NCD Mobile Phone Survey Analysis Plan

NCD Mobile Phone Survey
Analysis Plan

Version 2.0
Draft | November 2016

TABLE OF CONTENTS

List of Figures

1. INTRODUCTION

1.1 Overview

Noncommunicable diseases (NCDs) are the leading cause of death worldwide. Efficient monitoring and surveillance are cornerstones to track progress of NCD 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 to address the growing NCD burden.

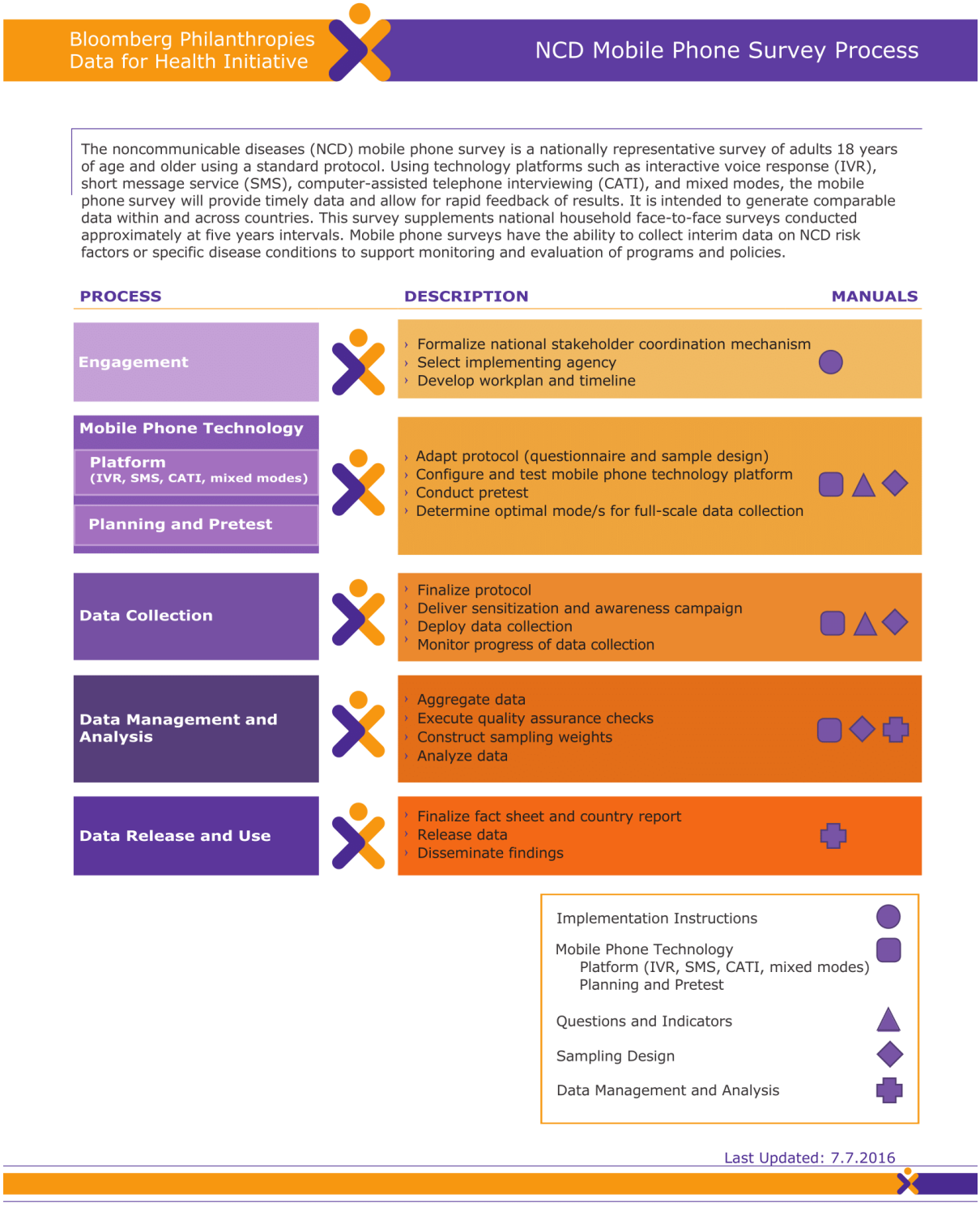
With increasing access and use of mobile phones globally, opportunities exist to explore the feasibility of using mobile phone technology as an interim method to collect data and supplement household surveys. Such technologies have the potential to allow for efficiencies in producing timely, affordable,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

The Bloomberg Data for Health 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 NCD surveillance, including the use of mobile phone surveys for NCDs. The main objectives of this component are to assess the feasibility, quality, and validity of nationally representative NCD Mobile Phone Surveys and propose a globally standardized protocol. The specific objectives are to:

- Implement mobile phone surveys in ten countries and support face-to-face STEPS surveys in six overlapping countries
- Compare findings from the two methodologies

The NCD Mobile Phone Survey is a nationally representative stratified 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 , data management procedures, and data collection using single or mixed-mode technology such as interactive voice response (IVR), short message service (SMS), and computer-assisted telephone interviewing (CATI). The implementation process consists of five stages: 1) engagement and orientation; 2) mobile phone technology and pretesting, which includes protocol adaptation and assessment; 3) data collection; 4) data management; and 5) data release and use. Details on each stage are presented in the NCD Mobile Phone Survey Process Chart (see **Figure 1**).

Figure 1. NCD Mobile Phone Survey Process Chart



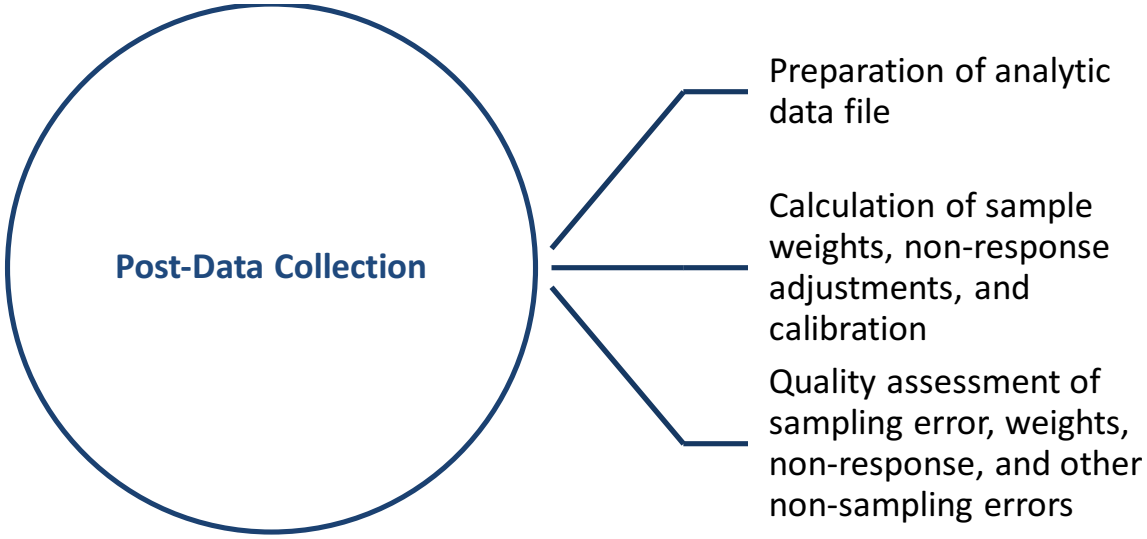
1.2 Purpose

The aim of this document is to provide guidance on evaluating and ensuring data quality and conducting analyses of the NCD Mobile Phone Survey. It describes the components of the quality assurance process that should occur after implementation of the NCD Mobile Phone Survey. Following these quality assurance guidelines are important facets of successful survey implementation. This manual also covers the general analytic guidelines on how to report the estimates, handle missing data and small sample sizes, and calculate population counts. Templates for country reports and fact sheets are included with a set of suggested tables. Lastly, guidelines for reporting results of repeated NCD survey data over time are provided.

2. 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 NCD Mobile Phone Survey data (see **Figure 2**). Countries are encouraged to add additional quality assurance activities to ensure that high-quality data are collected.

Figure 2. Quality Assurance Diagram



2.1 Quality Assurance: Post-data Collection

This section describes the quality assurance guidelines and procedures applicable to and recommended for 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 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.

2.1.1 Data Preparation and Cleaning Data for Sample Weight Calculations

This section provides guidelines on merging data files, validation of variables and skip patterns, and the creation of the final disposition codes.

Technical assistance from CDC is available for the use of the following statistical software packages for data management: EpiInfo, SAS, SPSS, R, and Stata.

2.1.2 Creating the Master Database File

Merge data into a single comma delimited file and remove mobile phone numbers to maintain participant confidentiality. The merged data file will contain the sample information, aggregated files from all interviews and attempted interviews, and the questionnaire database and data dictionary containing a description of the contents, formats, and structure of the database. A comma delimited, or CSV, file can be viewed with a Unicode-enabled text editor such as WordPad.

2.1.3 Clean and Validate the Merged Data File

Verify that variables have valid values and skip patterns worked correctly, and check on any fields that are unexpectedly blank. While many of these data quality checks are built into the data collection process, it is important to confirm that no errors were left undetected in the software programming by doing the following:

- Check that the questionnaire skip patterns have worked as specified in the questionnaire. The skip patterns from the final country questionnaire should be verified for all core variables.
- Check that skip pattern data are coded as 888. If other blank fields do exist, then output those records to an error file.
- Check each variable to ensure that invalid values are not present. Use the country's data dictionary (codebook) to confirm valid values for each variable in the dataset.
    - For example, the only valid responses for Q7 are 1 (daily), 2 (less than daily), 3 (not at all), # (refused), or the missing value code specific to the software used. If the end user prefers, simple frequencies can be run on each of the variables to see if there are any out-of-range values. Any records with invalid values can be flagged and output to the error file.
- Verify that all respondent ages range from 18 to 120 and all respondents reported a sex.

2.1.4 Assign Final Disposition Codes

The data file will include the final result codes for all the sample cases during data collection. Each mobile phone number (MPN) interview case should have one final result code and these codes are generally specific to the country or mobile network operator (MNO). Using the final result codes, assign final disposition codes. The final disposition codes are then used to calculate and report response rates and quality assurance measures.

Specifically, final disposition codes are as follows (see ***Appendix A.1*** for conversion of these result codes into final disposition codes for both IVR and SMS modes):

Table 1. Final disposition codes

1.0 Interview
- 1.10 Complete
- 1.20 Partial

2.0 Eligible, Non-interview
- 2.20 Non-contact

3.0 Unknown Eligibility, Non-interview
- 3.10 Unknown if housing unit
- 3.11 Not attempted or worked
- 3.12 Always busy
- 3.13 No answer
- 3.14 Telephone answering device (don't know if housing unit)
- 3.15 Telecommunication technological barriers (e.g., call blocking)
- 3.16 Technical phone problems
- 3.161 Ambiguous operator message
- 3.90 Other

4.0 Not Eligible
- 4.20 Fax/data line
- 4.30 Nonworking/disconnected number
- 4.31 Nonworking number
- 4.32 Disconnected number
- 4.33 Temporarily out of service
- 4.44 Pagers
- 4.50 Nonresidence
- 4.51 Business, government office, other organization
- 4.52 Institution
- 4.53 Group quarters
- 4.54 Person not household resident
- 4.70 No eligible respondent
- 4.80 Quota filled
- 4.90 Other

Refer to **Section 8.3** in the ***Sampling Design Manual*** for details on assigning the final disposition codes in addition to the calculations for response rates.

The following guidelines inform this process:

- Each case should be assigned one final disposition code for the MPN based on the final MPN result code.
- Only cases with a disposition code of 1.0 should be included in the final analytic dataset. Therefore, it is essential to assign the disposition codes correctly.
- Use cross tabs to check all final result codes against their disposition codes to identify any misclassification. If the two codes do not match as they should, it indicates a problem with the software code used to create the disposition codes.

2.2 Quality Measures: Sampling, Sampling Error, and Sample Weights

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

2.2.1 Weights Calibration Adjustments Among Cells

An important step in producing sample weights involves calibrating the weights to population counts by demonstrated relationships between key population characteristics and study outcomes, called calibration variables (e.g., sex and age, as suggested in the ***Sampling Design Manual***).

Background. Calibration adjusts for differences between the distributions of the sample and population. Adjustments are applied to all units of deliberately constructed cells. The goal is to increase the weights for those population subgroups that are underrepresented and decrease the weights for those population subgroups that are overrepresented. The more values of calibration adjustments deviate from 1.00 (high or low side), the greater the potential impact of sample imbalance on the bias of survey estimates.

Producing post-stratification adjustments. Calibration of post-stratification involves constructing adjustment cells by the cross-classification of the related characteristics. The post-stratification adjustment (PSA) in each of these adjustment cells is <1.00 if the members in that cell were overrepresented in the sample and >1.00 in those cells where the sample was underrepresented.

Reporting post-stratification adjustments. Indicate for each adjustment cell how they are defined by the variables used for calibration. For each cell, report the value of the post-stratification adjustment and its size relative to 1.00. An optimal table of these values will indicate all PSAs are close to 1.00 with some a little greater or less than 1.00.

2.2.2 Multiplicative Effect of Variable Sample Weights on the Precision of Survey Estimates

The ***Sampling Design Manual*** calls for a design where selection probabilities will vary somewhat because of potential clustering or multiplicity of MPNs requiring adjustments to sample weights. The ***Sampling Design Manual*** describes the factors that are used to adjust the sampling weights.

Background. Variation in sample weights can increase sampling error in survey estimates and, therefore, lead to larger estimates of variances/standard errors. This

multiplicative increase, referred to as $Meff_{Wts}$, depends on the degree of variability the weights are for the observations used to calculate the estimate.

Estimating $Meff_{Wts}$. The simple mean and variance of the weights are needed to compute $Meff_{Wts}$ for the data used to produce survey estimates. The value of $Meff_{Wts}$ for estimates is computed by first calculating the ratio of the variance and the square of the mean, and then adding one to this ratio.

Reporting $Meff_{Wts}$. Because $Meff_{Wts}$ applies to all estimates for a reporting domain for survey estimates (e.g., the overall population, age, sex), it should be reported for all the main population subgroups where estimates will be reported. This can be done in a table with a list of reporting subgroups and the associated values of $Meff_{Wts}$.

Interpreting $Meff_{Wts}$. Interpretation of the value of $Meff_{Wts}$ for a reporting domain is the following:

> Variation in sample weights increased the variance of all estimates (from the reporting domain) by a factor of ($Meff_{Wts}$).
>
> **Example.** Suppose that for male estimates $Meff_{Wts} = 3.0$:
>
> "Variation in sample weights increased the variance of all estimates from male respondents by a factor of 3.0."

While $Meff_{Wts}$ close to 1.0 is preferred, $Meff_{Wts} > 2.0$ might be viewed as substantial, and weight trimming or truncating strategies may be considered for outliers or extreme weights (Potter, 1988). Trimming the extreme weights can substantially reduce the overall variation in sample weights and can considerably improve the precision of the estimates. Trimming the weights, however, may also introduce some degree of bias in an estimate. The final decision as to whether one should trim the weights depends on finding a balance between the reduction in variances and the increase in bias as indicated by the gain in the mean square error. If weight trimming reduces $Meff_{Wts}$ but does not appreciably change weighted estimates for key study outcome measures, the trimming step may be justifiable.

2.2.3 Overall Design Effect on the Precision of Survey Estimates

Once the questionnaire data have been cleaned and the final sample weights have been calculated, sample data are ready to be reviewed before analysis findings are reported.

Background. The overall design effect for the estimate, or $Deff_o$, is the variance of a survey estimate from a complex sample design divided by the variance of a comparable estimate based on a simple random sample of the same size. For the

NCD Mobile Phone Survey, there is only one multiplicative component to $Deff_o$: the multiplicative of variable sample weights $Meff_{Wts}$ (see Section 2.2.2).

Estimating $Deff_o$. $Deff_o$ is estimate specific and can be reported directly with some survey analysis software packages, or it can be computed from the estimate and its variance when the estimate is a proportion or rate. Because $Deff_o$ will have many values across variables, summarize them with the median, minimum, and maximum values for reporting.

Reporting $Deff_o$. Estimates of $Deff_o$ should be reported for all key study outcomes.

Interpreting $Deff_o$. Interpretation of the estimated values and $Deff_o$ is the following:

> "The variance of the survey estimate (of the population characteristic), given the NCD Mobile Phone Survey sample design, is $Deff_o$ times greater than if simple random sampling had produced the same number of respondents."

Example. Analysis to estimate the current smoking prevalence rate produces the following from a sample where $Deff_o = 3.0$: "The variance of the survey estimate of current smoking prevalence rate, given the NCD Mobile Phone Survey sample design, is 3.0 times greater than if simple random sampling had produced the same number of respondents."

2.2.4 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 NCD Mobile Phone Survey recommends reporting the estimated MOE along with estimates for key survey measures. The ***Sampling Design Manual*** describes the two main features of NCD 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.

General background and instructions on how to compute these measures are 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. Although an estimate of MOE is not usually computed by survey analysis software, the information necessary to compute it 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 should be presented together. 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 NCD 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.

2.3 Quality Measures: Coverage, Non-response, and Other Non-sampling Errors

2.3.1 Patterns of Respondent Cutoff Rates

There will likely be some NCD Mobile Phone Survey interviews that are not complete. Respondents may decide they no longer wish to continue the interview and hang up, or a call may be interrupted because of a network issue. In either case, a partially completed interview is an indicator of respondent disengagement, which may be seen as a reflection on the respondents' attitudes toward the survey and potentially the quality of the data. The ***Sampling Design Manual*** contains a detailed discussion of respondent cutoff/response rates.

Data sources. The data file with final disposition codes should be used for these calculations. A disposition code of 1.0 indicates that the respondent completed at

least the demographic questions and one NCD question of the NCD Mobile Phone Survey interview. A disposition code of 3.90 indicates that the respondent did not consent before the demographic questions could be completed.

Calculation. A survey *respondent* is defined as any selected individual who is assigned a final disposition code of 1.0 and 3.90. Also define a cutoff rate *(COR)* to be:

$$COR = \frac{R_c}{R} = \frac{\#\ of\ respondents\ who\ partially\ complete\ the\ interview}{\#\ of\ respondents}$$

$$= \frac{P + NC + UO}{I + P + NC + UO} = \frac{1.20 + 2.20 + 3.90}{1.10 + 1.20 + 2.20 + 3.90}$$

Example. Calculating COR

In this example, we use 238,927 as the number of MPNs needed to dial from **Section 8.2** in the ***Sampling Design Manual***. If only 20% of the numbers are active, we will reach 47,786 potential respondents. Assuming a 50% eligibility rate (i.e., 23,893 not eligible) and a 30% response rate (16,725 non-interview), we are left with 7,168 respondents. Suppose we have the distribution across the final codes as shown in **Table 2**.

Table 2. Final Disposition Code Data for Example Survey

	Code	No.
Interview	1.0	**7,168**
Complete (I)	1.10	2,151
Partial (P)	1.20	5,017
Eligible, Non-Interview	2.0	**4,516**
Non-contact (NC)	2.20	4,516
Unknown Eligibility, Non-Interview	3.0	**12,209**
Unknown if Housing Unit (UH)	3.10	1,221
Other (UO)	3.90	10,988
Not Eligible		**23,893**
Total		47,786

The cut-off rate would be calculated as follows:

$$COR = \frac{P+NC+UO}{I+P+NC+UO}$$

$$COR = \frac{5{,}017+4{,}516+10{,}988}{2{,}151+5{,}017+4{,}516+10{,}988} = 0.90512$$

Uses. Values of COR should be computed directly for the sample as a final stage of quality assurance after data collection is completed. Values of COR could also be computed by the following:

- The week of data collection in which the interview took place
- Respondent age
- Respondent sex

Interpretation. Generally, the lower the value of COR, the better. While CORs are useful measures of data collection performance and overall survey quality, they are not the most critical component.

2.3.2 Item Non-response Rates for Fact Sheet Indicator Variables

For the NCD 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 should be computed for all indicators included in the country-specific NCD Mobile Phone Survey Fact Sheet (see **Section 5**). 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 (V^x):

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

The total number of unit-level non-respondents of x (M^x) will be obtained from an unweighted frequency of respondents with missing data for item x after appropriate cleaning to ensure proper skip patterns were followed. The total number of unit-level respondents will be obtained from the total unweighted frequency of responding males or females to the gender or age questions because this variable 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 888. INRRs below 5% are considered low.

2.4 Creation of Analytic Data File

After the sample weighting and all quality assurance checks have been completed, a new file should be created containing only cases with an individual level final disposition code of 1.0. Only cases with an individual level final disposition code of 1.0 will be considered to be respondents to the NCD Mobile Phone Survey.

This new file is called the *Analytic Data File* and should be used when conducting data analyses to create indicators for the Country Fact Sheet and Country Report.

3. GENERAL ANALYTIC GUIDELINE

3.1 Overview

The NCD Mobile Phone Survey Country Report and Fact Sheet are important documents that enable countries to present key findings and facilitate cross-country comparison. The Country Report provides detailed results in the context of each country's unique NCD 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 and reporting guidelines and recommendations for the Country Report and a template for the Fact Sheet.

3.2 Reporting Point Estimates and Confidence Intervals

The NCD Mobile Phone Survey employs a complex sampling design; therefore, analysis must account for stratification, multiplicity, naturally occurring clustering, 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 clustering and unequal probabilities of selection) or be overestimated (which can occur with stratification and multiplicity). 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. The commonly used variance estimation methods supported by statistical software for a two-phase sample design are Successive Difference Replication (SDR) and model-assisted estimation.

Currently, only two statistical software packages support these variance estimation methods, Stata and R. Stata employs SDR as the option *sdr* in the *svy* module. R offers the *twophase* function of the *survey* package, as well as within the functionality of the *Design* package.

Reporting Estimates in Subgroups

The suggested tables in **Section 6** 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.** Four broad age groups (18–29, 30-44, 45-59, and 60 years and older)

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 $p<0.05$.

3.3 Evaluating Missing Data

Typically, responses with "don't know" or "refused" are excluded from analysis for each specific indicator. See ***Questions and Indicators Manual*** for specific guidance on addressing missing data for each indicator.

When a sampled person refuses to answer a question, a "refused" response is assigned a value of #. A "don't know" response is assigned a value of 888. Failing to identify these types of missing data or treating the assigned values for "refused" or "don't know" as real values will distort analysis results. Therefore, the analyst must recode a "refused" response to 999 and a "don't know" response to 888 as missing values.

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.

3.4 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."

3.5 Computing Population Counts

Calculating population counts in addition to the prevalence of a health risk factor is often helpful to further understand its direct public health impact or burden. Here are the basic steps to calculating a population count:

1. Estimate the unadjusted (crude) prevalence of the NCD or NCD risk factor.
2. Determine the relevant population totals from the country's census or equivalent population projection.

3. Multiply the prevalence estimate of the NCD or NCD risk factor by the corresponding country census population total to obtain an estimate of the number of country citizens with the NCD or NCD risk factor.
4. Population counts should be reported to the nearest thousand, with a 95% CI computed from the prevalence estimate and the SE.

3.6 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 NCD 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 of the statistical software programs, such SAS, SUDAAN, and SPSS, do not offer procedures or modules for analyzing survey data with a two-phase sampling design. However, data from this design can be analyzed using Stata or R. Stata offers the *sdr* option in the *svy* module. R offers the *twophase* function of the *survey* package, as well as within the functionality of the *osDesign* package. Technical assistance is available for the use of 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).

4. COUNTRY REPORT TEMPLATE

4.1 Overview

The purpose of the Country Report is to present the major survey findings in a format that is easily understood. The recommended Country Report outline (see **Figure 2**) and Country Report template below have been developed to guide reporting. Countries may choose to organize their report in a manner to fit their needs. The Country Report will contain mainly descriptive analyses, however, further analyses may be conducted to build upon the reported findings.

Figure 2. Recommended Outline for Country Report

CONTENTS

4.2 Recommended Template for Country Report

Executive Summary

Include the following:

- Key outcomes of the survey

- Recommendations

1 Introduction

Include the following:

- NCD disease and mortality rates for [Country], [Year]
- Goals for reducing NCD disease and mortality for [Country], [Year]
- The goal of the NCD Mobile Phone Survey

1.1 Burden of NCD in [Country]

Include the following:

- Prevalence of NCDs
- NCD patterns from previous surveys
- Economic impact of NCDs

1.2 Current NCD Policies in [Country]

Include the following:

- National NCD legislation
- Current NCD initiatives in [Country]

1.3 Objectives of the NCD Mobile Phone Survey

The objectives of the NCD Mobile Phone Survey are:

- To systematically monitor NCDs and their related risk factors

2 Methodology

2.1 Study Population

Include the following:

- Description of sample and target population
- Eligibility criteria

2.2 Sampling Design

Include a description of the sample design.

2.3 Questionnaire

Include the following, which may need modification to reflect the country-specific questionnaire:

The [Country] NCD Mobile Phone Survey questionnaire covers the following topics:

- Demographics
- Tobacco Use
- Alcohol Use
- Diet (Fruit and Vegetable Consumption)
- Diet (Salt Consumption)
- Blood Pressure
- Diabetes

2.4 Data Collection

Include the following:

- Implementing partners

- Training/staff
- Timelines
- Pretest
- Technology platform used in data collection
- Languages included
- Data security
- Confidentiality and consent

2.5 Statistical Analysis

Include the following:

- Methods employed for estimating standard errors
- Software used for statistical analysis

For detailed descriptions of tables described below, see **Section 5**.

3 Results

3.1 Sample and Population Characteristics

Table 3: Demographic characteristics of adults age 18 years and older, [Country], [Year].

3.2 Tobacco Use

Table 4.1: Percentage of adults aged 18 years and older who currently smoked tobacco, [Country], [Year].
Table 4.2: Number of adults aged 18 years and older (in thousands) who currently smoked tobacco, [Country], [Year].
Table 4.3: Percentage of adults aged 18 years and older who currently used smokeless tobacco, [Country], [Year].
Table 4.4: Number of adults age 18 years and older (in thousands) who currently used smokeless tobacco, [Country], [Year].
Table 4.5: Percentage of adults aged 18 years and older currently using tobacco, [Country], [Year].

3.3 Alcohol Use

Table 5.1: Alcohol consumption status for adults aged 18 years and older - Overall, [Country], [Year].

3.4 Diet

Table 6.1: Fruits and/ or Vegetable consumption for adults aged 18 years and older, [Country], [Year].
Table 6.2: Fruit consumption for adults aged 18 years and older, [Country], [Year].
Table 6.3: Vegetable consumption for adults aged 18 years and older, [Country], [Year].
Table 6.4: Percentage of adults aged 18 and older who ate less than 5 servings of fruit and/or vegetables on average per day, [Country], [Year].
Table 6.5: Number of servings of fruit and/or vegetables on average per day, [Country], [Year].
Table 6.6: Salt Consumption among adults aged 18 years and older, [Country], [Year].

3.5 Blood Pressure

Table 7.1: Prevalence of diagnosed raised blood pressure or hypertension among adults aged 18 years and older, [Country], [Year].
Table 7.2: Prevalence of current treatment among adults aged 18 years and older with diagnosed raised blood pressure or hypertension, [Country], [Year].

3.6 Diabetes

Table 8.1: Prevalence of diagnosed raised blood glucose or diabetes among adults aged 18 years and older, [Country], [Year].
Table 8.2: Prevalence of current treatment among adults aged 18 years and older with diagnosed raised blood glucose or diabetes, [Country], [Year].

4 Discussion

- Include comparison of results to expectations, theoretical considerations, and/or previous work. Explain, if possible, differences.
- Include limitations of methodology, for example, sample size, deviations from intended methodology, response rates, etc.
- Include implications for future surveys and public health policy.
- Include summary closing statement.

References

Include list of references.

Appendix A: Questionnaire

Include the country-specific NCD Mobile Phone Survey questionnaire.

Appendix B: Sample Design

Include the details of the country-specific sample design.

Appendix C: Estimation of Sampling Errors

Include estimates of sampling errors.

Appendix D: Technical and Survey Staff

Include list of country team technical and survey staff.

Appendix E: Glossary of Terms

Include indicator definitions adapted from ***Questions and Indicators***. Follow the example below.

- Percentage who currently smoke tobacco: The number of respondents who answered "Daily" and "Less than daily" to currently smoking tobacco divided by the total number of respondents surveyed
- Percentage who currently smoke tobacco daily: The number of respondents who answered "Daily" to currently smoking tobacco divided by the total number of respondents surveyed.
- Percentage who currently use smokeless tobacco: The number of respondents who answered "Daily" and "Less than daily" to currently using smokeless tobacco divided by the total number of respondents surveyed.
- Percentage who currently use smokeless tobacco daily: The number of respondents who answered "Daily" to currently using smokeless tobacco divided by the total number of respondents surveyed.
- Percentage currently using tobacco: The number of respondents who answered "Daily" and "Less than daily" to currently using either smoked or smokeless tobacco divided by the total number of respondents surveyed.

5. FACT SHEET TEMPLATE

The example Fact Sheet is a presentation of survey findings in a format that emphasizes key highlights concisely (see **Figure 3**). The template presented here provides a brief overview, summary of the methodology and highlights from the survey. For further guidance on creation of indicators refer to ***Questions and Indicators Manual***.

Figure 3. Example Fact Sheet

Fact Sheet: [COUNTRY]

NCD Mobile Phone Survey Objectives

The noncommunicable diseases (NCD) Mobile Phone Survey is a nationally representative survey of adults 18 years of age and older. The mobile phone survey will provide timely data and allow for rapid feedback of results. It is intended to generate comparable data within and across countries. This survey supplements national household face-to-face surveys conducted approximately at 5-year intervals. Mobile phone surveys have the ability to collect data on NCD risk factors or specific disease condition to support monitoring and evaluation of programs and policies.

Methodology

The NCD Mobile Phone Survey uses a global standardized methodology. It includes information on respondents' background characteristics, tobacco use, alcohol use, diet, blood pressure, and diabetes. In [COUNTRY], the survey was conducted in [YEAR] by [COUNTRY IMPLEMENTING AGENCY], under the coordination of [MOH]. [MOBILE PHONE TECHNOLOGY PLATFORM] was used to produce nationally representative data. A total of XXXXX phone numbers were sampled and a total of XXXX individual interviews were completed, with an overall response rate of XX.X%.

Highlights

Tobacco Use

- XX.X% of men, XX.X% of women, and XX.X% overall (X million adults) currently smoke tobacco.
- XX.X% of men, XX.X% of women, and XX.X% overall (X million adults) currently smoke tobacco daily.
- XX.X% of adults (X million adults) currently use smokeless tobacco.
- XX.X% of adults (X million adults) currently use smokeless tobacco daily.
- XX.X% of adults (X million adults) currently use tobacco.

Alcohol Use

- XX.X% of adults (X million adults) currently drink alcohol.
- XX.X% of adults (X million adults) engage in heavy episodic drinking (6+ drinks on any occasion in past 30 days).

Raised Blood Pressure/Hypertension

- XX.X% of adults (X million adults) were previously diagnosed with raised blood pressure or hypertension by a doctor or other health worker.
- XX.X% of adults (X million adults) currently take medication for raised blood pressure prescribed by a doctor or other health worker.

Raised Blood Glucose/Diabetes

- XX.X% of adults (X million adults) were previously diagnosed with raised blood glucose or diabetes by a doctor or other health worker.
- XX.X% of adults (X million adults) currently take medication for diabetes prescribed by a doctor or other health worker.

Diet

- Adults on average consume fruit/vegetable on X number of days in a typical week.
- Adults on average consume X number of servings of fruit/vegetable per day.
- XX.X% of adults (X million adults) consumed less than five servings of fruits and vegetables per day.
- XX.X% of adults (X million adults) always or often add salt or salty sauce to their food before eating or as they are eating.
- XX.X% of adults (X million adults) always or often add salt or salty seasoning when cooking or preparing foods in the household.
- XX.X% of adults (X million adults) always or often eat processed foods that are high in salt.

	Overall (%)	Men (%)	Women (%)
Tobacco Use			
Tobacco Smokers			
Current tobacco smokers			
Daily tobacco smokers			
Smokeless Tobacco Users			
Current smokeless tobacco users			
Daily smokeless tobacco users			
Tobacco Users			
Current tobacco users			
Alcohol Use			
Alcohol Users			
Current alcohol users			
Heavy episodic drinkers			
Raised Blood Pressure/Hypertension			
Diagnosed with raised blood pressure/ hypertension			
Currently taking medication for raised blood pressure/ hypertension			
Raised Blood Glucose/Diabetes			
Diagnosed with raised blood glucose/diabetes			
Currently taking medication for raised blood glucose/ diabetes			
Diet			
Salt Consumption			
Always or often add salt or salty sauce to food before eating or as they're eating			
Always or often add salt or salty seasoning when cooking or preparing foods			
Always or often eat processed foods high in salt			
Fruit and Vegetable Consumption			
Consume less than five servings of fruits and vegetables per day			
Fruit Consumption	**Overall (mean)**	**Men (mean)**	**Women (mean)**
Average number of days fruits are consumed			
Average number of servings of fruit consumed per day			
Vegetable Consumption			
Average number of days vegetables are consumed			
Average number of servings of fruit consumed per day			

[SPACE FOR A GRAPH WITH FEATURED INDICATORS]

6. SUGGESTED TABLES

6.1 Overview

The tables that are included here are the minimum suggested tables for inclusion in the Country Report. For most indicators, reporting of percentages is advised to facilitate comparisons between estimates. In general, it is recommended that sample sizes be presented, along with percentages and their 95% CIs. The Country Report provides detailed findings that are likely to be of interest to a variety of audiences, but not all tables will be relevant for all countries. Some countries may choose to add additional indicators from other surveys they have conducted.

The subgroups recommended for cross-tabulations are presented in the suggested tables below. In addition to these subgroups, countries may choose to examine the indicators by other relevant demographic/background characteristics. In reporting cross-tabulations, careful consideration should be given to the precision of subgroup estimates. In some cases, subgroups may be collapsed if sample sizes are insufficient. Recommendations on the minimum sample size (n) needed for reporting were presented in Section 3.4.

In addition to the recommended tables, figures may be included as an effective way to present key findings in this report. Every figure should have a caption that describes the content in a few lines. The caption should be informative and start with a sequential figure number that is used for reference elsewhere in the Country Report.

6.2 Table Symbols, Notation, and Rounding

Bold or italicized rows in tables are headers that are not meant to have accompanying statistics. General footnotes that refer to the entire table should be designated using the term "Note:" Specific footnotes should be designated using numbers (e.g., 1, 2, 3). Report percentages to one decimal place, weighted counts to the nearest 1,000, and unweighted counts as integers.

Table 3: Demographic characteristics of adults age 18 years and older, [Country], [Year]

Demographic Characteristics		Weighted			Unweighted No. of Adults
		%	(95% CI)	No.	
Overall					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				
Male					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				
Female					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				

Table 4.1: Percentage of adults aged 18 years and older who currently smoked tobacco, [Country], [Year]

	Overall		Male		Female	
Age (years)	Current tobacco smoker	Current daily tobacco smoker	Current tobacco smoker	Current daily tobacco smoker	Current tobacco smoker	Current daily tobacco smoker
	% (95% CI)		*% (95% CI)*		*% (95% CI)*	
All						
18-29						
30-44						
45-59						
60 years and older						

Table 4.2: Number of adults aged 18 years and older (in thousands) who currently smoked tobacco, [Country], [Year]

	Overall		Male		Female	
Age (years)	Current tobacco smoker	Current daily tobacco smoker	Current tobacco smoker	Current daily tobacco smoker	Current tobacco smoker	Current daily tobacco smoker
All						
18-29						
30-44						
45-59						
60 years and older						

Table 4.3: Percentage of adults aged 18 years and older who currently used smokeless tobacco

	Overall		Male		Female	
Age (year)	**Current users of smokeless tobacco**	**Current daily users of smokeless tobacco**	**Current users of smokeless tobacco**	**Current daily users of smokeless tobacco**	**Current users of smokeless tobacco**	**Current daily users of smokeless tobacco**
	% (95% CI)	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*
All						
18-29						
30-44						
45-59						
60 years and older						

Table 4.4: Number of adults age 18 years and older (in thousands) who currently used smokeless tobacco

	Overall		Male		Female	
Demographic Characteristics	**Current users of smokeless tobacco**	**Current daily users of smokeless tobacco**	**Current users of smokeless tobacco**	**Current daily users of smokeless tobacco**	**Current users of smokeless tobacco**	**Current daily users of smokeless tobacco**
All						
18-29						
30-44						
45-59						
60 years and older						

Table 4.5: Percentage of adults aged 18 years and older currently using tobacco, [Country], [Year]

Demographic Characteristics		**Current tobacco users***	**Smoke ONLY**	**Smokeless ONLY**	**Both smoked and smokeless**
		% (95% CI)	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*
Overall					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				
Male					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				
Female					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				

* Indicates users of Smoked OR Smokeless tobacco

Table 5: Alcohol consumption status for adults aged 18 years and older - Overall, [Country], [Year]

Demographic Characteristics		Alcohol in the past 30 days	Heavy episodic drinking in the past 30 days*
		% (95% CI)	*% (95% CI)*
Overall			
Age (years)	All		
	18-29		
	30-44		
	45-59		
	60 years and older		
Male			
Age (years)	All		
	18-29		
	30-44		
	45-59		
	60 years and older		
Female			
Age (years)	All		
	18-29		
	30-44		
	45-59		
	60 years and older		

*Heavy episodic drinking is defined as 6 or more standard drinks in a single drinking occasion.

Table 6.1: Fruits and/ or Vegetable consumption for adults aged 18 years and older, [Country], [Year]

Demographic	Characteristics	No consumption of fruit and/or vegetables	Mean # of servings of fruit and/or vegetables on average per day	Mean # of days fruits and/or vegetables consumed in a typical week
		% (95% CI)	*mean (95% CI)*	*mean (95% CI)*
Overall				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Male				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Female				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			

Table 6.2: Fruit consumption for adults aged 18 years and older, [Country], [Year]

Demographic	Characteristics	No consumption of fruit	Mean # of days fruit consumed in a typical week	Mean # of servings of fruit on average per day
		% (95% CI)	*mean (95% CI)*	*mean (95% CI)*
Overall				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Male				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Female				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			

Table 6.3: Vegetable consumption for adults aged 18 years and older, [Country], [Year]

Demographic	Characteristics	No consumption of vegetables	Mean # of days vegetables consumed in a typical week	Mean # of servings of vegetables on average per day
		% (95% CI)	*mean (95% CI)*	*mean (95% CI)*
Overall				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Male				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Female				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			

Table 6.4: Percentage of adults aged 18 and older who ate less than 5 servings of fruit and/or vegetables on average per day, [Country], [Year]

Age (years)	Overall	Male	Female
	% (95% CI)	*% (95% CI)*	*% (95% CI)*
All			
18-29			
30-44			
45-59			
60 years and older			

Table 6.5: Number of servings of fruit and/or vegetables on average per day, [Country], [Year]

Demographic	Characteristics	No consumption of fruit and/or vegetables	1-2 servings	3-4 servings	5+ servings
		% (95% CI)	*% (95% CI)*	*% (95% CI)*	*% (95% CI)*
Overall					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				
Male					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				
Female					
Age (years)	All				
	18-29				
	30-44				
	45-59				
	60 years and older				

Table 6.6: Salt Consumption among adults aged 18 years and older, [Country], [Year]

Demographic Characteristics		Percentage of adults who always or often:		
		Add salt or salty sauce to their food before eating or as they are eating	Add salt or salty seasoning when cooking or preparing foods in the household	Eat process foods high in salt
		% (95% CI)	*% (95% CI)*	*% (95% CI)*
Overall				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Male				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			
Female				
Age (years)	All			
	18-29			
	30-44			
	45-59			
	60 years and older			

Table 7.1: Prevalence of diagnosed raised blood pressure or hypertension among adults aged 18 years and older, [Country], [Year]

Age (years)	Overall	Male	Female
	% (95% CI)	*% (95% CI)*	*% (95% CI)*
All			
18-29			
30-44			
45-59			
60 years and older			

Table 7.2: Prevalence of current treatment among adults aged 18 years and older with diagnosed raised blood pressure or hypertension, [Country], [Year]

Age (years)	Overall	Male	Female
	% (95% CI)	*% (95% CI)*	*% (95% CI)*
All			
18-29			
30-44			
45-59			
60 years and older			

Table 8.1: Prevalence of diagnosed raised blood glucose or diabetes among adults aged 18 years and older, [Country], [Year]

Age (years)	Overall	Male	Female
	% (95% CI)	*% (95% CI)*	*% (95% CI)*
All			
18-29			
30-44			
45-59			
60 years and older			

Table 8.2: Prevalence of current treatment among adults aged 18 years and older with diagnosed raised blood glucose or diabetes, [Country], [Year]

Age (years)	Overall	Male	Female
	% (95% CI)	*% (95% CI)*	*% (95% CI)*
All			
18-29			
30-44			
45-59			
60 years and older			

7.

Appendix A:
Post-data Collection

A.1 Final Disposition Codes

Final Disposition Codes for Random Digit Dialing Telephone Surveys	Code	Conversion for Cell Phone
1. Interview	1.0	
Complete (I)	1.10	
Partial (P)	1.20	Demographic questions completed plus one NCD Question
2. Eligible, Non-interview	2.0	
Non-contact (NC)	2.20	Consented and completed Demographic questions but broke off before NCD questions began
3. Unknown Eligibility, Non-interview	3.0	
Unknown if housing unit (UH)	3.10	
Not attempted or worked	3.11	
Always busy	3.12	Phone busy or network busy/down
No answer	3.13	
Telephone answering device (don't know if housing unit)	3.14	Voicemail
Telecommunication technological barriers (e.g., call blocking)	3.15	Call blocking
Technical phone problems	3.16	Bad audio quality (i.e., static, poor reception), Unable to connect because of network issues, Breakoff by respondent due to technical difficulties before Demographic questions began
Ambiguous operator message	3.161	
Other (UO)	3.90	Breakoff before Demographic questions were complete, Pressed 3 to refuse the interview, Unable to understand language of interview, Immediate hang up, Temporarily out of service, or Part-time fax/data line, Out of coverage area
4. Not Eligible	4.0	
Fax/data line	4.20	Dedicated fax/data line
Nonworking/disconnected number	4.30	
Nonworking number	4.31	
Disconnected number	4.32	
Temporarily out of service	4.33	
Pagers	4.44	
Nonresidence	4.50	
Business, government office	4.51	
Institution	4.52	
Group quarters	4.53	
Person not household resident	4.54	
No eligible respondent	4.70	Less than 18 years old
Quota filled	4.80	

Other	4.90	Phone or SIM (subscriber identity module) card not used

A.2 Margin of Error for Key Survey Estimates

Background

Measures of statistical precision of estimates are incorporated here to assist with analysis. For example, a sample mean will be distinctive for various samples of a similar size taken from the same population, resulting in mean estimates that are all slightly different than the true population parameter. **Sampling error** is a major behind the distinction between an estimate and the true population parameter. The sampling methodology recommended as a part of the NCD Mobile Phone Survey allow analysts to measure of the precision of estimates. To characterize these measures, the symbol *P* signifies the parameter being evaluated (e.g., the prevalence rate of persons who always or often eat processed foods that are high in salt). $\hat{P}$ represents the estimate of the population parameter.

An estimate of statistical precision is the variability of the estimate, composed as V(P^). The variability of estimates is a marker of how much sample estimates would differ among all the conceivable samples drawn from the same sample design. The **standard error of the estimate** is $SE(\hat{P}) = \sqrt{V(\hat{P})}$. The **relative standard error** of the estimate is:

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

$RSE(\hat{P})$ measures accuracy with respect to effect size. As a result, it is standardized (a.k.a., without unit) which makes it useful for comparing indicators.

The measure of precision countries should report in the NCD Mobile Phone Survey is the **MOE**, characterized as $MOE(\hat{P}) = [Z][SE(\hat{P})]$, where Z is a measure of the level of certainty for the measure and $SE(\hat{P})$ is the standard error of $\hat{P}$. $MOE(\hat{P})$ is interpreted as:

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

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

Data Sources

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

Computational Software

Estimates of population parameters and variability must account for sample design. Estimates must be weighted, and incorporate the use of stratification, naturally occurring

clustering, without-replacement sampling, and sample weights. Failure to do so contributes to biased estimates as well as inappropriate interval estimates and tests of significance. Therefore, NCD 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. It implies the use of software that incorporates a widely accepted approach to variance estimation.

Two statistical organizations globally have developed computer software to analyze data from complex samples like the NCD 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 NCD Mobile Phone Survey, namely naturally occurring clustering, the use of stratification, and varying selection probabilities (i.e., sample weights). Refer to **Section 3.2** in the ***Data Management and Analysis Plan Manual*** for a description of these statistical software packages.

The NCD Mobile Phone Survey analysts provide technical assistance for use of the following statistical software packages: Stata and R.

More Information

For additional details on the statistical definitions provided earlier, refer to the *Encyclopedia of Survey Research Methods* by Paul Lavrakas. The website sponsored by the Survey Research Methods Section of the American Statistical Association also has additional software information (available from URL: *www.hcp.med.harvard.edu/statistics/survey-soft*).

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 $\hat{P}$.

A.3 Estimates of Sampling Errors

The respondents in the NCD 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. Despite the fact that 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.
- **Sample size (n).** Total number of observations used to calculate the prevalence estimate (R).
- **Design effect *(Deff).*** The design effect 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. A *Deff* value of 1.0 indicates that the sample design is as efficient as a simple random sample, while a *Deff* value above 1.0 indicates the increase in the standard error due to the use of a more complex sample design. In general, for a well-designed survey, *Deff* usually ranges from 1 to 3. It is common, however, for *Deff* to be much larger, up to 7 or 8.
- **Relative standard error (RSE).** Also known as coefficient of variation (CV), this is the ratio of the standard error to the value of the indicator.
- **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 NCD 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 NCD Mobile Phone Survey [country] data, [statistical software version] was used. The Successive Difference Replication (SDR) and model-assisted variance estimation methods should be used for survey estimates. Analysts can use the output from the appropriate statistical software package to obtain the standard errors. Currently, there are only two software packages that support a two-phase stratified design. Stata employs the SDR method with the *sdr* option in the *svy* module. R employs the model-assisted method with the *twophase* function in the *survey* package, as well as within the functionality of the *osDesign* 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.

A.4 List of Indicators for Sampling Errors, NCD Mobile Phone Survey [country] [year]

Indicator	Estimate	Base Population
Current Tobacco Smokers	Proportion	Adults ≥ 18 years old
Current Daily Tobacco Smokers	Proportion	Adults ≥ 18 years old
Current Smokeless Tobacco Users	Proportion	Adults ≥ 18 years old
Current Daily Smokeless Tobacco Users	Proportion	Adults ≥ 18 years old
Current Tobacco Users	Proportion	Adults ≥ 18 years old
Current Alcohol Drinkers	Proportion	Adults ≥ 18 years old
Heavy Episodic Drinkers	Proportion	Adults ≥ 18 years old
Average number of days fruits are consumed	Proportion	Adults ≥ 18 years old
Average number of servings of fruit consumed per day	Proportion	Adults ≥ 18 years old
Average number of days vegetables are consumed	Proportion	Adults ≥ 18 years old
Average number of servings of fruit consumed per day	Proportion	Adults ≥ 18 years old
Consume less than five servings of fruits and vegetables per day	Proportion	Adults ≥ 18 years old
Always or often add salt or salty sauce to food before eating or as they're eating	Proportion	Adults ≥ 18 years old
Always or often add salt or salty seasoning when cooking or preparing foods	Proportion	Adults ≥ 18 years old
Always or often eat processed foods high in salt	Proportion	Adults ≥ 18 years old
Previously diagnosed with raised blood pressure/hypertension	Proportion	Adults ≥ 18 years old
Currently taking medication for raised blood pressure/hypertension	Proportion	Adults ≥ 18 years old who reported they were told by doctor or other health care worker that they have raised blood pressure or hypertension
Previously diagnosed with raised blood glucose/diabetes	Proportion	Adults ≥ 18 years old
Currently taking medication for raised blood glucose/diabetes	Proportion	Adults ≥ 18 years old who reported they were told by doctor or other health care worker that they have raised blood sugar or diabetes

A.5 Example Table for Reporting Sampling Errors

Indicator	Estimate (R)	Standard Error (SE)	Sample size (n)	Design Effect (*Deff* or *Deft*)	Relative Standard Error (SE/R)	Margin of Error (MOE)	Lower 95% Confidence Limit (R-1.96SE)	Upper 95% Confidence Limit (R+1.96SE)

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.