statistical software based on data from simple random sample is usually not appropriate to analyze survey data with a complex design. Ignoring the complex design can lead to biased estimates and overstated significance levels (Brogan, 1998).

APPENDICES

Appendix A: Margin of Error for Key Survey Estimates

Background

Basic definitions of measures of statistical precision from survey samples are included here to help guide the analyst. An estimate of a population parameter, such as a sample mean or sample proportion, is likely to be different for different samples (of the same size) taken from the population, and each estimate is likely to be different from the true population parameter. **Sampling error** is one of two reasons for the difference between an estimate and the true but unknown value of the population parameter. The other reason is non-sampling error. Probability samples like those used in the Mobile Phone Survey enable us to produce summary measures of sampling error that indicate the precision of estimates. There are several different summary measures of precision for survey estimates. To define these measures, use the symbol P to denote the population characteristic being estimating (e.g., the prevalence rate of persons who always or often eat processed foods that are high in salt). Use $\hat{P}$ to represent the estimate of P based on the sample.

The first common summary measure of precision is the **variance of the estimate**, written as $V(\hat{P})$. All other measures are somehow related to $V(\hat{P})$. The variance of survey estimates and all other related measures are quantitative indicators of how much sample estimates would vary among all the possible samples that the sample design used in the survey could produce. A second measure of precision is the **standard error of the estimate**, defined as $SE(\hat{P}) = \sqrt{V(\hat{P})}$. The standard error of an estimate is simply the square root of the variance of that estimate and is computed in the same units as the estimate $\hat{P}$. A third measure of the statistical precision is the **relative standard error** of the estimate, defined as:

$$RSE(\hat{P}) = {SE(\hat{P})}/{\hat{P}} = {\sqrt{V(\hat{P})}}/{\hat{P}}.$$

Because $RSE(\hat{P})$ measures precision relative to the size of what is being estimated, it is unit free and, thus, a more comparable indicator of precision among estimates than the $SE(\hat{P})$. The recommended measure of precision for countries to report for key estimates in Mobile Phone Survey is the **MOE**, defined as $MOE(\hat{P}) = [Z][SE(\hat{P})]$, where Z is a measure of the level of confidence for the measure and $SE(\hat{P})$ is the standard error of $\hat{P}$. Most national estimates in the Mobile Phone Survey could use Z = 1.96 for a 95% level of confidence. The estimated value $\left(MOE(\hat{P})\right)$ reported for $\hat{P}$ is interpreted as follows:

We are 95% confident that the reported value ($\hat{P}$) is within the amount MOE($\hat{P}$) of $\hat{P}$.

Mobile Phone Survey analysts are urged to report the value of MOE($\hat{P}$) for all key estimates.

Because measures of precision can be obtained in statistical software, Mobile Phone Survey countries should request precision estimates for all their generated estimates. Note that, just as with $\hat{P}$, estimates of V($\hat{P}$), SE($\hat{P}$), MOE($\hat{P}$), and RSE($\hat{P}$) are subject to sampling error. This is important because all these estimates may be imprecise if they are based on small samples.

Data Sources

The final weighted data file used for analysis should be used for these calculations.

Computational Software

Estimates of population characteristics ($\hat{P}$) and their corresponding values of MOE($\hat{P}$) must take into account statistically important features of the sample design. Estimated characteristics must be weighted, and associated precision measures must accommodate the use of stratification, naturally occurring clustering, without-replacement sampling, and sample weights. It is well known that failure to do so typically contributes to biased estimates (particularly of precision) and, thereby, inappropriate interval estimates and tests of significance. Therefore, Mobile Phone Survey country analysts are strongly urged to use analysis software that allows one to fully account for the sample design used to produce the survey data. This means using sample weights to produce all estimates (of $\hat{P}$ and MOE ($\hat{P}$) for descriptive analysis. It also means using software that follows a widely accepted approach to estimate variances and standard errors of survey estimates.

Statistical analysis software can analyze data from complex samples like the Mobile Phone Survey. These software programs will not only produce survey estimates (i.e., $\hat{P}$), but they can also produce estimates of precision (i.e., usually either V($\hat{P}$) or SE($\hat{P}$)) that appropriately account for key design features in the Mobile Phone Survey, namely the use of stratification, and varying selection probabilities (i.e., sample weights).

Computation

Output from the software listed earlier will report a value of $\hat{P}$, as well as its estimated variance, denoted by V($\hat{P}$), or its standard error, written as SE($\hat{P}$). From these reported values, one can compute the estimate MOE for $\hat{P}$, as

$$MOE(\hat{P}) = [t][SE(\hat{P})] = [t]\sqrt{v(\hat{P})}.$$

Interpretation

The value of MOE($\hat{P}$) reported for P is interpreted as follows:

We are 95% confident that the estimated value ($\hat{P}$) is within the amount MOE($\hat{P}$) of $\hat{P}$.

Appendix B: Estimates of Sampling Errors

The respondents in the Mobile Phone Survey in [country] make up just a single sample of all the conceivable samples that could have been chosen from the same population, utilizing the sample sampling design. *Sampling errors* are a measure of the precision between every single conceivable sample. Even though the degree of precision is not known precisely, it can be evaluated from the survey data.

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

- **Value (R).** Weighted prevalence estimate of the indicator.
- **Standard error (SE).** Sampling errors are measured by the standard errors for a particular estimate or indicator. Standard error of an estimate is the square root of the variance of that estimate.
- **MOE.** Margin of error is calculated as the product of the desired confidence measure and the standard error of the estimate. The level of confidence is usually based on a value (Z) of the standard normal distribution. For example, for a 95% level of confidence, use Z = 1.96.
- **Confidence limits (R±1.96SE).** Calculated to show the interval within which the true value for the population can be reasonably assumed to fall. 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 of the statistic in 95% of all possible samples of identical size and design.

Calculation of Standard Error

The Mobile Phone Survey [year] sample is the result of a two-phase stratified design, so it is necessary to use complex formulae for estimating standard errors. For the calculation of standard errors from Mobile Phone Survey [country] data, [statistical software version] was used. Analysts can use the output from the appropriate statistical software package to obtain the standard errors, such as Stata's svy module, SAS Survey procedures or R's survey package.

The results are presented in this appendix for the country as a whole and for sex. For each variable or indicator, the type of statistic (mean, proportion, or rate) and the base population are given in **Table A.4**. In addition to the standard error (SE) described earlier, **Table A.5** includes the value of the estimate (R), the sample size, the design effect (*Deff*), the relative standard error (RSE, MOE, and the 95% confidence limits (R±1.96SE) for each variable or indicator.

REFERENCES

Biemer PP, Lyberg LE. Introduction to survey quality. Hoboken (NJ): John Wiley and Sons, Inc.; 2003.

Potter, F. Survey of procedures to control extreme sampling weights. Proceedings of the Section on Survey Research Methods, American Statistical Association; 1988. p. 453–8.

Langkamp DL, Lehman A, Lemeshow S. Techniques for handling missing data in secondary analyses of large surveys. Acad Pediatr. 2010 May-Jun;10(3):205-10.

Brogan DJ. Pitfalls of Using Standard Statistical Software Packages for Sample Survey Data.

1998. Encyclopedia of Biostatistics. P Armitage and T Colton, eds. John Wiley, New York.

Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, Inoue M. Age standardization of rates: A new WHO standard. GPE Discussion Paper Series: No. 31. EIP/GPE/EBD World Health Organization 2001. Available from URL: www.who.int/healthinfo/paper31.pdf.

Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. Epidemiology 1995 Jul;6(4):356-65.

Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol. 2003 Oct 20; 3:21.

Gabler S, Hadar S, Lahiri P (1999). A model based justification of Kish's formula for design effects for weighting and clustering. Sur Meth 1999:25:105-106.

Leo B, Morello R, Mellon J, Peixoto T, Davenport S. Do mobile phone surveys work in poor countries? World Bank, Center for Global Development, Working Paper 398, 2015, April. Available from: *www.cgdev.org/publication/do-mobile-phone-surveys-work-poor-countries-working-paper-398*.

The American Association for Public Opinion Research. Standard definitions: Final dispositions of case codes and outcome rates for surveys. 8th edition. AAPOR; 2015. Available from: www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions2015_8theditionwithchanges_April2015_logo.pdf