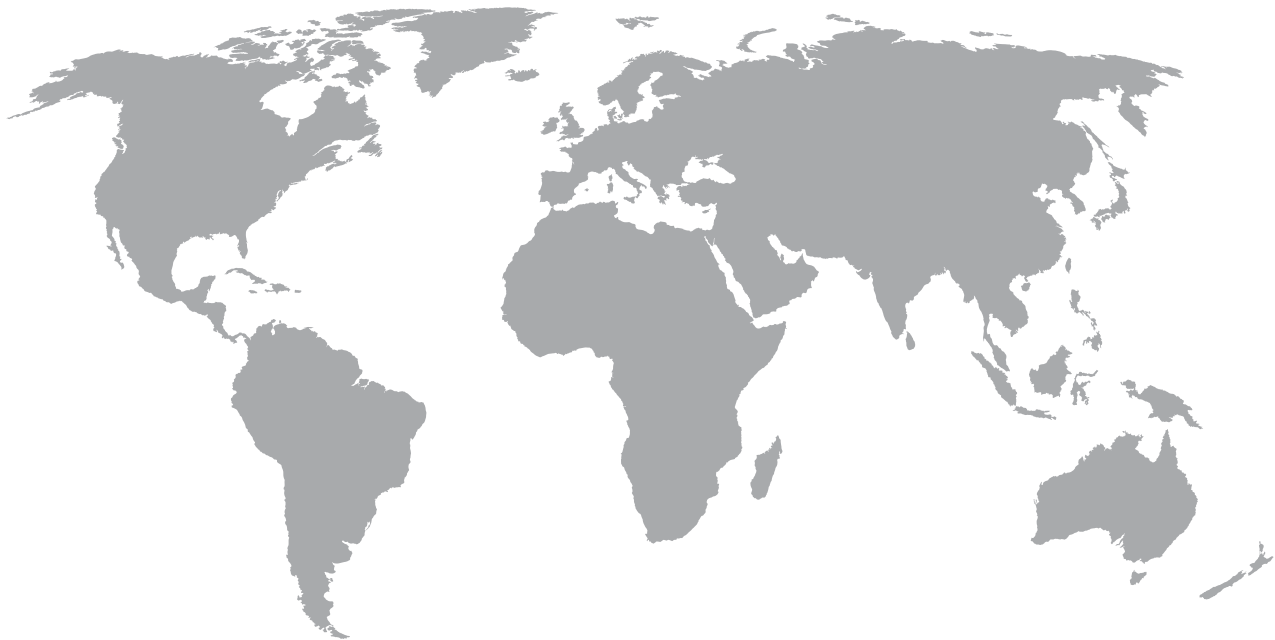




Global Adult Tobacco Survey (GATS)



Sample Design Manual

Global Adult Tobacco Survey (GATS) Sample Design Manual

September 2020

Global Adult Tobacco Survey (GATS) Comprehensive Standard Protocol

GATS Questionnaire

Core Questionnaire with Optional Questions
Question by Question Specifications

GATS Sample Design

Sample Design Manual
Sample Weights Manual

GATS Fieldwork Implementation

Field Interviewer Manual
Field Supervisor Manual
Mapping and Listing Manual

GATS Data Management

Programmer's Guide to General Survey System
Core Questionnaire Programming Specifications
Data Management Implementation Plan
Data Management Training Guide

GATS Quality Assurance: Guidelines and Documentation

GATS Analysis and Reporting Package

Fact Sheet Templates
Country Report: Tabulation Plan and Guidelines
Indicator Definitions

GATS Data Release and Dissemination

Data Release Policy
Data Dissemination: Guidance for the Initial Release of the Data

Suggested Citation

Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Sample Design Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.

Acknowledgements

GATS Collaborating Organizations

- United States Centers for Disease Control and Prevention (CDC)
- CDC Foundation
- Johns Hopkins Bloomberg School of Public Health (JHSPH)
- RTI International
- World Health Organization (WHO)

Financial Support

Financial support was provided by the *Bloomberg Initiative to Reduce Tobacco Use* through the CDC Foundation with a grant from Bloomberg Philanthropies.

Disclaimer: The views expressed in this manual are not necessarily those of the GATS collaborating organizations.

Contents

Chapter	Page
1. Introduction	1-1
1.1 Overview of the Global Adult Tobacco Survey	1-1
1.2 Use of this Manual	1-2
1.3 Deciding on a GATS Sample Design	1-3
2. Survey Design Objectives	2-1
3. Target Population and Sample Frame.....	3-1
3.1 Definitions of the GATS Target Population and Study Eligibility.....	3-1
3.2 Area Sampling Frame	3-4
3.3 Country Master Samples.....	3-4
4. Basic Survey Design Specifications	4-1
4.1 Sample Design Features.....	4-1
4.2 Sample Sizes and Expected Precision	4-3
4.3 Data Collection Procedures	4-6
5. Overview of the Sampling Approach	5-1
6. Forming Primary Sampling Units (PSUs)	6-1
7. First Stage of Sampling: Selecting PSUs	7-1
7.1 Countries Selecting Geographic Areas (i.e., Segments) in One Stage of Selection	7-1
7.2 Countries Selecting Geographic Areas in Several Stages of Selection.....	7-4
8. Selecting Households.....	8-1
8.1 Creating a Household Listing for a Sampled Segment by Visiting the Segment Prior to Selecting the Household Sample	8-1
8.2 Creating a Household Listing for a Sampled Segment Using Administrative Records	8-2
8.3 Selecting the Sample of Households	8-3
8.4 Optionally Assigning Households to the Male and Female Household Groups	8-3
8.5 Preparing an Electronic Sample File for Handheld Computer Administration	8-5
8.6 Addressing Missed Households and Other Apparent Errors in the Household Listing	8-5
9. Selecting an Eligible Resident Within Each Screened Household	9-1
10. Determining Sample Sizes at Each Stage of Selection and Reporting Final Response Rates	10-1
10.1 Sample Sizes at Each Stage of Selection.....	10-1
10.2 Example Sample Size Computation	10-4
10.3 Response Rates.....	10-5
11. Computing Final Sample Weights for Analysis	11-1

12. Bibliography	12-1
12.1 Sampling: Simple and Brief.....	12-1
12.2 Sampling: Introductory	12-1
12.3 Sampling: More Advanced Treatment	12-1
12.4 Practical Manuals	12-2
12.5 Other Papers	12-2
12.6 GATS Manuals	12-2
Appendix A: Rationale for the GATS Recommended Sample Size	A-1
Important Estimates in GATS	A-1
Power and Sample Size to Detect Change Over Time.....	A-1
Specifications, Assumptions, and Approach.....	A-1
Appendix B: Sample Design Options for New GATS Countries	B-1
OPTION N1: Stand-Alone Design.....	B-1
OPTION N2: Smaller Stand-Alone Design	B-2
OPTION N3: Integrated Design	B-3
Appendix C: Sample Design Options for Repeat GATS Countries.....	C-1
OPTION R1: Partially Overlapping Design	C-3
OPTION R2: Identical Structure but Independent Selection Design	C-4
OPTION R3: Completely Independent Selection Design	C-5
Appendix D: Survey Design Checklist.....	D-1

Exhibits

Number		Page
Exhibit 3-1.	GATS Residency Examples.....	3-3
Exhibit 5-1.	Overview of GATS Sample Design.....	5-2
Exhibit 8-1.	Summary of Process to Address Missed Households on Household Listing.....	8-6
Exhibit 10-1.	Household Questionnaire Final Disposition Codes	10-6
Exhibit 10-2.	Individual Questionnaire Final Disposition Codes	10-6

1. Introduction

Tobacco use is a major preventable cause of premature death and disease worldwide, with approximately 1.4 billion people age 15 years or older using tobacco¹. Furthermore, more than 8 million people die each year due to tobacco-related illnesses². If current trends continue, tobacco use may kill a billion people by the end of this century, and it is estimated that more than three quarters of these deaths will be in low- and middle-income countries³. An efficient and systematic surveillance mechanism is essential to monitor and manage the epidemic.

The **Global Adult Tobacco Survey (GATS)**, a component of Global Tobacco Surveillance System (GTSS), is a global standard for systematically monitoring adult tobacco use and tracking key tobacco control indicators. GATS is a nationally representative household survey of adults 15 years of age or older using a standard core questionnaire, sample design, and data collection and management procedures that were reviewed and approved by international experts. GATS is intended to enhance the capacity of countries to design, implement and evaluate tobacco control interventions.

In order to maximize the efficiency of the data collected from GATS, a series of manuals has been created. These manuals are designed to provide countries with standard requirements as well as several recommendations on the design and implementation of the survey in every step of the GATS process. They are also designed to offer guidance on how a particular country might adjust features of the GATS protocol in order to maximize the utility of the data within the country. In order to maintain consistency and comparability across countries, following the standard protocol is strongly encouraged.

GATS manuals provide systematic guidance on the design and implementation of the survey.

1.1 Overview of the Global Adult Tobacco Survey

GATS is designed to produce national and sub-national estimates among adults across countries. The target population includes all non-institutionalized men and women 15 years of age or older who consider the country to be their usual place of residence. All members of the target population will be sampled from the household that is their usual place of residence.

GATS uses a geographically clustered multistage sampling methodology to identify the specific households that Field Interviewers will contact. First, a country is divided into Primary Sampling Units, segments within these Primary Sampling Units, and households within the segments. Then, a random sample of households is selected to participate in GATS.

The GATS interview is composed of two parts: *Household Questionnaire* and *Individual Questionnaire*. These questionnaires are administered using an electronic data collection device.

¹ World Health Organization. WHO report on the global tobacco epidemic, 2019: Offer help to quit tobacco use. Geneva, Switzerland: World Health Organization; 2019. <https://apps.who.int/iris/bitstream/handle/10665/326043/9789241516204-eng.pdf?ua=1>

² GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Seattle, WA: Institute for Health Metrics and Evaluation; 2018.

³ Mathers, C.D., and Loncar D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Medicine*, 2006, 3(11):e442.

The GATS interview consists of two parts: the *Household Questionnaire* and the *Individual Questionnaire*. The *Household Questionnaire* (household screening) and the *Individual Questionnaire* (individual interview) will be conducted using an electronic data collection device.

At each address in the sample, Field Interviewers will administer the *Household Questionnaire* to one adult who resides in the household. The purposes of the *Household Questionnaire* are to determine if the selected household meets GATS eligibility requirements and to make a list, or roster, of all eligible members of the household. Once a roster of eligible residents of the household is completed, one individual will be randomly selected to complete the *Individual Questionnaire*. The *Individual Questionnaire* asks questions about background characteristics; tobacco smoking; electronic cigarettes; smokeless tobacco; cessation; secondhand smoke; economics; media; and knowledge, attitudes, and perceptions about tobacco.

1.2 Use of this Manual

This manual is designed to offer both requirements and recommendations, as well as suggested guidelines for countries to follow as they develop an appropriate sample design for their implementation of GATS. This chapter provides background information on GATS. Subsequent chapters are summarized below:

- *Chapter 2* summarizes the survey objectives.
- *Chapter 3* presents a definition of the target population for the survey and provides discussion of the area sample frame that will be used. Clearly defining the target population and sample frame is imperative. One can better understand the extent of sample coverage of the target population when a clear definition of the target population is established. This definition also will be used to determine survey eligibility for sampling and data collection. This chapter also provides a definition of a “household”⁴.
- *Chapter 4* presents the basic features of the sample design that are required to achieve between-country comparability as well as established levels of statistical quality. Other design requirements and recommendations are also presented in the following chapters.
- *Chapter 5* presents a brief overview of the recommended approach for GATS samples.
- *Chapters 6-9* present a summary of the sample selection process at each stage of the sample design.
- *Chapter 10* offers suggestions on determining an adequate sample size to select at each stage of the design.
- *Chapter 11* provides a brief overview of the process of producing a sample weight for each respondent. These weights are typically adjusted for nonresponse and calibrated to target population counts before they are used for data analysis.
- *Chapter 12* provides a list of references that might be useful to statisticians as they develop the specific GATS sample design for their country.

⁴ It should be noted that while we use the terms “household” and “housing unit” (or “dwelling unit”) somewhat interchangeably in this manual, they have somewhat different meanings. A household is a person or group of people who live together in a housing unit. Thus, housing unit refers to the physical structure (e.g., single-family home, apartment, mobile home, group of rooms, etc.) in which the person or persons comprising the household reside.

- *Appendix A* presents the rationale for the recommended overall sample size for GATS surveys.
- *Appendix B* offers a general set of “N Options” for the sample design in any of the new GATS countries after the issuance of this manual’s second edition in 2010.
- *Appendix C* offers a general set of “R Options” for countries that have completed a survey round of GATS and are about to conduct a repeat round.

1.3 Deciding on a GATS Sample Design

This manual provides a general statistical template for the probability samples that should be used for any GATS. The purpose of this section is to lay out several basic options for each of two types of design settings that are relevant at this point in the development of the GATS system of surveys in which several countries have already done a GATS and others are looking to complete a GATS for the first time. The two types of design settings are: (i) new countries hoping to conduct their first round of GATS, and (ii) countries wishing to repeat GATS by conducting another survey round.

1.3.1 Sample Design Options for Countries Conducting First Round of GATS

The design options offered to new countries are delineated according to: (i) whether or not the national sample used for GATS is statistically derived from an existing national sample (i.e., whether or not the GATS sample design is “integrated” with an existing sample design), and (ii) whether or not sample sizes need to be sufficient to (besides measuring gender differences) also make urban-rural comparisons. A GATS design that is integrated with a highly regarded existing sample and data collection infrastructure may imply lower survey cost than a stand-alone sample design (i.e., one that is developed exclusively for GATS). Moreover, a GATS sample designed only to adequately measure differences between two gender categories can be smaller than a sample for which one wishes to have adequately precise estimates for four population subgroups formed by the cross-classification of gender by urbanicity.

The following three design options considered in this setting are:

- **OPTION N1** with a standard stand-alone design and overall sample size of 8,000, which is sufficient to produce estimates for the four gender-by-urbanicity population subgroups.
- **OPTION N2** with a smaller stand-alone design and overall sample size of 4,000, which is sufficient to produce estimates for the two gender groups and the two urbanicity subgroups separately, but not for the four gender-by-urbanicity population subgroups.
- **OPTION N3** with a sample design that is integrated with a highly regarded existing national sample design and whose respondent sample size may be at the standard level (8,000) or smaller.

See *Appendix B* for more detail on these options.

1.3.2 Sample Design Options for Countries Repeating GATS

For countries repeating GATS, the focus is on how the sample for the upcoming (repeat) round relates to the sample from the initial (baseline) round. The options here are delineated according to: (i) whether or not the two samples have any planned overlap, and (ii) whether or not the structures (i.e., defined by the sampling units in each stage and stratum definitions) of the designs for the repeat and baseline samples are the same.

The following three design options considered in this setting are:

- **OPTION R1** where there is planned overlap between the samples in the baseline and repeat rounds.
- **OPTION R2** where the structures of the baseline and repeat round samples are identical but the two samples are chosen independently.
- **OPTION R3** where the structures of the baseline and repeat round samples are not identical and the two samples are chosen independently.

See *Appendix C* for more detail on these options.

2. Survey Design Objectives

Survey design requirements and recommendations for GATS have been developed so that high quality estimates can be generated for each country as a whole as well as for two analysis groups defined by urbanicity and gender.

As will be noted in *Chapter 4*, the targeted respondent sample size is typically defined so that approximately half of the respondent sample will reside in urban areas and half in rural areas. This allocation of the sample will be accomplished by explicitly stratifying the sample by urbanicity (i.e., urban vs. rural) at the earliest possible stage(s) of the design. As an option, the sample design has also been defined so that the mixture of males and females in the respondent sample can be controlled by randomly assigning selected households to male or female gender groups through a process called *gender randomization*. For example, if a household is randomly assigned to the male group, then only eligible male residents will be listed and selected from the household for interview. Similarly, if the household is randomly assigned to the female group, then only eligible female residents will be listed and selected from the household. Gender randomization is done to accomplish either or both of the following: (i) to pre-determine the gender of selected household members so that matching the gender of the field interviewer and respondent can be easily accomplished in cultures where a gender mismatch should be avoided, or (ii) to disproportionately sample males or females depending on sample size needs by gender.

Gender randomization will be optional to countries implementing GATS. Therefore, a setting will be installed in the program code to enable countries to establish whether or not gender-specific household rosters will be created in participating sampled households. The default option for this programming parameter will be to identify eligible household residents as if gender randomization had **not** been done.

Designing the survey to generate precise cross-sectional estimates by gender and urbanicity is preferred for GATS, primarily to allow comparisons of estimates by these domains between different countries participating in this effort. Some countries may have other high-priority demographic domains for reporting survey estimates. For example, a country may be interested in generating precise estimates for some geographic regions — where regions may be individual or groups of states, provinces, or any sort of geographic entities defined by landmarks, political beliefs or other cultural factors. These types of additional data requirements imposed by a country are acceptable, provided estimates of acceptable statistical precision can be produced for them and they do not compromise the statistical quality of estimates for major domains of interest in intra- and inter-country analyses.

Recommendation:
The design for the standard version of GATS seeks to obtain precise estimates jointly by urbanicity and gender.

It is worth noting that the success of generating estimates by geographic regions, generally implies that overall sample sizes must be substantially larger than if no regional estimates are required. In later sections of this manual it will be noted that the GATS sample should be designed to obtain a minimum number of 8,000 respondents per country with roughly half the sample allocated to urban areas and half to rural areas. If any country is interested in obtaining regional estimates and would like to compare estimates between urban and rural areas within each region, it is recommended that the survey be designed to obtain 8,000 respondents *in each region* (half allocated to the urban areas and half to the rural areas). If a country is interested in obtaining regional estimates and is **not** interested in comparing

estimates between urban and rural areas of the region, then it is recommended that the survey be designed to obtain 4,000 respondents *in each region*. These minimum sample sizes will generate estimates that meet the expected precision requirements of the GATS.

3. Target Population and Sample Frame

3.1 Definitions of the GATS Target Population and Study Eligibility

In general, the target population of the GATS should include all eligible individuals residing in all geographic areas within a GATS country. *In rare instances* it may be necessary to exclude some geographic areas. Areal exclusions will be considered, provided a significant proportion of the country's population does not reside in these areas. For example, extremely remote areas or areas that cannot be visited due to war, political unrest, etc. may be considered for exclusion from the GATS target population. To propose an exclusion, a country should (i) explicitly define which geographic areas of the country are included in and excluded from the target population, and (ii) provide an estimate of the percent of the population that resides in those areas being excluded. This percentage will provide an indication of the potential bias effect of any areal exclusions on estimates that are otherwise intended to be national in scope.

The target population for GATS surveys should include the civilian, non-institutionalized population of men and women, 15 years of age or older, and living in the included areas of the country, who meet GATS residency requirements both at the country and household levels. This definition is important when completing the *Household Questionnaire*, which includes a section where all eligible members of the household are listed and one is randomly chosen to complete the *Individual Questionnaire*. Eligible residents of just the designated gender group for the household should be listed if gender randomization of households is done.

GATS Target Population:
Includes all civilian, non-institutionalized men and women 15 years of age or older, who live in the included areas of the country, and who meet GATS residency requirements for the country as well as for the household where they are sampled.

Individuals, 15 years of age or older, who are explicitly **excluded** from the survey, are those who, at the time that the *Household Questionnaire* is completed, are:

- non-citizens visiting the country for a few weeks (e.g. tourists, in the country to see friends/relatives, etc.);
- citizens in the military who indicate that their usual place of residence is either on or off of a military base; OR
- citizens who are institutionalized — including people residing in hospitals, prisons, nursing homes, college residence halls, and other such institutions, who will not be sampled in GATS.

3.1.1 Country Residence Requirement for GATS

Individuals are viewed as residents of a country implementing GATS if they are: (i) citizens of and residing in the country, or (ii) non-citizens living in the country, but who consider the country to be their **usual** country of residence (i.e., they have lived in the GATS country for at least half of the time during the 12 months prior to completing the *Household Questionnaire*).

3.1.2 Household Residence Requirement for GATS

All civilian, non-institutionalized men and/or women, 15 years of age or older, who are enrolled in a school and live in a dormitory while the school is in session, but who consider the selected household to be their place of residence otherwise, automatically meet the household residence requirement for GATS¹.

All other civilian, non-institutionalized men and women, 15 years of age or older, in a sampled household meet the household residence requirement if the sampled household is considered to be their usual place of residence at the time that the *Household Questionnaire* is completed. A sampled household is the *usual place of residence* for an otherwise-eligible person living there, if that person either has no other residence, or has multiple residences but has lived at the sampled household for at least half of the time in the past 12 months.

In addition, a person who has recently moved to the sampled household to make it his/her sole residence is considered to be a member of that household if he/she does not plan to return to his/her previous household. Conversely, a person who recently moved out of the sampled household, and has no known intention of returning, is no longer considered a member of the sampled household.

Except for students living in dormitories, the usual place of residence for household members with more than one residence is the place where they spent at least half of the time in the past 12 months. Procedurally, it is important to note that if the residents randomly selected to complete an *Individual Questionnaire* are temporarily away from the household at the time of selection, every possible effort should be made to complete the interview with them at a later time. They should not be arbitrarily (or even randomly) replaced by another eligible member of the household.

Exhibit 3-1 presents some specific examples regarding application of country and household residency requirements in completing household rosters of non-institutionalized men and/or women, 15 years of age or older. It includes an indication as to whether or not they would be eligible for selection to complete an *Individual Questionnaire*.

Some non-citizens of a GATS country may meet the country residence requirement. Non-citizens qualify as “usual” residents of a GATS country if they have resided in that country for at least half of the time during the past 12 months.

All students living away in dormitories meet the household residence requirement for GATS.

A “usual” member of a sampled household is any otherwise-eligible resident who: (i) has no other residence, or (ii) has multiple residences but has been living in the selected household for at least half of the time during the past 12 months.

¹ Since school dormitories will not be included on the lists used to select households, but students temporarily living away from home are included in the GATS target population, we must consider persons living in these dormitories to be residents of the household of which they are a part off when they are not in school.

Exhibit 3-1. GATS Residency Examples

Example of person for whom country and household residency must be determined	Should the interviewer consider this person eligible to be selected for the <i>Individual Questionnaire</i> and include this person on the household roster? ²
A citizen is a student currently living in a dormitory at school, but she returns to the sampled household when school is not in session.	Yes. She meets the country eligibility criterion; and since dormitories are not sampled, the only way for this person to have a chance to be selected is to consider her to be a resident of the sampled household.
A citizen is a student and he lives in a school dormitory during the school year, but he is at home on vacation.	Yes, and for the same reason as in the previous example.
A temporary farm worker is a citizen who considers the sampled household to be “home,” but he has been working elsewhere in the country for nine of the past 12 months.	No. This person meets the country criterion since he is a citizen, but he has multiple residences and has not lived at the sampled residence more than half of the time during the past year.
A foreign citizen whose only residence has been the sampled household for the past 10 months while he completes a temporary work assignment for his foreign employer.	Yes. This person does meet the country residence criterion, since he has lived in the country for at least half of the time in the past year; he also meets the household residence criterion since this is his only residence.
A citizen has just moved into her newly purchased sole residence, which is the sampled household.	Yes. She meets the country residency criterion and intends to stay at what would be considered her only residence, thus enabling her to meet the household residence criterion.
At the time that the household roster is completed, a citizen is living in her vacation home, which was selected as the sampled household and where she has stayed two days per week in the past year.	No. She is a citizen and thus meets the country residence criterion, but she has multiple residences and has not lived at the sampled residence more than half of the time during the past year.
A citizen owns two homes and the one selected for GATS is her primary residence where she has lived all but one month during the past year; however, she is living at the other (vacation) home at the time that the roster is completed.	Yes. She is a citizen and thus meets the country residence criterion, and she has multiple residences and has lived at the sampled household more than half of the time during the past year.

In summary, at the initial visit made by the interviewer to each selected household, the interviewer will create a roster of all eligible residents who consider the selected household to be their usual place of residence at the time the roster is completed. Moreover, all eligible household residents should be included on the roster, and all reasonable efforts should be made to interview them if they are selected to complete the *Individual Questionnaire*.

The *GATS Field Interviewer Manual* will provide country-specific information on who is included and excluded from the target population, as well as a discussion of when to follow-up with selected individuals who are currently not residing in their usual place of residence.

² Assume that the reference person in each example meets all other eligibility criteria for GATS.

3.2 Area Sampling Frame

In general, the sampling frame for a survey sample is the list of all sampling units in the survey population from which the sample is to be drawn. There are two types of frames used in most face-to-face surveys of human subjects: list frames and cluster (e.g., area) frames. A list frame is a simple list of population members (e.g., a list of students in a city's schools; a list of files to be randomly selected from a room full of file cabinets). A cluster frame is a list of entities that provide indirect reference to individuals by grouping them somehow (e.g., a list of city's schools, each being a grouping of students; or a list of the room's file drawers, each containing many files). This method is called indirect because a sample of clusters must be selected first, and then the sample of individuals is identified from the individuals linked to the selected clusters.

For GATS, as with most population studies, a comprehensive list of all survey-eligible individuals in a country simply does not exist, thus making sampling from a list frame virtually impossible. Consequently, an indirect cluster sample frame must be used to select the sample. In summary, GATS samples will be drawn in several stages with an initial random selection of area clusters drawn within each country. A complete list of all households within each selected cluster will be constructed and a sample of households will be randomly selected from the list. Finally, one individual will be randomly selected from each selected household for interview. This type of design assumes that each selected survey-eligible individual can be linked to one, and only one, household in the country.

3.2.1 Definition of a "Household"

The previous section provided a formal definition of a survey individual for the GATS. Equally important is to provide a formal definition of a household. While this definition can be changed somewhat by a country, in general we suggest that countries use the definition of a household as defined by the United Nations in their report ***Principles and Recommendations for Population and Housing Censuses***. The report defines a household³ as:

"Either a one-person household, defined as an arrangement in which one person makes provision for his or her own food or other essentials for living without combining with any other person to form part of multi-person household or a multi-person household, defined as a group of two or more persons living together who make common provision for food or other essentials for living. The persons in the group may pool their incomes and have related or unrelated persons or a combination of persons both related and unrelated. This arrangement exemplifies the housekeeping concept. In an alternative definition used in many countries exemplifying the so-called household-dwelling concept, a household consists of all persons living together in a housing unit."

The complete, final definition of a household that a country plans to use should be included in the country's GATS sample design proposal, and later, in its report of GATS findings.

3.3 Country Master Samples

National statistics offices (NSOs) in most countries implementing GATS select initial samples of geographic area clusters ad hoc specifically for GATS surveys. Geographic clusters may be formed by

³ This definition of a household was taken from <https://unstats.un.org/unsd/demographic-social/sconcerns/family/#docs>

administrative units such as villages or wards or most often census enumeration areas (EAs). NSOs in some countries may preselect a large number of EAs that can provide subsamples of area clusters for first stage sampling of more than one survey or survey round. These are known as master samples. NSOs then independently subsample their master samples to select the first stage geographic clusters for GATS. Often countries create multiple replicates of independent samples from master samples that can then be selected to form survey subsamples. Depending upon the size of the replicates and that of the GATS, one or more replicate samples are then randomly drawn (*for additional information on master sample design see Chapter 5 of Household Sample Surveys in Developing and Transition Countries, 2005*). GATS sample designs must explicitly address this step of creating master samples. Probabilities of selecting initial master samples of geographical clusters must be provided as well as those of selecting GATS subsamples and sample units in all subsequent sampling stages. As few countries have utilized master samples for GATS, details on incorporating master samples into GATS sample designs and sample weighting are only provided generally in GATS manuals. As the design of subsampling from master samples varies across countries, close coordination between country statistical offices, country GATS coordinators, focal points and statisticians at CDC, and SRC members, is essential in order to ensure the accuracy of country sample designs.

4. Basic Survey Design Specifications

The sampling unit in each stage of selection refers to the entities that are selected for the survey. In this survey, the final sampling units are the household and one individual residing within the selected household. In general, the sample for GATS has been designed to be selected using a multi-stage, geographically clustered design. A multi-stage design is defined as a sample design that entails progressively selecting subsamples from a previously selected sample until the ultimate sampling units are selected. In the GATS, existing geo-political areas will be selected at the early stages of the design; households, and ultimately individuals within households, will be selected at the latter stages of the design. Additional details of the design are presented in the next sections.

As noted in *Chapter 1*, there are certain requirements and recommendations that should be followed in order to maximize the comparability of the results between countries that are conducting GATS. However, each country has the option of introducing design enhancements that would allow them to increase the usability of the results from this survey (e.g., selecting the sample to ensure precise estimates by region). In this chapter, we present some of the basic survey design requirements. Any design enhancement that a country wishes to introduce will generally be acceptable provided it does not interfere with these basic requirements. Additional specific requirements on the sample design are presented in the remaining sections of this manual.

In summary, some of the required basic survey design features of the GATS are presented in *Sections 4.1 to 4.3*.

4.1 Sample Design Features

Requirements related to the sample design include the following:

1. Random selection must be used in each sampling stage so that every member of the target population has a non-zero chance of being selected into the sample.
2. The probability of selection for every unit (household and individual) selected at each stage of the design must be known and retained on the final analytic files for the survey. For example, if the sample is selected in four stages (primary sampling unit, or PSU, selection; segment¹ selection; household selection; and individual selection) then the final analytic files must contain the corresponding four probabilities of selection. At a minimum, each country should create and maintain two files:
 - a. Household-Level File: This file will have a record for every household *selected* for GATS. Each record should include a unique household identification number; the geographic probabilities of selection (e.g. the PSU and segment probabilities of selection); and the probability of selecting the household within each selected segment, all strata variables, and cluster identifiers. An indicator of the results of gender allocation (to male or female group), along with the corresponding probability of allocation to the allocated gender group, should also be included in

REQUIREMENT:

At each stage of selection, every unit must be selected randomly and probabilities of selection at each stage must be recorded.

¹ A segment is defined as a smaller geographic area within a particular primary sampling unit. One or more intermediate sampling stages may be required to select segments within PSUs.

this file, if gender allocation is used by the country. Finally, countries may wish to consider creating a flag variable (derived from *Household Questionnaire* final disposition code) indicating whether the selected household was eligible or not for GATS, and a second flag variable indicating whether or not the household has completed the household roster.

- b. **Person-Level File:** This file will have a record for every individual selected for interview from successfully screened households. Each record should include all the variables in the household file, plus a unique individual identifier and the probability of selecting the individual. Additionally, countries may wish to consider creating a flag variable (derived from *Individual Questionnaire* disposition code) indicating whether the selected individual was eligible or not for the GATS and a second flag variable indicating whether or not, based on the final disposition code for the individual, the selected individual has completed the *Individual Questionnaire*.

3. Countries should create unique identifiers for every sampling unit selected at each stage of the design, and these identifiers must be retained in the final analytic file. For example, if a sample of individuals is selected in four stages, then the analytic file should contain unique identifiers at the PSU, segment, household, and individual levels. These identifiers are needed to verify the probability of selections at each stage of the design and may also be needed to estimate variances of the estimates produced from the survey. Except for first stage sampling unit and stratification identifiers, countries should not make these identifiers available in a public use file in order to preserve the confidentiality of each respondent in the survey.

REQUIREMENT:

Create unique identifiers for every sample unit at each selection stage.

4. As noted earlier, the sample should be selected using a multi-stage, geographically clustered sample design. Details of this type of sample design will be provided in the next sections of this manual. It will be noted in these sections that for the first stage of selection, countries should partition their country into geographic, compact, non-overlapping areas. These areas will be referred to as the primary sampling units, or PSUs, for the GATS. These areas can be formed using some natural, political, economic, or other such geographic boundaries such as states, counties, provinces, villages, or census sectors. Note that one of the main reasons for clustering in a sample design is to minimize travel requirements of the data collectors by concentrating the work in geographic pockets, or clusters.

5. In general, there should be at least 250–500 households within each PSU (some countries may have considerably more), and at least 100 PSUs should be selected at the first stage of the design. This minimum PSU sample size should be sufficient to provide adequate geographic representation of a country while still offering the desired reduction in data collection costs that are sought in a clustered design. It is important to keep the number of sample PSUs as large as possible since the smaller the PSU sample size, the larger the average respondent sample size per PSU, and the relatively less precise survey estimates will be compared to an unclustered respondent sample of the same size.

REQUIREMENT:

Select at least 100 PSUs, regardless of the overall respondent sample size.

6. The number of PSUs that a country is partitioned into at the first stage of the design should be large enough (>1,000) so that the sampling rate for the PSU selection will not be greater than about 10%.

Large geographic regions can be used to form strata, but they should not be used as PSUs. Additional information on forming PSUs is provided in *Chapter 6* of this manual.

4.2 Sample Sizes and Expected Precision

Requirements and recommendations related to other measures of sample size are based on the following indicators of statistical quality that were established for GATS findings:

1. GATS should be designed to produce estimates that meet the following precision requirements:
 - Estimates computed at the national level, by urbanicity, by gender and by the cross of gender and urbanicity should have a 95% confidence interval with a margin of error of 3 percentage points or less for tobacco use rates of 40%.
 - Sample sizes for the initial round of the GATS, as well as possibly subsequent rounds, should be sufficiently large to accommodate the following requirements for tests to detect differences between survey rounds based on independently chosen samples at each round:
 - *Tobacco use prevalence rate reduced from 40% to 34% with 80% power and to 33% with 90% power, Type I error of 0.05, and a two-sided alternative.*
 - *Secondhand exposure rate reduced from 80% to 70% with >90% power, Type I error of 0.05, and a two-sided alternative.*
 - *Quit attempt rate increased from 10% to 20% with >90% power, Type I error of 0.05, and a two-sided alternative.*
2. The **design effect** associated with any particular estimate from a survey is defined as the inflation of variance due to complex survey design features such as unequal weighting and clustering. Mathematically, it 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. While it is theoretically possible to achieve a design effect less than one, in practice the complex design features of a survey nearly always has a detrimental effect on precision of the estimates. Therefore, for most studies, the design effects will be greater than one.

Recommendation:

A respondent sample size of 2,000 is recommended for each key reporting domain to meet GATS standards of statistical quality set for the domain estimate. This corresponds to a recommended overall respondent sample size of at least 8,000 for national estimates that are to be jointly reported by gender and urbanicity. Interview at least 8,000 respondents for country-level estimates by urbanicity and gender.

Assuming a design effect of 2.00 for estimates computed at the national level, by urbanicity, by gender, and by the cross of gender and urbanicity, the minimum sample size needed to attain the GATS standards of statistical quality just described is 2,000 respondents. When applied to each of the four groups defined by the cross of urbanicity and gender, this results in a minimum recommended respondent sample of 8,000, which is set as the recommended overall sample size when the cross-classification of gender-by-urbanicity at the national level defines the key reporting domains for GATS findings. Additional detail on how this respondent sample size was computed is provided in *Appendix A*.

3. If any country is interested in obtaining regional estimates and would like to produce estimates jointly by gender and urbanicity in each region, then the survey should be designed to obtain 8,000 respondents in each region (half allocated to the urban areas and half to the rural areas). If a country is interested in obtaining regional estimates, but only wishes to report estimates separately by gender and urbanicity, then the survey should be designed to obtain 4,000 respondents in each region.
4. If any country introduces enhancements that would result in the expected design effect to generally exceed 2.00, or has empirical data that indicates that most design effects are likely to be greater than 2.00, then the sample size within the group should be appropriately adjusted upward. For example, if a country assumes a design effect of 2.5, then their design must be adjusted to yield:

$$8,000 \cdot \frac{2.50}{2.00} = 10,000 \text{ respondents.}$$

5. The design of the survey should correctly reflect anticipated levels of nonresponse and ineligibility in determining how many households must be *selected* in order to yield the recommended number of respondents indicated above. Nonresponse and ineligibility may be observed at both the household level and the selected individual level. For example, a household adult resident asked to complete a roster of eligible residents, or a person selected for interview, may refuse to participate (nonresponse). Similarly, a selected household may prove to be vacant or a selected person may indicate they are less than 15 years old and therefore ineligible. A more comprehensive definition of nonresponse and ineligibility can be found in the *GATS Field Interviewer Manual*.

Recommendation:

A respondent sample size of at least 8,000 is recommended for each region when estimates are to be reported jointly by gender and urbanicity in each region. When estimates are to be reported separately by gender and urbanicity in each region, 4,000 respondents per region is sufficient.

REQUIREMENT:

GATS sample design should address potential nonresponse and ineligibility at each stage.

If, for example, a country's survey is designed to achieve 8,000 respondents and it expects to observe the following:

Rate	Comment	Assumption
Household Eligibility Rate	Ineligible households include those that are vacant.	90%
Household Screening Rate	Accounts for households with no adult, 15 years of age or older. If gender randomization is used, this also accounts for households selected for the male household group and no males reside in the household. Similarly, this accounts for households selected for the female household group and no females reside in the household.	95%
Household Response Rate	Accounts for eligible screened household where a <i>Household Questionnaire</i> (the household roster) is not completed.	98%
Individual Eligibility Rate	In some cases, an individual might be interviewed for GATS and later determined to be ineligible; e.g., they are in fact, younger than 15 years old. This rate accounts for this.	98%
Individual Response Rate	Accounts for those eligible household residents who are selected but do not complete the <i>Individual Questionnaire</i> .	85%
Individual Response Rate for Females	An individual response rate when the household is randomly allocated to be a "female household."	90%
Individual Response Rate for Males	An individual response rate when the household is randomly allocated to be a "male household;" past studies suggest males will respond at slightly lower rates than females.	80%

Then, in this example, the survey must be designed to achieve a selected sample of

$$\frac{4,000}{.90 \cdot .98} = 4,535 \text{ selected females,}$$

$$\frac{4,000}{.80 \cdot .98} = 5,102 \text{ selected males, and}$$

$$\frac{4,535 + 5,102}{.90 \cdot .98 \cdot .95} = 11,501 \text{ selected households.}$$

Additional guidelines for determining an appropriate sample size at each stage of the GATS sample design are provided in *Chapter 10*.

4.3 Data Collection Procedures

Requirements related to the data collection procedures include the following:

1. Each country must collect the data via a face-to-face data collection method. In addition, each country should use the core part of the GATS questionnaire that is presented in the *GATS Core Questionnaire and Optional Questions* manual.
2. Each country must develop data collection procedures to achieve a *combined response rate* of 80% or greater. At each stage of the design, the response rate is defined as the total number of respondents divided by the total number of eligible individuals. The combined response rate is defined as the product of the stage-specific response rates. In GATS then, the combined response rate is defined as the product of the household and individual-level response rates. Specific information on defining response rates in GATS is presented in *Chapter 10* of this manual as well as in *GATS Quality Assurance: Guidelines and Documentation*.
3. In some studies, sample substitutions are allowed. This refers to the practice of replacing eligible nonrespondents or ineligibles with another randomly selected unit from the sample frame. Sample substitutions can lead to biased estimates and therefore are ***not allowed*** with GATS.
4. Sample supplementation refers to the practice of introducing an additional sample into the survey. This is done to increase the final respondent sample size when the levels of ineligibility and/or nonresponse are higher than anticipated in a particular stratum or area. Sample supplementation is allowed provided there is sufficient time in the data collection schedule to allow the new cases to proceed through all phases of data collection in a typical manner.
5. A maximum of one respondent per household should be selected for an interview. Since handheld computers are used for household rostering and data collection in all GATS data collection operations, the handheld device will use a random number generator to select the respondent and display the name of the chosen individual on the screen. Details on within-household sampling are presented in *Chapter 9*.

REQUIREMENT:

Obtain combined response rate of at least 80%.

REQUIREMENT:

No substitutions for eligible nonrespondents or ineligibles.

If any country implementing the GATS has questions regarding the design requirements and recommendations presented in this section, or if a country would like to request a deviation from them, it should contact CDC or WHO.

5. Overview of the Sampling Approach

The GATS sample should be selected using a multi-stage, geographically clustered design to ensure adequate coverage of the entire target population while simultaneously minimizing data collection costs. Costs are reduced because the geographic clustering feature of the sample design will lower the travel needed for interviewers to visit each selected household. Where master samples are available, country statisticians should work closely with their country CDC focal point as well as SRC members to modify the GATS protocol accordingly (see *Section 3.3* above and *Section 2: GATS Sample Weights* manual) The design outlined in *Exhibit 5-1* and detailed in this manual assumes that a country's primary sampling units are directly selected from a sampling frame that includes all geographically defined units in the country.

The first stage(s) of the design will involve selecting established geo-political area units. This selection process will be highly dependent on the particular country that is fielding GATS. Some countries may have suitable data so that only one stage of geographic selection is needed. Other countries may need multiple stages of geographic selection, where each stage is designed to randomly subsample geographic areas from within larger areas selected at a previous stage.

The ultimate goal is to select at least 100 primary sampling units and at least 400 ultimate geographic units in the survey. Ultimate geographic units, also referred to as **sample segments**, are the smallest geographic areas selected for GATS. Note that if a country is selecting geographic areas in only one stage of selection, then we would expect them to sample about 400 segments directly from their initial area frame since the segment is actually the primary sampling unit in this survey design. Also, some country sample designs may call for intermediate stages of sampling to select segments within sample PSUs.

Geographic areas should be selected at each stage with probability proportionate to size (PPS). The size measure can be constructed from a recent census and/or administrative records and, in general, should equal the total eligible population, 15 years of age or older, or the number of households in the area. If these counts are not available, corresponding total population or household counts can be used as size measures.

If regional estimates are desired and a country is interested in comparing rural areas of the region with urban areas, then a country should select at least 100 PSUs and at least 400 segments from each region with probability proportionate to size. If regional estimates are desired but a country is not interested in comparing rural areas of the region with urban areas, then at least 50 primary sampling units and at least 200 segments should be selected from each region with probability proportionate to size.

After the final phase of geographic area sampling is completed (i.e., the segment sampling stage), separate listings of all households that reside in each sample segment should be produced. These

REQUIREMENTS:

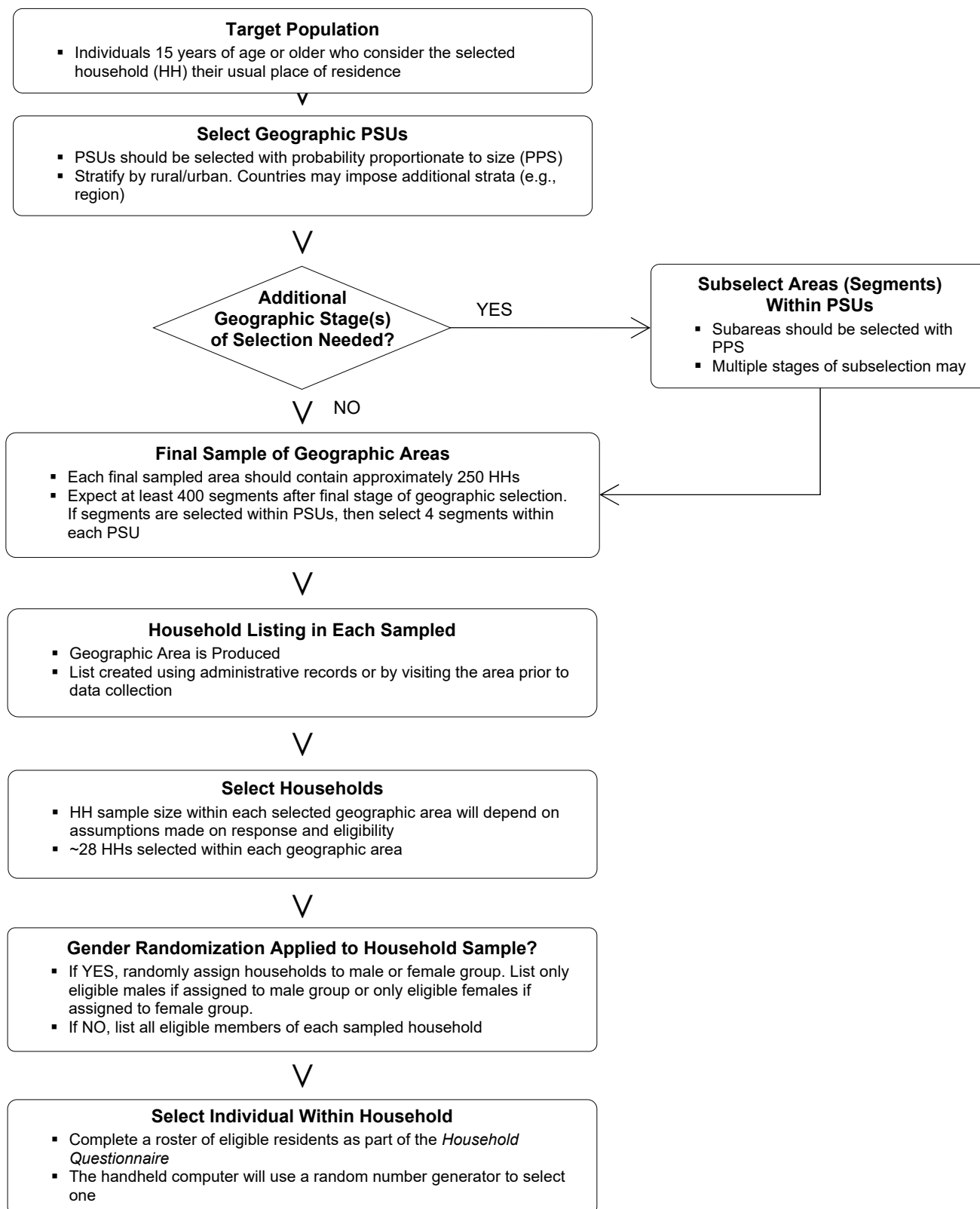
Select at least 100 geographic PSUs and 400 segments.

Select geographic areas using a random (e.g. systematic) PPS technique. Size is eligible population count (preferred) or total households.

Select at least 100 PSUs and 400 segments from each region if country wants to compare estimates by urbanicity within the region. Otherwise select 50 PSUs and 200 segments per region.

Households within segments should be randomly selected with equal probability and without-replacement.

Exhibit 5-1. Overview of GATS Sample Design



listings can be constructed using administrative records, by the field supervisor responsible for visiting selected households in the area, or by other staff specially trained in counting and listing households in a selected area. An example of the kind of household listing that should be constructed for each selected area, as well as some recommendations on how to create this list during a visit to the area, can be found in the *GATS Mapping and Listing Manual*.

The GATS protocol for listing households within selected segments differs from that of some national surveys that may have been conducted in GATS countries. The objective of our mapping and listing protocol is to identify all households in selected segments irrespective of whether the structures are currently occupied or not. Contrary to other surveys that might have been conducted in GATS countries, those listing households according to the GATS protocol are not required to talk with household members or neighbors to collect names or determine if residences or structures are occupied at the time of the listing. The household frame used to select the GATS sample of households includes all identified structures and residences in which households may reside, without consideration at the time of the listing, for whether the identified structures are currently occupied or, if not, when occupants may return. This allows time between listing and the GATS interview in which new residents can occupy selected structures. Additionally, it relieves that demand on GATS listers to communicate with household members and neighbors, placing that responsibility on interviewers who will visit selected households on up to four occasions spread over several days to better determine occupancy. Countries must take occupancy at the time of the interview into consideration when determining the number of households that will need to be selected in each segment to achieve sample size needs.

The household listing within each sample segment will serve as a sample frame for the selection of households within each selected segment. Households should be randomly selected from the list with equal probability and without-replacement. Simple random sampling or systematic sampling after a random start can be used.

After the household sample is selected in each segment, and if the country has decided to do gender randomization, each sampled household should be randomly assigned to either the male household group or the female household group. Households in the male group will have only survey-eligible males rostered during the screening visit and subsequently will have only males selected for the GATS interview. Similarly, households in the female group will have only survey-eligible females rostered during the screening visit and subsequently will have only females selected for the GATS interview. (For more on gender randomization, see *Chapter 2*.)

Interviewers will visit each selected household. If this visit reveals that additional households have emerged in an area or if the interviewer sees that the selected household is in fact a multiple household structure (such as an apartment building), then some subsampling may be required by the interviewer in consultation with the field supervisor. This type of subsampling is not likely to be needed very often.

After a listed address is selected and confirmed to be a single household, the interviewer will visit the household and create a roster of all survey-eligible males and or females (depending on if gender randomization is used) who would consider the sampled household to be their usual place of residence. (For more on criteria for residency and membership in the GATS target population, see *Section 3.1*.)

The questionnaire program on the handheld devices will use a random number generator to select one individual from within the household. Only these selected individuals will be administered the GATS *Individual Questionnaire* since no substitution for nonresponse is allowed.

It is important that countries retain sampling information at each stage of sample selection. This includes detailed descriptions of how master samples (if used), samples of PSUs, households and individuals were drawn, computational formula used to compute selection probabilities at each stage and the corresponding selection probabilities. This will allow independent verification that sampling probabilities and corresponding weights were computed correctly. This information should be submitted to the DCC soon after samples are drawn and before the data collection starts. It will provide statisticians at the DCC an early opportunity to verify the sampling process. The DCC will also act as a depository for all sampling information. See *Section 7* of this manual and the *GATS Sample Weights Manual* for the computational formulas of sample probabilities and weights using PPS selection.

A more detailed discussion of each stage of the sampling process is provided in the next few sections.

6. Forming Primary Sampling Units (PSUs)

The first task in designing the GATS is to determine the most appropriate definition of a primary sampling unit (PSU) and the associated sampling frame. This will be highly dependent on the country implementing GATS — some countries may be smaller and have recent census data, thereby enabling them to form PSUs of sufficient size for the interviewers to work directly. In other words, the sample of 400 segments can be selected in one stage of selection¹. Other countries may need to define larger PSUs and then use subsequent stages of selection to arrive at a set of geographic clusters that are of approximately equal and suitable size for the interviewers to work.

In general, the primary purpose of defining and selecting PSUs in the first stage is to reduce the costs associated with travel by the field interviewers. There are various guidelines which should be applied when forming PSUs.

PSUs should be clearly defined, compact and mutually exclusive areas defined by clear boundaries. Some countries may wish to use county or census tracts as PSUs. Examples of such entities include the counties/municipal districts or census sectors.

The number of PSUs that a country is partitioned into should be large enough (>1,000) so that the sampling rate for the PSU selection will not be greater than about 10%. Large geographic regions such as provinces or states can be used to form strata, but they **should not** be used as PSUs.

Recommendation:

Expect at least 1,000 PSUs on initial sample frame.

It is recommended that countries that have not subdivided into census sectors consider selecting the geographic areas for GATS in multiple stages. Selecting the sample in multiple stages will force the sample to be concentrated in small subareas within the country, defined by the boundaries of the PSUs that are selected at the first stage before segmentation. This will likely result in much lower data collection costs compared to having a sample of segments that are widely dispersed.

The sample of PSUs should be selected from a list of all PSUs in the country. For example, for a survey of the Russian Federation, the PSU sampling frame should include all of the Russian islands to the north and east of the Russian mainland (e.g. New Siberian Islands, Sakhalin Island), and not just the Russian mainland.

There may be some exceptions to this. Extreme remote areas or areas that cannot be visited due to war, political unrest, etc., can be excluded from the PSU frame provided a significant proportion of the country's population do not reside in those areas that are excluded. Each country should explicitly define which geographic areas of the country are included and excluded from their initial geographic frame and provide an estimate of the percent of the population that resides in those areas being excluded. This

REQUIREMENT:

Design documentation should include an estimate of the percent of country's population that is not included in the target area, if portions of the country must be excluded from GATS.

¹ If segments can be selected from a country in one stage of selection, then the segments are actually the primary sampling units or PSUs in the survey.

percentage will provide an evaluation of the potential for coverage bias in the final GATS estimates that will be generated from the country's data.

In general, the final geographic areas selected for GATS should contain approximately 250 households. If PSUs cannot be formed so that they contain only about 250 households, then the country should consider selecting the geographic areas in multiple stages of selection.

Recommendation:

Final geographic areas selected (segments) should have ~250 HHs.

An estimate of the number of members of the target population (residents, 15 years of age or older) is needed for PPS selection of PSUs. This estimate can be obtained from sources such as administrative records or the last census conducted in the country. If the estimated eligible population is not available, then a country can use some measure that is highly correlated with the survey-eligible population such as the total number of households in the area. Identifying suitable data for each PSU on the frame is important because it will be the size measure used to select geographic areas into GATS.

The complete list of PSUs must be easy to process or manipulate, i.e. sortable, for sample selection. A machine-readable form of this listing is highly preferable so that PSU selection procedures can be easily and accurately implemented.

Some countries may wish to define PSUs in a manner that significantly deviates from the guidelines suggested above. This deviation should be discussed with a country's point of contact at CDC or WHO before the sample is selected. In general, to assure the application of standardized sampling procedures, it is important that each country that administers GATS operate within the guidelines presented above and in this manual.

REQUIREMENT:

Any deviation from this plan to define PSUs should be discussed with CDC or WHO.

7. First Stage of Sampling: Selecting PSUs

The recommended sampling process for countries that will select their ultimate sample of geographic areas in one stage is somewhat different than for those countries that will select areas in several stages. Guidelines for both of these scenarios are presented below.

7.1 Countries Selecting Geographic Areas (i.e., Segments) in One Stage of Selection

Countries that plan to select ultimate geographic areas (i.e., the area segments) in only one stage of selection should follow the following guidelines:

1. Segments should be formed so that they can be reasonably worked by a small number of interviewers. In general, a segment should be formed so that it can be worked by two to four interviewers. If segments are too large, then the cost savings in the data collection process will be lost since the interviewer travel time may be substantial.
2. Each segment must be clearly designated as rural or urban.
The urbanicity designation should be used to define two explicit strata at the first stage of the design. The choice of additional explicit stratification variables is limited by the availability of classification information for each geographic segment in the sampling frame of all segments. Stratification variables associated with tobacco use are ideal choices as the variances of sample estimates are reduced when such association increases. Stratification variables are often chosen because their categories form important subpopulations for estimating key smoking variables. These are referred to as analytic “domains” of the survey. A region stratification variable would be appropriate for stratification if regional differences in tobacco use are anticipated (variance lowering effect) even if sample sizes are insufficient for regional estimation of smoking variables.

REQUIREMENT:
Stratify segment frame by urbanicity.
3. If a country is seeking to select their geographic areas in one stage of selection, then the total number of units to select should be approximately 400. In general, we expect a country will select 28–30 households per segment so this would equate to a segment sample size of roughly 400 depending on the assumptions made about nonresponse and ineligibility. Additional information on determining an appropriate segment sample size is provided in *Chapter 10*.
4. Since the size of segments used for sample design will likely vary by a considerable amount, particularly since the segments are the PSUs in this design scenario, segments should be selected with probability proportionate to a size (PPS) measure. In general, this size measure should be defined as the expected total number of survey-eligible individuals in the segment. As noted in *Chapter 6*, if these data are not available, then a country can use some size measure that is highly correlated with the number of survey-eligible individuals, such as the count of households in the area.

5. The segment sample should be selected with PPS. There are several standard techniques¹ that can be used to select the segments with PPS including the Hanuray-Vijayan algorithm that selects units without replacement [see Vijayan (1968)] and Sampford's Method that selects units without replacement [this is an extension of Brewer's Method and is discussed in Cochran (1977) and Sampford (1967)]. There are also several sequential PPS sample selection techniques that can be used that select units with-replacement or with minimal replacement (see for example, Chromy, 1979). The PPS minimal replacement, sequential selection technique is similar to a standard sequential, with-replacement PPS technique except that it ensures first stage units (in this case segments) will not be selected an inordinate number of times. Specifically, if a segment has survey-eligible people and one wishes to select segments, then this technique will select a segment with probability:

$$p_i^{(1)} = \frac{I \cdot N_i}{\sum_i N_i}$$

where N_i is the size measure, I is the number of segments chosen, and $\sum_i N_i$ is the sum of size measures for all segments. We use the superscript (1) on $p_i^{(1)}$ to indicate this is the first stage of sample selection.

In addition, with this minimal replacement selection methodology, each segment i is guaranteed to be selected either $\text{int}(p_i^{(1)})$ or $\text{int}(p_i^{(1)}) + 1$ times, where $\text{int}(p_i^{(1)})$ refers to the integer portion of $p_i^{(1)}$.

Regardless of the method used, it is important that every geographic unit on the frame has some nonzero probability of selection and that the probabilities of selection are retained in the final analysis file. Sample weights for the ultimate respondents to GATS will be created from these probabilities of selection².

REQUIREMENT:

Select segments with probability proportionate to size. Size is equal to population or HHs.

REQUIREMENT:

Every segment must have a nonzero probability of selection. Retain probabilities on final data file.

6. PPS systematic selection of PSUs is frequently used by countries conducting GATS surveys since its application is both intuitive and relatively simple. Separate PSU samples within explicitly defined strata (e.g., defined by characteristics like region, urbanicity, etc.) are chosen from an ordered frame with probabilities proportional to their size measures, typically population or household counts. The ordering of the PSUs within each explicitly defined stratum may be strategic (e.g., by physical proximity to assure broad geographic representation) or arbitrary (i.e., in a sense, random). When the former, and frame sorting criteria correlate with key survey outcome measures, PPS systematic selection contributes to reduce the variance of survey estimates. To account for this contribution in variance estimation, a common strategy first suggested by Kish (1965) for simple systematic

¹ Many of these sample selection techniques are now available in widely used commercial, statistical software packages. For example, SAS introduced the SurveySelect Procedure in SAS Version 8 that will perform many of these PPS selection algorithms.

² If the sampling rate is large (e.g. one is selecting 10% or more of a frame) then the joint probabilities of selection should also be retained for more accurate variance estimation purposes.

sampling is to pair selected PSUs following the order of selection and to specify these “pseudo stratum pairs” as part of setup in using variance estimation software for complex samples (e.g., SUDAAN, STATA, SAS, etc.). In effect, these pairs are treated as “implicit strata” within each of the explicit strata, and the total number of such strata for the entire sample will equal the total number of selected PSUs divide by two. Choosing to ignore a correlated order effect of PPS systematic sampling by only specifying the explicit strata for variance estimation will likely result in an overstatement of the variances (or standard errors) of survey estimates, since the PSUs from a PPS systematic sample tend to be more diverse than those chosen from a randomly sorted PSU frame.

7. The construction and use of pseudo stratum pairs of PSUs chosen by PPS systematic selection should be done as follows. First, during the development of the sample design, make sure that PSU sample allocation among explicit strata calls for choosing an even number (i.e., 2, 4, 6, 8, 10, etc.) of PSUs in each explicit stratum. Second, as the PPS systematic sample is selected within each explicit stratum, be sure to note the order of PSU selection (i.e., note which PSU was selected 1st, 2nd, 3rd, 4th, 5th, 6th, etc.). Then next, form pseudo strata for each explicit stratum by pairing the 1st and 2nd selected PSUs, the 3rd and 4th selected PSUs, the 5th and 6th selected PSUs, etc. Dealing with an odd number of PSUs in an explicit stratum where sorting criteria correlate with key study outcomes is more complicated and should be avoided if at all possible, since at least two selected PSUs are required for each pseudo stratum, and the only way to achieve that is for one pseudo stratum to have three selected PSUs and all others to have two. The complication with having three selected PSUs in one pseudo strata and two in the rest is that the set of pseudo strata are no longer equal in size, but the variance estimation formulae used in most software packages for analysis of complex samples presume equal-sized pseudo strata in order for their computational formulae to work. The resulting sets of PSU-pairs from all explicit strata define the set of pseudo strata to be used for analysis. Finally, appropriately specify this set of pairs as the PSU strata in the setup for the survey data analysis software that is used to produce variances for GATS the estimates.

8. If a with-replacement (or with minimal replacement) PPS selection technique is chosen to select the sample of first stage units (i.e., segments), then there is some chance that a segment may be selected more than once. In fact, if a sequential PPS with minimal replacement algorithm is used (see for example, Chromy, 1979), a segment is guaranteed to be selected at least once if the size of the segment is sufficiently large. If a segment is selected multiple times (say α_i times) then α_i samples of households should be selected from each selected segment and the 1,..., α_i group number of each selected household should be retained in the final analytic file. The process for handling these situations in latter stages of the design is discussed in *Chapter 8*.

Recommendation:

When the target respondent sample size is 8,000, select ~200 first stage units (segments) from urban areas and ~200 from rural areas.

9. As noted earlier, a requirement for the GATS is that the first stage of selection consists of at least 100 PSUs. It was also noted earlier that countries that wish to select geographic areas in only one stage should select a considerably larger number of first stage units — at least 400. Since a survey where findings are to be presented nationally jointly by gender and urbanicity is designed to achieve 4,000 respondents from rural areas and 4,000 respondents from urban areas, a country should select half the sample of segments (~200) from the urban stratum and half (~200) from the rural stratum. Some deviation from this segment sample allocation to the rural and urban strata is acceptable, particularly if an equal allocation will greatly increase the expected design effect. In any case, however, the sample must be designed to obtain 4,000 individual respondents from the rural areas and 4,000 from the urban areas. In countries aiming to present national findings separately (but not jointly) by gender and urbanicity, only 200 segments would be needed to yield the recommended sample size of 4,000 respondents overall.
10. Selection probabilities should be computed and submitted to CDC as soon as the GATS sample of PSUs has been drawn. See computations of probabilities in *Item 5* of this section above and in *Section 3.3* of the *GATS Sample Weights Manual*.

Recommendation

When the target respondent sample size is 4,000, select 200 segments. Survey should be designed to obtain 4,000 individual respondents from urban areas and 4,000 from rural areas.

7.2 Countries Selecting Geographic Areas in Several Stages of Selection

Countries that plan to select geographic areas in several stages of selection should follow many of the same guidelines as noted above. In summary:

1. PSUs will generally be defined by larger geographic areas. For the GATS, it is recommended that at least 1,000 PSUs be formed within a country and it is required that a random sample of at least 100 PSUs be selected in order to ensure adequate geographic coverage of a country.
2. The stratification variable(s) used at the first stage of collection will depend on the country. In some cases, it may be feasible to clearly designate a PSU as being either rural or urban. If this is the case, then approximately the same number of PSUs within strata defined by rural or urban should be selected in order to optimally obtain the desired 4,000 respondents in rural areas and 4,000 respondents in urban areas. If PSUs cannot easily be classified into rural and urban, then stratification on urbanicity can be imposed at a later stage of geographic selection. In any case, the survey should be designed to achieve 4,000 individual respondents in rural areas and 4,000 respondents in urban areas. The choice of additional explicit stratification variables is limited by the availability of classification information for each geographic segment in the frame. Variables related to tobacco use are ideal options for stratification variables as the variances of sample estimates would be reduced with the degree of this relationship. Stratification variables may also be chosen because their categories form important subpopulations for planned estimation of key smoking variables. These are referred to as analytic “domains” of the survey. A region variable would be appropriate for stratification if regional differences in tobacco use are anticipated (variance lowering effect) even if sample sizes are insufficient for regional estimation of smoking variables.

Recommendation:

Whenever possible, select half of the PSUs from urban areas and half from rural areas.

3. As described in *item 5* of *Section 7.1*, the PSUs should be selected with PPS. The size measure used to select PSUs should be an estimate of the total number of survey-eligible people that reside in the PSU. If these data are not available, some other size measure can be used that is highly correlated with the population count, such as the total number of households in the PSU.

REQUIREMENT
 Select PSUs and segments with probability proportionate to size. Preferred size measure is equal to the count of those 15 years of age or older; otherwise, size is equal to count of HHs.
4. If a sequential PPS algorithm is used to select the first stage sample, pseudo stratum pairs should be formed and used for variance estimation as described in *Item 6* of *Section 7.1* above. Sequential random sampling spreads the across the stratum thereby providing an implicit stratification based upon the ordering of the frame.
5. If a with-replacement or with minimal replacement selection technique is used to select the sample, there is some chance that a PSU may be selected more than once. In fact, if the sequential PPS with minimal replacement algorithm is used, a PSU is guaranteed to be selected at least once if $p_i^{(1)} > 1$. If a PSU is selected multiple times (say α_i times) then the number of geographic subareas selected at the next stage next stage of selection from the PSU should be α_i times the number selected in a PSU that was chosen only once.
6. When PSUs are relatively large geographic areas, some geographic subsampling will need to occur in order to obtain a sample of geographic areas that are of suitable size for a small number of interviewers to work with minimal travel. In general, the ultimate geographic sampling unit (i.e., the segment) should contain approximately 250 households.

Recommendation:
 Ultimate geographic sampling unit, i.e., the segment, should contain ~250 households.
7. The number of stages of selection within a selected PSU may differ, depending on the size of the PSU. For discussion purposes, we will assume at the final stage of geographic selection that an area will be selected that is generally referred to as a **segment**.
8. Segments within a PSU should be selected using a PPS approach that can be similar to the PPS approach used to select the PSUs. The definition of the size measure used to select the segments should be equivalent to the size measure definition used to form PSUs.

REQUIREMENT:
 Select subareas within PSUs (e.g., segments) with probability proportionate to size. Preferred size measure is equal to population count of those 15 years of age or older; otherwise, number of HHs.
9. To see the benefit of using the same size measure definition at both the PSU and segment stages of selection, suppose PSU i has $k = 1, \dots, \Omega_i$ segments.

Furthermore, suppose the sample frame indicates PSU i has $N_i = \sum_k N_{ik}$ people 15 years of age or older. Then with most standard PPS selection approaches, the PSU will be selected with probability

$\frac{I \cdot N_i}{\sum_i N_i}$ where I is the number of PSUs selected. Suppose we wish to select K_i segments within PSU i and segments are selected with probability proportionate to the N_{ik} 's. For most designs we would expect K_i to equal 4. Then the probability of selecting the segment is $\frac{K_i \cdot N_{ik}}{\sum_k N_{ik}} = \frac{K_i \cdot N_{ik}}{N_i}$ and the unconditional probability of selection from both stages reduces to:

$$\frac{I \cdot N_i}{\sum_i N_i} \cdot \frac{K_i \cdot N_{ik}}{\sum_k N_{ik}} = \frac{I \cdot N_i}{\sum_i N_i} \cdot \frac{K_i \cdot N_{ik}}{N_i} = \frac{I \cdot K_i \cdot N_{ik}}{\sum_i N_i}$$

Notice how the N_i term in the numerator of the PSU probability of selection cancels with the N_i term in the denominator of the segment probability of selection. Because of this cancellation, the final unconditional probability of selection for the segment is equivalent to what would have been obtained if the segment was selected in a one-stage design (instead of a two-stage design).

10. The number of subareas to select within a PSU will depend on the size of the PSU. For example, if a with-replacement method is used in which the PSU is randomly selected four times during the selection process, then the number of subareas selected from the PSU should be four times the number of subareas selected from a PSU chosen once. In general, one should always select at least two subareas from within a previously selected area. If the geographic areas are selected in two stages and 100 PSUs are selected at the first stage, then, on average, there should be four segments selected from each PSU.

Recommendation:
Select on average, four segments from each PSU, with a minimum of two.

11. After all geographic stages of selection are complete, the total number of segments selected from urban areas should equal the total number of segments selected from rural areas. In general, approximately 200 segments should be selected from the urban areas and 200 segments from the rural areas.

Recommendation:
After all geographic stages of selection are complete, in a country where national estimates are to be presented jointly by gender and urbanicity, there should be ~200 segments selected from urban areas and ~200 segments selected from rural areas.

12. Selection probabilities should be computed and submitted to the CDC as soon as the GATS sample of PSUs and segment(s) has been drawn. See computations of probabilities in *Item 8* of this section above and in *Section 3.3* of the *GATS Sample Weights Manual*.
13. In general, we expect a country will select approximately 28–30 households in each segment so this would equate to a total segment sample size of roughly 400. Additional information on determining an appropriate segment sample size is provided in *Chapter 10*.

8. Selecting Households

After the geographic areas are selected for the GATS survey, the next stage of the GATS design involves selecting households from the final sample of area segments previously chosen.

There are several ways to construct a suitable household sampling frame and select the sample of households for this stage of the design. Some countries may find one method will work best for all their sampled segments, while other countries may need to use a mixture of methods — perhaps one method for segments located in urban areas and a second method for segments located in rural areas. Some countries may also wish to use an updated master sampling frame that may have been constructed for the last census, a recent labor force survey, or other similar national survey. The methods used by a country will depend on the household data that are available for a segment, the quality of this data, and the ease in which interviewers and supervisors can adjust any errors in this data during their initial visit to the segment. In general, the main goal of this step in the sampling process is to construct the most complete and accurate listing of all households currently located in each sample segment regardless of whether residences are occupied or not.

Achieve high response rates and efficiency by always completing the individual interview as soon as possible after completing the household screening.

REQUIREMENT:

A country's final sample design must include a description of the household frame construction process, and indications of steps that will be taken to assure that the listings of households will be as complete and accurate as possible.

Below are descriptions of methods a country may wish to consider for constructing a new household sample frame for each segment and for selecting the sample of households. A country's final sample design should include a detailed description of how a country plans to form a suitable sample frame for the selection of households, how the country will address imperfections in the frame that are not represented on existing lists (such as newly built households in the area), and how the sample will be selected from each area. This chapter also provides recommendations on how to handle households that are discovered by the interviewer or supervisor and are missing from the household listing (see *Section 8.5*).

8.1 Creating a Household Listing for a Sampled Segment by Visiting the Segment Prior to Selecting the Household Sample

The preferred method of selecting a household sample within a segment is to first create a listing of households in the segment. This household listing becomes the sample frame for the household selection stage of the design and is constructed by visiting the geographic area and enumerating all households that appear within the geographic boundaries of the segment prior to the onset of data collection. This method produces a household sample frame that is current and is particularly beneficial for areas that are experiencing significant growth. If a segment is

Recommendation:

The preferred method of creating a list of households for each segment is to have an enumerator list all households in area prior to data collection.

determined to have a much larger number of households than originally anticipated, it is recommended that the country's GATS sample design team initiate another geographic stage of selection and subselect an area within the segment. Again, the ultimate goal is to have segments that are not too small or too large—they should contain approximately 250 households.

During the enumeration visit where a field worker will create the list of households in the sampled segment, the households in the list should be ordered in a *continuous path of travel order* as indicated in the *GATS Mapping and Listing Manual*. Presenting the households in the continuous path of travel order is very important since the order can be used by interviewers and supervisors to identify any missed households on the household listing. If an area is located in an extremely remote or inaccessible area of the country, or some logical continuous path of travel does not exist for the segment, then a country may wish to have the enumerator record the location of households on a map of the segment.

Recommendation:

Whenever a new household listing is created within each sample segment for GATS, it is preferable to have households listed in a continuous path of travel order.

It is worth noting that in order to adequately deal with households that were not listed in a segment that may occur regardless of the method used to create the household listing, the household listing must either be:

- a list of households that are ordered in a logical, continuous path of travel through the segment, or
- the households within the segment must be noted on a map.

REQUIREMENT:

Households must be either listed on a sheet in a continuous path of travel order or be noted on a map of the segment.

Suggested methods for handling missed households on the household listing are presented in *Section 8.5*.

8.2 Creating a Household Listing for a Sampled Segment Using Administrative Records

A second method of creating a household listing for a sample segment is to use administrative records, such as a previous census. This method of creating a household listing can be less costly than constructing a listing via a personal visit to the area by an enumerator. However, the quality of the list can be much lower, particularly if the administrative records used to create the list are old. A country should contact their WHO/CDC representative if they plan to use an individual-level directory for creating their household listing. Some things to consider when creating a sample frame using administrative records include:

1. As noted previously, one should consider the age of the administrative data. Older data may yield household listings with numerous inaccuracies, particularly for areas that are prone to change, such as growing urban areas.
2. Consider the ease with which one could either make a list of households that represent a logical path of travel through the area and/or the ease with which one could note the households on a map of the segment. Having either a logical path of travel or a map of the households will be important for the field staff, particularly since they are likely to find many missed housing units in the area.

3. If multiple administrative sources are used to make a household listing, one should consider the ease with which one could subtract duplicates from the sources. This may need to be done by visual inspection. For example, one source may list a house as “110 Main Street, Apt D” and a second source may list a house as “110 Main St, #D.”

8.3 Selecting the Sample of Households

In the previous two sections, it was noted that the household listing can be created from a separate field listing of households or from existing administrative records, such as from the most recent census. Either way, a listing of households (or household addresses) is the end-result. A within-segment simple random sample or systematic sample of households is then separately chosen within each sample segment. The recommended number of selected households in each segment by either method is around 28. Additional details on determining the sample size are provided in *Chapter 10*.

REQUIREMENT:

Sampling from a list of households should be done using a simple random or systematic, without-replacement method.

Simple random sampling calls for random selection without replacement from the list, while systematic sampling involves sampling every K -th household on the list after a random starting point, where the sampling interval (K) for each segment is based on the ratio of the total number of households in the segment and the designated number of selected households for the segment (e.g., 28). Most sampling textbooks describe sample selection by these two methods. Achieving the designated number of selected households is more difficult with systematic sampling when the total number of households in the segment is not an integral multiple of the segment's designated household sample size. Kish (1965, Section 7.5) describes a way to deal with this problem.

The conditional selection probability for any household selected in each segment depends on which of the two selection methods is used. *If simple random sampling is used*, this probability is simply the household sampling rate (i.e., the ratio of household sample size to total number households on the list frame). *When systematic sampling is used* to select households, this selection probability one divided by the selection interval (K). See the *GATS Sample Weights Manual* for more details on computing this and other GATS selection probabilities.

Once the household sample has been drawn, probabilities of selection should be submitted to the CDC for each selected PSU either in a single step or multiple steps.

8.4 Optionally Assigning Households to the Male and Female Household Groups

As an optional design consideration, some countries may wish to randomly partition sample households into two groups. One set would be assigned to the male household group and the remaining sample would be assigned to the female household group. Those households assigned to the male group will have only eligible males, 15 years of age or older, rostered during the screening portion of the interview and therefore only a male will be selected from within the household for the GATS individual interview. Similarly, those households assigned to the female group will have only eligible females, 15 years of age or older, rostered during the

Design Option:

Randomly assign selected households to either the male household group or the female household group.

screening portion of the interview and therefore only a female will be selected from within the household for the GATS individual interview. This partitioning of the sample into male and female household groups may be done for any of the following reasons:

1. In some countries, respondent recruitment will be less complicated and thus participation rates will be higher if the gender of the field interviewer matches that of the respondent. This gender randomization of the household sample within each segment to the male and female groups enables a country to achieve this gender match.
2. In some countries where female tobacco use prevalence rates are dramatically lower than corresponding male rates, it may be necessary to disproportionately sample females at a higher rate than males. This kind of over-sampling by gender can be accomplished most readily by gender randomization of households.
3. Based on the previous experience and studies, males tend to respond at lower rates than females; this partitioning of the sample will allow a country to efficiently account for this response propensity differential by randomly assigning a slightly larger number of households to the male group compared to the female group.

We caution countries to only choose gender randomization where there is a clear advantage in data collection. This approach may increase household nonresponse as households with only male residents encountered may be designated for female selection and vice versa.

We recommend that the random assignment of sampled households to male and female groups be done using a simple random sampling process. For instance, as noted in number 3 above, in most countries one would expect the males to respond at lower rates than females, consequently a greater number of sampled households should be assigned to the male group compared to the female group.

Suppose $\bar{H}_{ik}$ households are selected within a segment k

located in PSU i and we wish to assign $\bar{H}_{male,ik}$ of these to the male group and $\bar{H}_{female,ik}$ of these to the female group. So $\bar{H}_{male,ik} + \bar{H}_{female,ik} = \bar{H}_{ik}$. Then a simple random sample (without-replacement) of $\bar{H}_{male,ik}$ should be selected from the $\bar{H}_{ik}$ households and assigned to the male group. The remaining $\bar{H}_{ik} - \bar{H}_{male,ik} = \bar{H}_{female,ik}$ should then be assigned to the female group. For example, suppose 28 households are selected within a segment and 15 of these should be assigned to the male group. Then 15 households should be randomly selected from the 28 and these should be given a probability of selection into the male group of 15/28. The remaining 13 households would be assigned to the female group and given a probability of selection equal to 13/28.

Recommendation:

Random partitioning of sampled households to male/female group should be done using a simple random sampling method.

REQUIREMENT:

The probability of assigning a household to the male or female group must be retained on the analysis file used to create sample weights.

As with all other stages of selection, the probability of assigning a household to the male or female group should be retained in the final GATS analytic file.

8.5 Preparing an Electronic Sample File for Handheld Computer Administration

When handheld computers are used for data collection, an electronic version of the sample must be prepared. This file should contain the information needed to load cases into the handheld computer. Sampling or their IT support staff should be prepared to create a file derived from the master sample that will contain the following information for each dwelling unit to be surveyed:

- ID information (case ID, PSU, region).
- Locator Information.
- Designation of gender randomization grouping as Male or Female (if design option is incorporated).
- ID of the Field Interviewer (FI) to receive the cases (optional).

See the *GATS Programmer's Guide to General Survey System* for more information about the layout and exact content needed for this file.

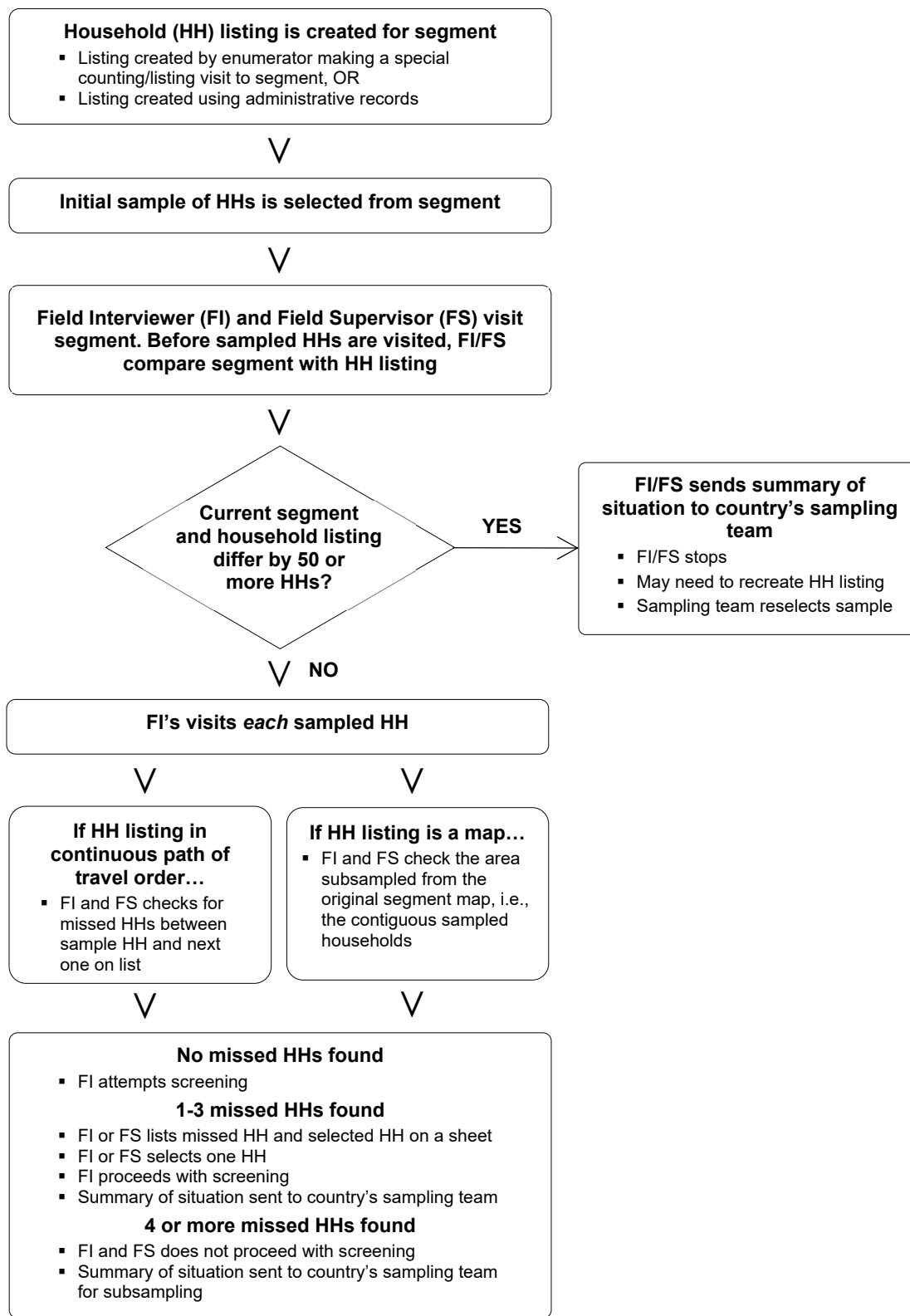
8.6 Addressing Missed Households and Other Apparent Errors in the Household Listing

Any of the following may cause existing household addresses to be missing from the list frame used to choose the within-segment household sample in the GATS: (i) the frame may be several years old so that newly constructed residential housing is missing; (ii) administrative address lists may exclude entire segments of the residential population (e.g., non-voters), and (iii) a newly created list may be of unknown quality if the household listing was done by staff who lacked prior experience or training in household enumeration. In any event, the use of list frames with a seriously high rate of incomplete coverage can lead to biased estimates from GATS samples, particularly when the smoking behavior of individuals living in households that are included on these frames collectively differs from the behavior of those living in households that were missed on the list frame used for household selection. Practical preventive strategies are available to minimize the impact of frame coverage bias. This section presents a summary of the process that should be used to address missed households or other errors found in the household listing at the time the field interviewer and field supervisor visits the segment just prior to contacting each sampled household. A summary of this process is presented in *Exhibit 8-1*.

Recommendation:

Each GATS country should devise a strategy to deal with the likelihood that the sampling frames used for household sampling are at least partially incomplete.

Exhibit 8-1. Summary of Process to Address Missed Households on Household Listing



After the initial sample of households is selected, an interviewer or supervisor should visit the area prior to making any contact with a household. The interviewer/supervisor should carry with them a copy of the household listing. During this initial visit, the interviewer/supervisor should do the following:

1. Identify any large groupings of households that are located within the boundaries of the segment and that are not included on the household listing from which the GATS sample was drawn. If this occurs, the interviewer or supervisor should record the missed housing units and send the information back to the sample design team prior to visiting any household. Although it will depend on the area being sampled, in general if an interviewer or supervisor finds 50 or more housing units, then they should send an updated listing of the area back to the country's sample design team prior to conducting any work. If an interviewer or supervisor finds less than 50 housing units, sending an updated listing may not be necessary. The half-open interval technique (described below) may be sufficient to account for fewer than 50 missed households in a segment.

Recommendation:
If interviewer/supervisor finds 50 or more missed HHs in a segment, then the household listing should be recreated and the sample redrawn.

2. For each sampled household, the interviewer and supervisor should employ the commonly used, **half-open interval** technique for identifying households that might have been missed on the household listing for the area. This methodology can only be used if the household listing reflects a continuous path of travel as noted in *Section 8.1*.

Recommendation:
The half open interval technique should be used to deal with missed HHs when HH listing reflects a continuous path of travel.

In summary, for each sampled household, the interviewer or supervisor should look to see if any household exists between the sampled household and the household right underneath it on the household listing. If one to three households appear to be missing, the interviewer or supervisor should record the sampled unit and the various missed household(s) on a separate sheet. The interviewer/supervisor should then randomly select one household from within this set. The originally selected household may be selected again, or one of the missed households may be selected. If more than three households appear to be missing, this information should be sent back to the country's sample design team prior to making any contact with a household in the area. The sample design team will decide if taking a random sample of one unit is sufficient to address the missed households or if taking a greater sample is warranted.

Note that if one to three households are missing from the household listing, then the sheet that the interviewer/supervisor used to select a household should be sent to the sample design team so they can record this information. Among other things, the link between the questionnaire and the newly selected household may need to be adjusted in the final analytic file in cases where the originally selected household was not selected again.

Recommendation:
When one to three missed HHs are found, interviewers or supervisors will list missed HHs on sheet and select one. This sheet must be returned to country's design team.

3. If a household listing does not reflect a continuous path of travel and the sample of households was selected as a contiguous cluster, then the interviewer and supervisor should identify households that might have been missed within the geographic boundaries of the contiguous cluster chosen. If a contiguous group of one to three households appears to be missing within these boundaries, the interviewer or supervisor should identify a selected household nearest to the group of one to three. The nearest sampled household, as well as the missed households, should be listed on a separate sheet and one household should be randomly selected from this sheet. This newly selected household will take the place of the nearest sampled household. Again, this sheet should be sent back to the sample design team so that the additional units as well as the replaced household can be properly documented in the GATS analytic files. If an interviewer or supervisor finds a contiguous group of more than three missed households then they should record all the missed households on a sheet and send this information back to the sample design team before any contact is made with a sampled household. In this case, the sample design team may want to consider sub selecting an area within the contiguous group of households.

Recommendation:
Interviewer/supervisor should look for missed HHs in the contiguous cluster selected within the segment when household listing is a map.
4. The interviewer and supervisor should inspect the selected dwelling unit to see if there appears to be more than one household living in the unit. For example, a house may have been converted to several apartments. When this occurs, the interviewer or supervisor should create a listing of all the households in the selected dwelling unit on a separate sheet and randomly select one of the households for participation into the GATS survey. This sub-sampling step adds another stage of selection to the design, and it requires that the number of households at the dwelling unit address be recorded, since the conditional sub-sampling probability is one divided by the number of households at the selected address. Again, if a single dwelling unit has more than three missed households, then the information should be sent back to the sample design team in the country prior to any contact with a household.

REQUIREMENT:
Interviewer and Supervisor should identify any selected dwelling unit with more than one household and choose one of them at random.

More detailed recommendations on how the interviewers and supervisor should address these situations are presented in the *GATS Field Interviewer Manual* and the *GATS Field Supervisor Manual*.

From a sample design perspective, these types of issues are common in household surveys and should be addressed on an individual basis. If an interviewer or supervisor finds a small number of missed housing units or finds a selected dwelling unit that contains several households, it will be most efficient if the interviewer or supervisor takes a random sample of the missed households while visiting the segment. If a large number of households are missed, then interviewers and supervisors are instructed to convey this information to the country's sampling team and await further instructions. In the latter case, one can address the situation in one of three ways:

1. The statisticians may simply want to add the missed housing units to the household listing and reselect the sample.

2. The statisticians may simply want to add the missed housing units to the end of the household listing and select a sample of households from those that were missed, using the same sampling rate that was used to select the initial sample of households from the segment. This method would be particularly advantageous if an interviewer made contact with some of the sampled households prior to sending the information back to the country's sampling team.
3. If a very large number of missed households are identified, the statistician may wish to partition the segment into smaller geographic areas, randomly select an area, and then select a new sample of households that reside in this subarea. This process is sometimes referred to as sub-segmentation.

In all cases, it is imperative that any sub-selection be done using some sort of random process, and that the probabilities of selection during this sub-selection process be calculated and retained on the final analytic file for the survey.

9. Selecting an Eligible Resident Within Each Screened Household

A random selection method should be used to select an eligible individual at random from within sampled households. (See GATS eligibility criteria for individuals in *Section 3.1*.) In summary, this method of selection proceeds as follows:

1. For each household selected into the GATS sample, an interviewer will knock on the door and attempt to identify an individual in the household who is 18 years old or older and knowledgeable about household residents. This individual will be considered the **screening respondent**.
2. An interviewer will ask the screening respondent several questions about the residents of the household. Among other things, their goal is to create a roster of all non-institutionalized household residents, 15 years of age or older, who consider the sampled household to be their usual residence, and whose gender matches the male/female group assignment of the sampled household (if gender randomization was applied to the country's household sample). The eventual list of eligible household residents on the roster should be ordered from oldest to youngest.
3. Once the roster of all eligible residents of the household is entered into the handheld device, the computer will use a random process to select one name on the roster and that name will be displayed on the screen. More specifically, the handheld device will generate a random number after the household roster is completed. Suppose, for example, that the random number generated for household j is r_j and

Every household with one or more eligible residents will have one of them randomly selected from a household roster.

suppose the total number of people rostered in the household is κ_j . The handheld device will be programmed to select the individual corresponding to the integer portion of $(r_j \cdot \kappa_j) + 1$. For example, if the randomly generated number is 0.365789 and there are three eligible individuals on the roster, the product of 0.365789 and 3 is 1.097367, and adding one to this yields 2.097367. Thus, the individual in the second position on the roster is selected. Note that by using this method of selection, all eligible respondents have an equal chance of selection and the probability of selection for the randomly chosen resident within household is the inverse of the number of eligible residents rostered in the household. To avoid selection bias by this process, the handheld device is programmed to allow only running the respondent selection once per household (to avoid the problem of interviewers hitting the "select" command until a preferred resident is chosen). The resident selected by the handheld and the resident completing the *Individual Questionnaire* should be the same person.

10. Determining Sample Sizes at Each Stage of Selection and Reporting Final Response Rates

In this chapter, we provide guidance on determining the selected sample sizes needed at each stage of selection for GATS. We also provide guidance on computing the final eligibility and response rates for GATS. The definition of the response and eligibility rate to use is particularly important since having all countries use consistent definitions of these rates will allow GATS to compare these critical data quality measures between countries.

10.1 Sample Sizes at Each Stage of Selection

Recall that PSUs and/or segments should be stratified by urbanicity. Some countries may wish to impose additional explicit stratification in order to obtain a specified sample size in other domains of interest — for example by region. In this section we have provided an example assuming a national sample size of 8,000 respondents and where gender randomization is a feature of the design. Countries with other respondent sample sizes and features should modify their calculations accordingly.

Within each of the geographic strata, suppose the index s refers to a geographic stratum. Furthermore, suppose:

R_s^{Males} = Total number of male respondents desired from within the stratum. If a country is only forming strata by urbanicity, then $R_s^{Males} = 2,000$ for s =urban and s =rural. It is required that $\sum_s R_s^{Males} = 4,000$.

$R_s^{Females}$ = Total number of female respondents desired from within the stratum. If a country is only forming strata by urbanicity, then $R_s^{Females} = 2,000$. As with the males, it is required that $\sum_s R_s^{Females} = 4,000$.

$\mathcal{E}_s^{HH}$ = Estimated household eligibility rate. This is defined as the total number of households that will be selected and determined to be eligible for the GATS divided by the total number of households selected. A household may be declared ineligible if it is unoccupied, has no eligible residents (of either gender if gender randomization is not used, or of the assigned gender if gender randomization is used), or otherwise does not exist at the time of data collection. In general, we would expect this rate to be approximately 90%, although this will depend on the country, the likelihood of households with all residents of the same gender (when gender randomization is used), and the age of the data used to form the household listing for the areas. The GATS listing protocol calls for the identification of all household structures regardless of whether residences are occupied or not. This may impact on the eligibility rate encountered in the GATS survey compared with other national surveys conducted in countries. This should be

taken into consideration in determining the estimate of eligibility used to compute the total number of households to select within stratum.

τ_s = Estimated household screening rate. This is the percent of households that were successfully screened and have at least one survey-eligible individual. This rate accounts for households with no eligible residents 15 years of age or older. This also accounts for households that were selected for the male group but have no males living in the household, when gender randomization is used. Similarly, this accounts for households that were selected for the female group but have no females living in the household. In general, we would expect the fraction of households that have at least one eligible individual for the GATS to be very high. Most countries should assume a rate of 95% or higher.

ϵ_s^{Person} = Estimated individual-level eligibility rate. This is defined as the total number of people that will be selected and determined to be eligible for the GATS divided by the total number of people selected. For instance, an individual would be declared ineligible if it was determined that they were 14 years of age or younger during the interview. In general, we would expect this rate to be very high — approximately 98% or higher.

Note: The ϵ_s^{Person} accounts for people who were selected and later determined to be ineligible for the survey. The τ_s accounts for households where the roster was not completed because no survey-eligible people were identified by the household resident responding to the screening questions.

ρ_s^{HH} = Estimated household response rate. This is defined as the total number of households with an individual that will respond to the roster questions divided by the total number of *eligible* households selected. In general, we would expect this rate to be approximately 98% for most countries.

ρ_s^{Males} = Estimated individual-level response rate for males. This is defined as the total number of males that will respond to the GATS divided by the total number of *eligible* males selected. In general, we would expect this rate to be approximately 80% for most countries.

$\rho_s^{Females}$ = Estimated individual-level response rate for females. This is defined as the total number of females that will respond to the GATS divided by the total number of *eligible* females selected. In general, we would expect this rate to be approximately 90% for most countries.

Then

M_s^{Males} = $\frac{R_s^{Males}}{\epsilon_s^{Person} \cdot \rho_s^{Males}}$ is the total number of selected males needed in stratum s from screened households.

$M_s^{Females} = \frac{R_s^{Females}}{\epsilon_s^{Person} \cdot \rho_s^{Females}}$ is the total number of selected females needed in stratum s from screened households.

And

$H_s = \frac{M_s^{Males} + M_s^{Females}}{\epsilon_s^{HH} \cdot \tau_s \cdot \rho_s^{HH}}$ is the total number of households that should be selected from stratum s .

We further recommend that the within-cluster sample size be restricted to about 28 selected households per ultimate geographic area (usually segment) selected. In general, the optimal cluster sample size depends on the outcome measure being considered and the extent of homogeneity within the cluster.

Recommendation:

Select 28 households per segment.

Therefore, if the desired number of households to select within each ultimate geographic areas is 28, then the total number of areas to select at the final stage of the geographic selection process is approximately equal to $\frac{H_s}{28}$.

Note that if a country is selecting geographic areas with one stage of selection, then the total number of PSUs to select is $\frac{H_s}{28}$, which should roughly equal 400.

As noted earlier, if a country is selecting geographic areas in multiple geographic stages of selection, then the total number of PSUs to select might be about 100 and a country should ultimately select a sample of approximately four segments per PSU, or approximately 400 segments.

10.2 Example Sample Size Computation

To illustrate the sample size computations mentioned in the previous section, suppose within some stratum that a country assumes:

Total Number of Male Respondents in the Stratum (Note: This sample size should equal 4,000 or more over all strata)	R_s^{Male}	2,000
Total Number of Female Respondents in the Stratum (Note: This sample size should equal 4,000 or more over all strata)	R_s^{Female}	2,000
Total Number of Respondents in the Stratum (Note: This sample size should equal 8,000 or more over all strata)	R_s^{Person}	4,000
Household Eligibility Rate	ϵ_s^{HH}	90%
Household Response Rate	ρ_s^{HH}	98%
Percent of Households with at Least One Survey-Eligible Individual	τ_s	95%
Individual Eligibility Rate	ϵ_s^{Person}	98%
Individual Response Rate	ρ_s^{Person}	85%
Individual Response Rate for Females	ρ_s^{Female}	90%
Individual Response Rate for Males	ρ_s^{Male}	80%

Then the total number of selected people within households that are successfully screened are:

$$M_s^{Females} = \frac{R_s^{Females}}{\epsilon_s^{Person} \cdot \rho_s^{Females}} = \frac{2,000}{.98 \cdot .90} = 2,268 \text{ selected females and}$$

$$M_s^{Males} = \frac{R_s^{Males}}{\epsilon_s^{Person} \cdot \rho_s^{Males}} = \frac{2,000}{.98 \cdot .80} = 2,551 \text{ selected males.}$$

And the total number of households to select from within this stratum is

$$H_s = \frac{M_s^{Males} + M_s^{Females}}{\epsilon_s^{HH} \cdot \tau_s \cdot \rho_s^{HH}} = \frac{2,551 + 2,268}{.90 \cdot .95 \cdot .98} = 5,751 \text{ selected households.}$$

Please note in the above example that gender randomization was done. As a standard GATS survey where there is no gender randomization is done, the total number of persons within households that are successfully screened are:

$$M_s^{Person} = \frac{R_s^{Person}}{\epsilon_s^{Person} \cdot \rho_s^{Person}} = \frac{4,000}{.98 \cdot .85} = 4,954 \text{ selected persons.}$$

And the total number of households to select from within this stratum is

$$H_s = \frac{M_s^{Person}}{\varepsilon_s^{HH} \cdot \tau_s \cdot \rho_s^{HH}} = \frac{4,954}{.90 \cdot .95 \cdot .98} = 5,912 \text{ selected households.}$$

10.3 Response Rates

Response rates and eligibility rates should be computed for the household sample, individual sample and the combined sample using the formulas noted below. It is important that these formulas be used so that response rates and eligibility rates can be compared between countries. These rates should be computed for the entire sample (both household and individual), as well as by strata. The strata include rural/urban, male/female and any other strata that a country is using in their design (e.g., perhaps region).

REQUIREMENT:

Final eligibility and response rates should be computed using standard GATS formulas specified in this section.

It should be noted that the definition of response rates as defined below is consistent with a standard definition of response rates as suggested by the American Association of Public Opinion Research (AAPOR). For GATS, we are utilizing the definition of what AAPOR refers to as response rate RR1 in the 2009 version of the AAPOR report which can be found at http://www.aapor.org/Standard_Definitions/1818.htm. When a sizable portion of nonresponding sample cases are of unknown study eligibility and the percentage of eligibles among those of known eligibility is relatively low, it is recommended that response rate RR3 be used instead of RR1. Computation of these rates is more fully described in the *GATS Sample Weights Manual*.

Two sets of operational disposition codes will play a role in computing response rates at the household level (for efforts to get someone in the household to complete a *Household Questionnaire*) and at the individual level (for efforts to get a randomly selected resident in the household to complete an *Individual Questionnaire*). Multiple attempts by interviewers may be required to successfully complete these two levels of response. One of a set of household- and individual-level result codes, respectively, will be entered into the program after each of these attempts (see *GATS Field Interviewer Manual* for a complete listing of these two sets of result codes). Once data collection has been completed, a household-level and individual-level final disposition code will be determined for each sampled household (see *GATS Quality Assurance: Guidelines and Documentation* for further information about assigning disposition codes). In most instances, at either of the two levels of effort, the final disposition code will match the result code recorded on the final attempt to complete that part of survey recruitment.

The final disposition codes at the household- and individual-levels of effort are defined in the following two exhibits:

Exhibit 10-1. Household Questionnaire Final Disposition Codes

Code	Description
1	Completed Household Questionnaire, One Individual Selected —Household questionnaire is considered complete if the roster is complete
2	Completed Household Questionnaire, No One Selected —No one aged 15+ —No one considers the household to be their usual place of residence —Household was assigned to males and no males reside in household or household was assigned to females and no females reside in household (if gender randomization was used)
3	Completed Part of Household Questionnaire, Could Not Finish Roster (Incomplete Interview)
4	Household Questionnaire Not Complete, Could Not Identify an Appropriate Screening Respondent —No one 18+ at home —Available 18+ household member incompetent —These households may or may not have survey-eligible residents
5	Nobody Home
6	Household Refusal
7	Unoccupied House
8	Selected Address is Not a Household
9	Other Household Nonresponse

Exhibit 10-2. Individual Questionnaire Final Disposition Codes

Code	Description
11	Completed Individual Questionnaire —Questionnaire completed at least through question E1
12	Incomplete —Questionnaire started, but broken off before question E1
13	Selected Individual was Later Determined to be Survey Ineligible —Age < 15 —Individual does not consider household their usual place of residence —Incorrect gender (if gender randomization was used)
14	Selected Respondent Not Home
15	Selected Respondent Refusal
16	Selected Respondent Incompetent
17	Other Individual Nonresponse

Suppose [*] is the number of cases in GATS with final disposition code *. Note that each questionnaire can have up to two final disposition codes, one that reflects the final status at the household (screening) level and a second code that reflects the final status at the individual (interview) level. Specifically,

- If a questionnaire has a final household disposition code of 1, then the questionnaire must also have a second final disposition code associated with it that reflects the final status of attempts to interview the selected individual resident of the household using the *Individual Questionnaire*.
- If a questionnaire has a final household disposition code of 2, 3, 4, 5, 6, 7, 8, or 9 then there will be no final disposition code assigned to it that reflects the status of the selected individual resident.

Then using the notation noted in *Section 10.1*,

$$\begin{aligned}\mathcal{E}_s^{HH} &= \text{Estimated household eligibility rate.} \\ &= \frac{[1] + [3] + [4] + [5] + [6] + [9]}{[1] + [2] + [3] + [4] + [5] + [6] + [7] + [8] + [9]}\end{aligned}$$

$$\begin{aligned}\rho_s^{HH} &= \text{Estimated household response rate.} \\ &= \frac{[1]}{[1] + [3] + [4] + [5] + [6] + [9]}\end{aligned}$$

Notice that [3] is not included in the numerator of the household response rate. Therefore, a household screening questionnaire that is only partially complete (i.e., the roster could not be finished) will be considered a nonrespondent to the GATS. Note also that [2] is not to be included in either numerator or denominator of the response rate since households with no eligible residents are considered ineligible for the survey. These would be households with no eligible residents, 15 years of age or older who at the time of household screening consider the household to be their usual residence. If gender randomization is done, code [2] would also include households in a particular gender group that have no eligible residents in that gender group.

$$\begin{aligned}\tau_s &= \text{Estimated percent of households with at least one survey-eligible individual.} \\ &= \frac{[1]}{[1] + [2]}\end{aligned}$$

$$\begin{aligned}\mathcal{E}_s^{Person} &= \text{Estimated individual-level eligibility rate for males and females combined (if} \\ &\text{gender randomization is not used); calculated separately for each gender group of responding} \\ &\text{households if gender randomization is used.} \\ &= \frac{[1\ 1] + [1\ 2] + [1\ 4] + [1\ 5] + [1\ 6] + [1\ 7]}{[1\ 1] + [1\ 2] + [1\ 3] + [1\ 4] + [1\ 5] + [1\ 6] + [1\ 7]}\end{aligned}$$

$\rho_s^{Person} =$ Estimated individual-level response rate for males and females combined.

$$= \frac{[1\ 1]}{[1\ 1] + [1\ 2] + [1\ 4] + [1\ 5] + [1\ 6] + [1\ 7]}$$

Based on the description above and **RR1** AAPOR response rate, “Incomplete” interviews (code 12) are considered eligible nonrespondents to the GATS.

And the combined response rate for the GATS, for males and females combined, is:

$$\rho_s^{Total} = \rho_s^{HH} \cdot \rho_s^{Person}$$

It should be noted that the final response rates for a country’s GATS may be slightly different than those reported by a country during data collection, because final disposition codes will be assigned to the final result codes that were entered during data collection.

Combined response rate for the GATS is defined as the product of the household and individual response rates.

11. Computing Final Sample Weights for Analysis

Sample weights are numerical measurements that are essential to producing and evaluating estimates from sample survey data. They are intended 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 GATS are typically computed in the following three basic steps.

Sample weights are important statistical measurements that must be computed and used for analysis of GATS data.

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 stages that led to selecting the respondent (e.g., for a four-stage sample of households with no gender randomization of household, the probability of first selecting the PSU in which the respondent resides, times the probability of choosing the respondent's SSU given that its PSU was chosen, times the probability of choosing the respondent's segment given that its SSU was chosen, times the probability of choosing the respondent's household given that its segment was chosen, and times the probability of selecting the respondent within the respondent's household).
2. **Nonresponse Adjustment** — The base weight is multiplied times the inverse of the household- and individual-level response rates for a subset of the selected eligible sample members that are similar to the respondent with respect to characteristics that (hopefully) correlate with key study outcome measurements and the propensity to respond in GATS.
3. **Calibration** — The adjusted base weight is then multiplied times 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, gender, and level of education). These calibrated weights become the final adjusted sample weights that should be used for all analyses of the GATS 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. Computational details for each step in this process are given in *GATS Sample Weights Manual*.

Sample weights are computed in a three-step process.

As noted in *Chapter 4*, the probability of selection for each stage of the sample design must be retained on the final analytic file for each selected household and individual. These factors should be computed, stored, and the process followed in computing them carefully documented at the time that the sample in each stage of selection is chosen.

12. Bibliography

12.1 Sampling: Simple and Brief

Henry, Gary T., *Practical Sampling*, Sage Publications, Newbury Park, 1990.

Kalton, G., *Introduction to Survey Sampling*, Sage Publications, 1983.

Stuart, A., *Basic Ideas of Scientific Sampling*, 2nd Ed., Griffin's Statistical Monographs and Courses, No. 4, Hafner Press, New York, 1976.

12.2 Sampling: Introductory

Barnett, Vic., *Sample Survey Principles and Methods*, Oxford University Press, 2002.

Chaudhuri, Arijit, *Survey Sampling: Theory and Methods*, M. Dekker, New York, 2005.

Cochran, W.G., *Sampling Techniques*, 3rd Ed., John Wiley and Sons, New York, 1977.

Foreman, E. K., *Survey Sampling Principles*, M. Dekker, New York, 1991.

Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., and Tourangeau, R., *Survey Methodology*, Wiley and Sons, New York, 2004.

Jaeger, Richard M., *Sampling in Education and the Social Sciences*, Longman, New York, 1984.

Jessen, Raymond J., *Statistical Survey Techniques*, Wiley, New York, 1978.

Kish, L., *Survey Sampling*, Wiley, New York, 1965.

Konijn, H. S., *Statistical Theory of Sample Survey Design and Analysis*, American Elsevier Pub. Co., New York, 1973.

Levy, P. S. and Lemeshow, S., *Sampling of Populations – Methods and Applications*, Wiley, New York, 1999.

Lohr, S., *Sampling: Design and Analysis*, Duxbury Press, Pacific Grove, 1999.

Raj, D., *Sampling Theory*, McGraw-Hill, Inc., New York, 1968.

Scheaffer, R.L., Mendenhall, W., and Ott, L. *Elementary Survey Sampling* (6th Edition), Duxbury Press, 2006.

Thompson, Steven K., *Sampling*, Wiley, New York, 2002.

Yates, Frank, *Sampling Methods for Censuses and Surveys*, Macmillan, New York, 1981.

12.3 Sampling: More Advanced Treatment

Cassel, Claes-Magnus, Sarndal, Carl-Erik, and Wretman Jan H'akan., *Foundations of Inference in Survey Sampling*, Wiley, New York, 1977.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G., *Sample Survey Methods and Theory*, Vol. I (Concepts and Discussion) and Vol. II, 1953.

Krewski, D., *Current Topics in Survey Sampling*, Academic Press, New York, 1981.

Murthy, M.N., *Sampling Theory and Methods*, Statistical Publishing Society, Calcutta, 1967.

Namboodiri, N. K., *Survey Sampling and Measurement*, Academic Press, New York, 1978.

Sarndal, Carl-Erik, Swensson, Bengt, and Wretman, Jan H'akan, *Model Assisted Survey Sampling*, Springer-Verlag, New York, 1992.

Sukhatme, P. V. et al., *Sampling Theory of Surveys with Applications*, Iowa State University Press, 1984.

Sukhatme, P.V. and Sukhatme, B.V., *Sampling Theory of Surveys with Applications*. Revised 2nd Ed., Iowa State University Press, Ames, 1970.

Wolter, K.M., *Introduction to Variance Estimation*, Springer-Verlag, New York, 1985.

12.4 Practical Manuals

- Grosh, M. and Munoz, J. A Manual for Planning and Implementing the Living Standard Measurement Survey Study. LSMS Working Paper No. 126. World Bank, Washington, DC. 1996.
- Macro International. Sampling Manual. DHS Basic Documentation No. 6. Calverton, MD., 1996.
- SAS Institute. SAS/STAT User's Manual, Version 8. Cary, NC, 1999.
- United Nations. Household Sample Surveys in Developing and Transition Countries. New York, NY, 2005.
- United Nations. Principles and Recommendations for Population and Housing Censuses, Revision 1. Series M, No. 67, Rev. 1 (United Nations publication, Sales No. E.98.XVII.1). (para. 2.61).

12.5 Other Papers

- Chromy, J.R. (1979). "Sequential Sample Selection Methods." Proceedings of the American Statistical Association, Section on Survey Research Methods, pp. 401-406.
- Folsom, R.E., F.J. Potter, and S.R. Williams (1987). "Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains." Proceedings of the American Statistical Association, Section of Survey Research Methods, pp. 792-796.
- Kavee, J., Liu, J., Iannacchione, V.G. (2001). "An Optimal Allocation Method for Two-Stage Sampling Designs with Stratification at the Second Stage". Proceedings of the American Statistical Association 2001, Section on Survey Research Methods.
- Liu, J., Iannacchione, V.G., and Kavee, J.D. (1998). "Estimating Variance Components for a Two-Stage Design with Second-Stage Strata Nested Within PSUs". Proceedings of the American Statistical Association 1998, Section on Survey Research Methods, pp.657-661.
- Sampford, M.R. (1967) On Sampling Without Replacement With Unequal Probabilities of Selection. Biometrika, 54, pp 499-513.
- Vijayan, K. (1968), "An Exact π_{ps} Sampling Scheme: Generalization of a Method of Hanurav," Journal of the Royal Statistical Society, Series B, 30, 556 - 566.

12.6 GATS Manuals

- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Core Questionnaire with Optional Questions*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Field Interviewer Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Field Supervisor Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Mapping and Listing Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Quality Assurance: Guidelines and Documentation*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Sample Weights Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2020.

Appendix A: Rationale for the GATS Recommended Sample Size

Chapter 4 of this manual indicates that the minimum acceptable household sample size within each country, for each round of GATS, is 8,000¹. This number was determined by members of a GATS planning committee during an early phase in the development of the GATS research protocol. It was based solely on power and precision requirements for key GATS estimates by level of urbanicity (i.e., urban or rural) and gender. This appendix provides the rationale behind the recommended sample size.

Important Estimates in GATS

Statistical measures of power and precision were considered and thus developed in the context of the types of estimates that are required from GATS. Two types of estimates were thought to be important in the analysis of GATS data. One type indicates change in various tobacco use measures over time (i.e., between various pairs of survey rounds). These change measures will be important since successive rounds of GATS will be conducted as tobacco use control programs are being implemented, and thus must be evaluated for their effectiveness. A second type of estimate will be used to profile tobacco use in the population at the time that individual rounds of the GATS are being conducted.

Power and Sample Size to Detect Change Over Time

Estimates of change from one round to another as well as round-specific estimates will be produced for various behavioral/use rates. The following three rates were thought to be most important for planning purposes:

- **Tobacco Use Prevalence Rate** — The percentage of individuals currently using tobacco products;
- **Secondhand Smoke Rate** — The percentage of individuals who are exposed to smoke that is produced by other smokers; and
- **Quit Rate** — The percentage of current smokers who have tried to quit smoking.

Producing change or round-specific estimates of tobacco use by gender for urban and rural segments of the target population was thought to be sufficiently important that sample sizes should be roughly the same for each of these four subgroup estimates in each round of GATS. While sample sizes for male and female respondents were expected to be similar, it was decided that the composition of GATS samples would be controlled by stratification so that roughly half of male and female respondents would come from rural areas and the other half from urban areas. Methods for achieving equal sample sizes for these four important subgroups are presented in the manual.

Specifications, Assumptions, and Approach

The statistical power to detect change between any two rounds of GATS is determined by two main features of the sample design; i.e., (i) by the combined multiplicative effect on the variance of estimates due to cluster sampling, stratification, and disproportionate allocation among strata, as measured by the

¹ Note that a sample of 8,000 responding households corresponds to a sample of 8,000 responding individuals since one eligible resident will be selected per household.

design effect (or *Deff*); and (ii) by whether or not the samples in the two rounds are selected independently from one another. As noted in the *Chapter 4, Deff*, the multiplicative increase in the variance of sample estimates, is almost always greater than one (Kish, 1965). *Deff* was assumed to be 2.00 in both the power and precision calculations below. Moreover, the precision and power calculations use standard formulae for the variance of the difference of estimates from independent samples. An approximation for the cumulative of a standard normal distribution, suggested by Polya (1945) and described in Section 13.5 of Johnson and Kotz (1970), was used to calculate Type II error and power.

The relationship between how samples are chosen among rounds also affects power and precision. Compared to round-specific samples that are chosen completely independently of one another, samples of the same size that are designed to overlap in some way will have greater power to detect change. The amount of increased power depends on how the samples overlap. Samples that are forced to overlap in the latter stages of selection and beyond will generally realize a greater increase in power than those that overlap only in the earlier stages. For example, if the samples in two comparison rounds are chosen from the same PSUs but the household samples within the common PSUs are independently chosen, the amount of increased power will be less than if the samples in the two rounds include the same households². Since it was not known in the early stages of planning GATS whether it would be practical to design overlapping samples, the power calculations assumed (conservatively) that the sample in each round of GATS would be selected independent of the sample chosen in other rounds. Thus, the power figures given below in this appendix will be a bit conservative if GATS ultimately elects to overlap samples among rounds somehow.

For countries not giving priority to reporting regional estimates, specifications for all GATS rounds require that samples used to produce estimates for each of the four urbanicity by gender population subgroups be of sufficient size to detect for:

- **Tobacco Use Prevalence Rates** — A decrease in tobacco use prevalence rates from 40% to 34% with >80% power and to 33% with >90% power, given Type I error at 0.05 and a two-sided alternative hypothesis;
- **Secondhand Smoke Rates** — A decrease in secondhand smoke rates (among the assumed 60% of sample respondents who are non-smokers) from 80% to 70% with >90% power, given Type I error at 0.05 and a two-sided alternative hypothesis;
- **Quit Rates** — An increase in quit rates (among the assumed 40% of sample respondents who are smokers) from 10% to 20% with >90% power, given Type I error at 0.05 and a two-sided alternative hypothesis.

Corresponding GATS precision requirements for the same country setting were that round-specific estimates of tobacco use prevalence rates of 40% must have a margin of error of 3 percentage points or less.

Findings

The seven figures presented below contain the values of power and precision that were used to determine the recommended household sample size of 8,000. More specifically, findings from these

² Note that in a stable population, two rounds with the same sampled households must have been chosen from the same PSUs, SSUs, etc., so they have complete overlap in all sampling stages.

figures helped determine the sample size that would be needed in a GATS country whose main analysis priorities were to produce estimates to specification for urbanicity by gender subgroups. Since having equal sample sizes for these four subgroups achieves this requirement with the smallest total sample size, four times the sample size meeting the GATS precision and power specification for the four subgroup estimates would indicate the recommended total sample size. Since (as will be seen below) a respondent sample size of approximately 2,000 was found to meet the three power requirements and the single precision requirement, $8,000 = 4 \times 2,000$ was adopted as the recommended minimum sample size requirement for GATS in which regional estimates are not required. In those countries where GATS-quality estimates are required for the four subgroups within regions, 8,000 was recommended as the minimally acceptable sample size *for each region*.

Figures A.1 to A.3 display the respondent sample sizes needed to detect change on a test of difference with Type I error of 0.05 and a two-sided alternative between two rounds of GATS for the three key tobacco use rates defined above, with “Time A” referring to the first round in sequence and “Time B” referring to a subsequent round. The required power level was set at 80% for a tobacco use prevalence rate, and at 90% for the secondhand smoke and quit rates. In each figure a series of possible Time B rates (including the one set by the power specifications for GATS) are plotted in comparison with the corresponding Time A rate specification, thus varying the size of detectable rate difference.

Three key findings in these first three figures led to the recommended respondent sample size of 2,000 for each of the four urbanicity by gender subgroups. First, notice in *Figure A.1* that the required respondent sample size for a decrease to a Time B tobacco use prevalence rate of 34% with 80% power is 2,024. Furthermore, the required respondent sample size in *Figure A.2* for a decrease to a Time B secondhand smoke rate of 70% with 90% power is 1,295, while in *Figure A.3* the required respondent sample size for an increase to a Time B quit rate of 20% with 90% power is 1,317. Thus, a sample size of 2,000 will provide nearly adequate power for comparisons of tobacco use prevalence, especially if Deff turns out to be less than 2.00 and given the more conservative assumption of sample independence and a two-sided alternative hypothesis of difference. With a sample size of 2,000 for each of the four key subgroups, clearly adequate power is indicated to detect differences between rounds in the secondhand smoke and quit rates. For these reasons, a respondent sample size of 2,000 was considered adequate to meet the power requirements for each of the urbanicity by gender subgroups.

Figures A.4 to A.6 contain plots of the level of statistical power corresponding to respondent sample size of 2,000 for each urbanicity by gender subgroup as a whole. For instance, note in *Figure A.4* that a respondent sample size of 2,000 would be able to detect a decrease in tobacco use prevalence from 40% at Time A to 34% at Time B with 79.52% power and a change from 40% to 33% with 90.50%, thus suggesting that GATS power requirements for these comparisons will be met. In addition, with a sample of 1,200 non-smokers out of a total subgroup sample of size 2,000, the power to detect a decrease in the secondhand smoke rate from 80% at Time A to 70% at Time B from *Figure A.5* is 98.26%, which is well above the required 90%. Moreover, with a sample of 800 smokers out of a total subgroup sample of size 2,000, the power to detect an increase in the quit rate from 10% at Time A to 20% at Time B from *Figure A.6* (98.11%) is also well above the required 90%. These findings thus further confirm the ability of a subgroup sample size of 2,000 to meet GATS power requirements.

Precision and Sample Size to Estimate Prevalence Rates for Each Round

Finally, **Figure A.7** indicates the respondent sample size requirements to estimate tobacco use rates between 10% and 90% with a margin of error (MOE) based on a 95% level of confidence. Separate plots are given for MOEs ranging from two to six percentage points. According to GATS precision requirements, MOEs for round-specific estimates of tobacco use prevalence of 40% for urbanicity by gender subgroups must be no greater than 3%. A sample of 2,049 meets this requirement precisely, with a sample of 2,000 implying a MOE of 3.04%.

References

1. Johnson, N.L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions – 1*, Houghton Mifflin Company, Boston, MA.
2. Kish, Leslie (1965) *Survey Sampling*, John Wiley and Sons, New York.
3. Polya, G. (1945). "Remarks on computing the probability integral in one and two dimensions," *Proceedings of the First Berkley Symposium on Mathematical Statistics and Probability*, pp.63-78.

FIGURE A.1

**Overall Sample Size Each Time to Detect Decrease in
TU Prevalence from 40% at Time A with 80% Power**

Deff=2.00; 2-Sided Alternative

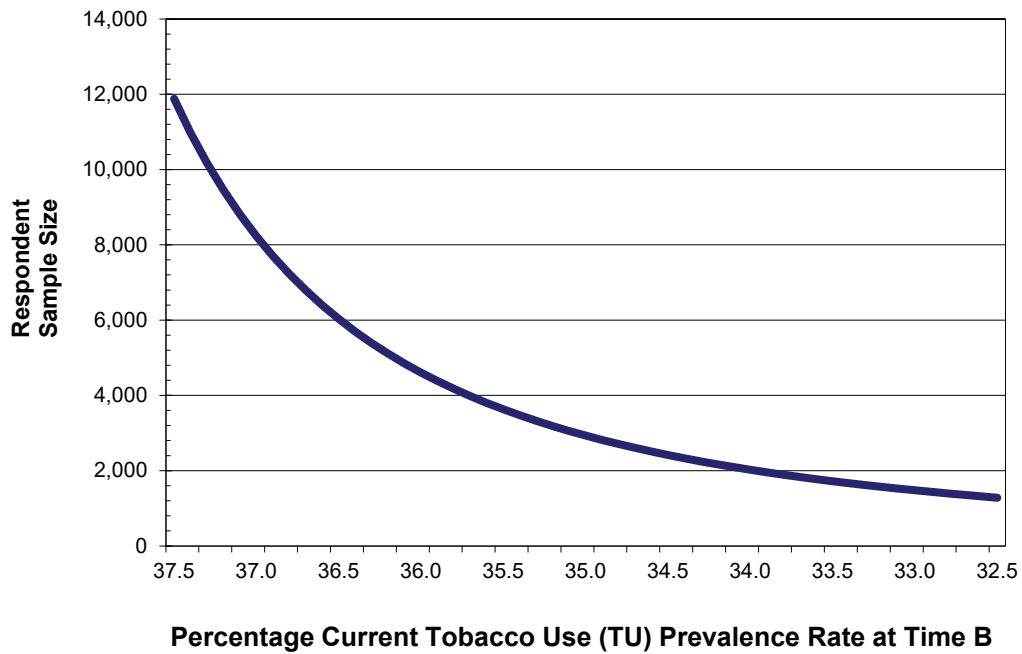


FIGURE A.2

**Overall Sample Size Each Time to Detect Decrease
in SHS Rate from 80% at Time A with 90% Power**

Deff=2.00; 2-Sided Alternative

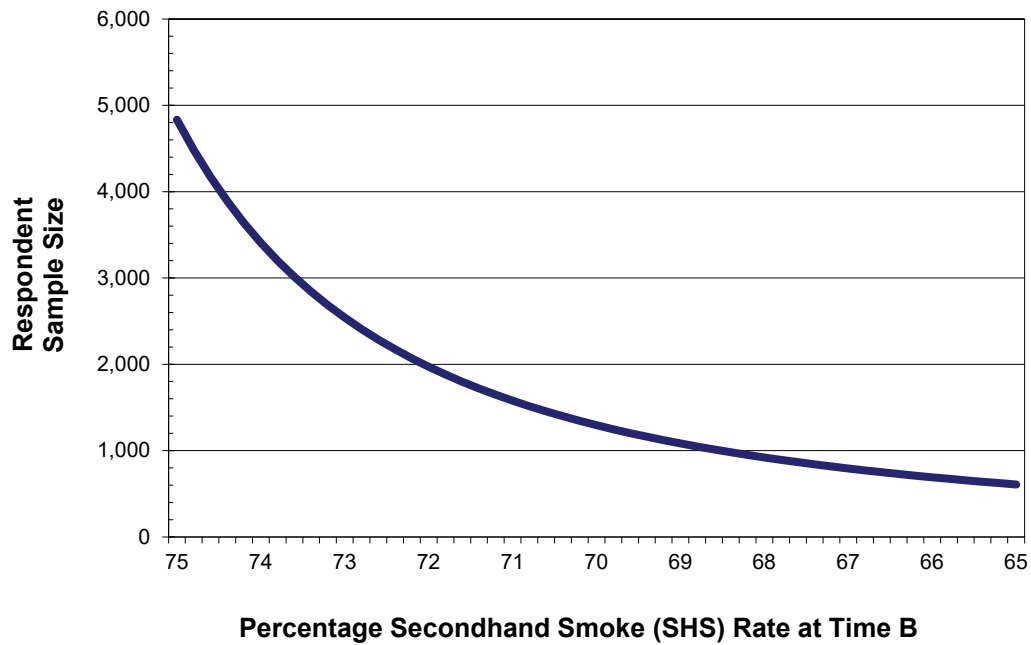


FIGURE A.3

Overall Sample Size at Each Time to Detect Increase in Quit Rate from 10% at Time A with 90% Power

Deff=2.00; 2-Sided Alternative

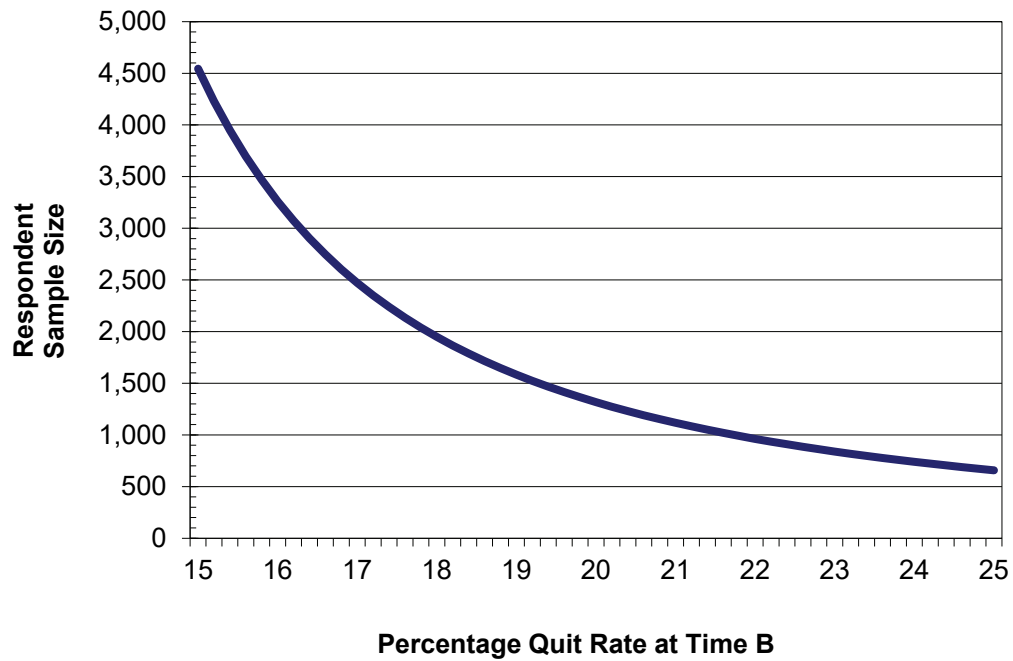


FIGURE A.4

**Power (in %) to Detect a Decrease from a
Time A Tobacco Use Prevalence Rate of 40%**

Time A and Time B Sample Sizes = 2,000; Deff=2.00; 2-Sided Alternative

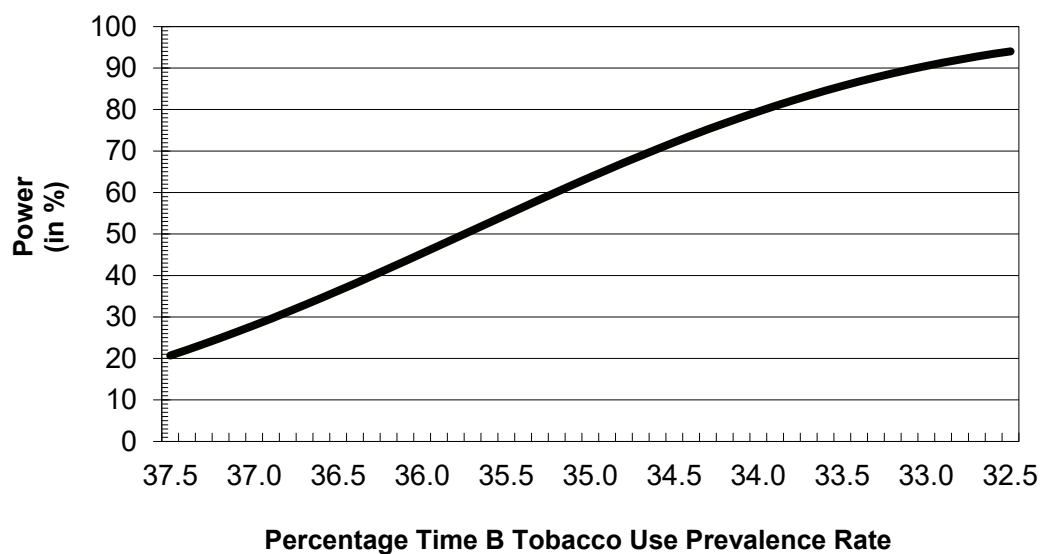


FIGURE A.5

**Power (in %) to Detect a Decrease from a
Time A Secondhand Smoke Rate of 80%**

Time A and Time B Non-Smoker Sample Sizes = 1,200; Deff=2.00; 2-Sided Alternative

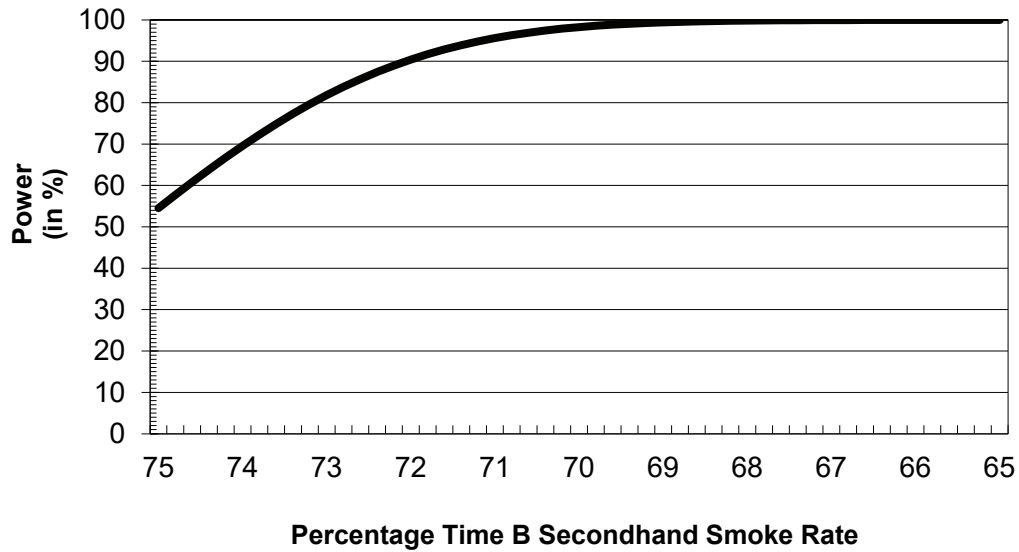
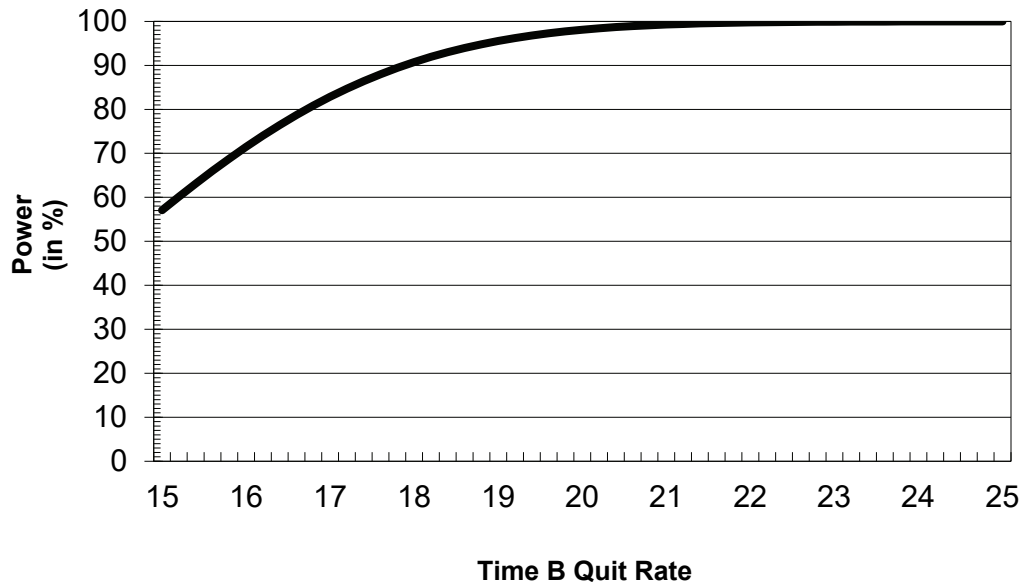
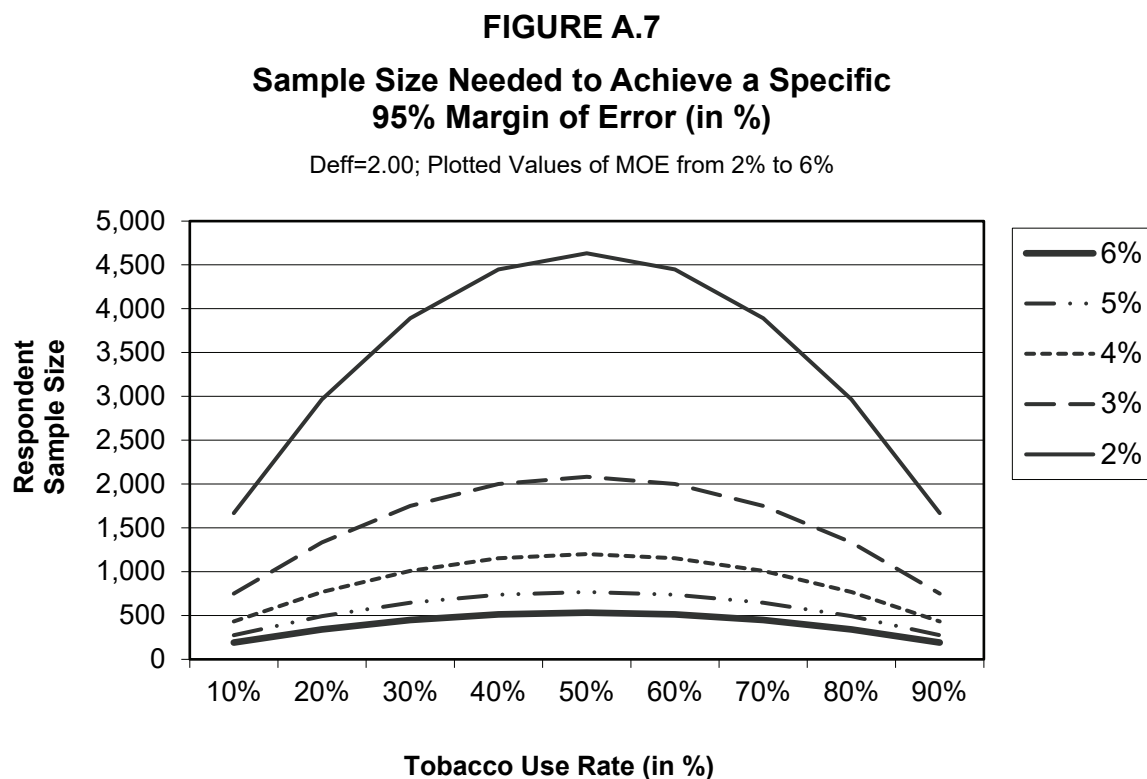


FIGURE A.6

**Power (in %) to Detect an Increase from
a Time A Quit Rate of 10%**

Time A and Time B Smoker Sample Sizes = 800; Deff=2.00; 2-Sided Alternative





Appendix B: Sample Design Options for New GATS Countries

This appendix describes three design options (OPTIONS N1-N3) that might be considered in conjunction with the development of sampling approaches for countries implementing GATS for the first time or as a baseline as compared to the countries that have already implemented GATS at least once. A short description of each option, along with as a listing of some of its advantages and disadvantages, are given below. GATS staff will work with in-country technical staff to decide on one of these options, seeking to find answers to questions that will more effectively facilitate the development of reputable sample designs.

The following are general considerations in developing the set of design options for new GATS country samples:

- All designs should be patterned after the specifications laid out in earlier sections of this manual.
- More specifically, all sample designs should use random selection in a way that the selection probability for all members of the resulting sample in each stage of selection can be determined. In so doing, the design will produce a **probability sample** of GATS respondents.
- PSUs should be recognized geo-political areal units, for which there are at least 1,000 in the country, and at least 100 PSUs should be sampled in the first selection stage overall, or for each region of the country for which GATS-quality estimates are desired.
- For designs accommodating the production of GATS-quality estimates jointly by gender and urbanicity designation, it is recommended that there be at least 8,000 respondents.
- The sample of households should be chosen in a minimum of two sampling stages.
- Integrating the sample design for GATS with another reputable national survey sample is acceptable and encouraged, especially if it will save on data collection costs.

OPTION N1: Stand-Alone Design

A GATS stand-alone design (i.e., a sample is designed and selected for GATS only), with the design of the sample following the standard GATS sampling protocol for structure and size. An overall respondent sample size of $n=8,000$ is recommended, since $n=2,000$ are needed to meet GATS statistical quality standards (to estimate round-specific tobacco use rates and to detect change in these rates from one round to another) for each of four gender-by-urbanicity population subgroups.

- The male and female tobacco use prevalence rates are presumed to be relatively similar.
- $n=8,000$ respondents per region if GATS-quality regional estimates are required.

Advantages

- Most familiar approach, since it is the design that is described in the *GATS Sample Design Manual* and frequently used in Phase I countries.
- It is the best route to assure general design comparability with findings *from other GATS countries*.

Disadvantages

- Size of the sample may make it too expensive if available funding for data collection is limited.
- Requires in-country partner organization(s) with strong prior sampling and survey experience. Thus, more technical assistance may be needed if there is limited prior experience in conducting in-person national household surveys.

OPTION N2: Smaller Stand-Alone Design

GATS stand-alone following the standard GATS sampling protocol for structure but having a smaller sample size that is sufficient to produce estimates of acceptable GATS quality for the two gender groups and the two urbanicity subgroups separately. An overall respondent sample size of $n=4,000$ is recommended since GATS-acceptable sample sizes of $n=2,000$ can be realized marginally for each of two gender or urbanicity key subgroups.

- It is not necessary to produce estimates of adequate precision for the four gender-by-urbanicity population subgroups.
- The male and female tobacco use prevalence rates are presumed to be relatively similar.
- The sample may need to be disproportionately sampled by urbanicity if a 50:50 split is needed for the urban:rural sample size.
- See *Section 7* of this manual to clarify sample size requirements and recommendations for each stage of sampling, particularly noting the minimum number of sample PSUs.
- $n=4,000$ is recommended per region if regional estimates of GATS quality are needed.

Advantages

- Shorter and less expensive data collection phase because of smaller sample size.
- An adequate route to assure comparability with findings from other GATS countries.

Disadvantages

- More limited quality of overall, urban-rural, and other smaller subgroup findings than OPTION N1 if nothing is done to disproportionately sample by urbanicity.
- Requires in-country partner organization(s) with strong prior sampling and survey experience. Thus, more technical assistance may be needed if there is limited prior experience in conducting in-person national household surveys.
- It is less likely (compared to OPTION N1) that regional estimates will approach acceptable quality if the overall sample size is limited to $n=4,000$ respondents.

OPTION N3: Integrated Design

Integrate the GATS sample with an existing, highly reputable, national household sample. The existing sample may be a country's master sample or another survey (e.g., DHS, a major ongoing social survey conducted by the country's main statistical agency, a highly regarded national health survey, etc.). Sample sizes may vary depending on analysis needs (see OPTIONS N1 and N2).

- Definition: sample design integration occurs when two or more samples share at least a portion of their selected sampling units from one or more of their selection stages.
- GATS would typically be the “recipient” design of the sample from the existing sample design, meaning that all or a random subsample of the households from the existing sample is chosen for GATS.
- Integration of sample designs may imply having data collection for two surveys to be completed by one survey organization more or less simultaneously.
- Like a stand-alone design, this design option can be smaller or larger depending on the population subgroups (defined by gender and/or urbanicity categories) for which adequately precise estimates are required. The GATS sample size proposed under this option, will depend on available resources and the size of the household sample with which the GATS sample is integrated.

Advantages

- More and more countries have existing master household samples and are using them to do various national surveys to accommodate their growing survey information needs. Moreover, many of these master samples are designed following established principles of area probability sampling, and are developed in consultation with those who are experienced in applying these principles.
- The average cost per respondent to integrate GATS sampling and data collection with a master/existing household sample and its associated data collection operation may be significantly less because of process efficiencies (i.e., one survey data gathering infrastructure to collect multiple sets of data, perhaps simultaneously).
- The country may be more likely to sustain the GATS survey system since in carrying out this option, GATS will be working with those who are able to design and implement an established and presumably high-quality governmental data system.
- Requires the complete cooperation of those who have produced the existing sample (to work out a feasible plan for integration, to get weights computed for the final GATS sample, to effectively document the GATS sample, to make the two sample designs compatible for analysis, etc.).
- The quality of the final GATS sample depends on the quality of the design and selection of the existing sample.

Disadvantages

- If subsampling is involved, one challenge is to design the subselection approach so that it yields a subsample that meets the GATS standards (e.g., the difficulty encountered in developing the PSU subsampling approach in Viet Nam, so that the net effect was to produce a PPS sample of PSUs).
- Unable to produce base weights in the sample weighting process if selection probabilities are unknown for the existing sample.
- Likely to require (perhaps much) more technical assistance and support if in-country technical staff are inexperienced or less-equipped to handle the complexities of two-phase sampling.
- GATS is constrained by the design features and data collection timetable of the existing sample (e.g., definitions of strata).
- Comparability with other GATS countries will occur to the extent that the survey design of the existing sample is consistent with the standards set by the *GATS Sample Design Manual*.
- Analysts of GATS data have the technical challenge of accommodating the two-phase sample design, if the GATS sample is a random subsample of the existing sample.
- Requires the complete cooperation of those who have produced the existing sample (to work out a feasible plan for integration, to get weights computed for the final GATS sample, to effectively document the GATS sample, to make the two sample designs compatible for analysis, etc.).
- The quality of the final GATS sample depends on the quality of the design and selection of the existing sample.

Appendix C: Sample Design Options for Repeat GATS Countries

This appendix describes possible directions in sample designs in which subsequent rounds of GATS could go in repeat countries. Although this discussion applies to any pair of rounds in a country where multiple rounds of GATS are completed, “baseline” round will generally refer to the first round in time sequence and “repeat” round will refer to a subsequent round.

Three options (R1–R3) are presented and briefly discussed for designing the sample for a repeat round of GATS. These options presume that any of the N options (or some variation on them) discussed in *Appendix B* would have been used in the baseline round. It will become apparent that the R options discussed here vary according to degree to which the composition of the repeat round sample depends on the composition of the baseline sample. As with new countries, GATS staff will collaborate with in-country technical staff of repeat countries and GATS Sampling Review Committee to decide on one of the indicated options.

The following are some issues to consider in developing the set of sample design options for repeat GATS country samples:

- As in all rounds of GATS in participating countries, the design should be patterned after the specifications, as laid out in the *GATS Sample Design Manual* which generally calls for the following design features:
 - *Random selection in all sampling stages to produce known selection probabilities for all sample members (i.e., the sample design must produce a “probability sample” of GATS respondents).*
 - *A minimum of 100 sample PSUs.*
 - *At least two sampling stages are required to choose households.*
 - *Integrating the sample design with another reputable national sample design is encouraged, especially if this worked well in the baseline round of GATS.*
 - *The overall respondent sample size should be at least 8,000 if GATS-quality estimates are required for population subgroups jointly defined by gender and urban-rural categories.*
 - *If resources are less than those needed to produce this sample size, the quality implications of smaller respondent samples size should be discussed with members of the GATS Sampling Review Committee so that there is a clear understanding of what the resulting sample will produce.*
- The discussion of options should begin with a detailed review of the sample design that was used in the baseline round of GATS. This is needed to orient the design discussion to what might be possible in the repeat round.
- Comparability among GATS rounds of data collection is achieved in a country if the designs in each round meet the basic standards of the *GATS Sample Design Manual*, which allows some latitude in the specifics of how each GATS randomized probability sample is chosen. Thus, it is allowable for design details to vary among rounds (e.g., stratum definitions, sample sizes, choice of sampling units, etc.). There are, however, some statistical and practical advantages to following the exact same sample design in each round. For instance, using the same household

sample in each round: (i) creates planned sample overlap that will produce somewhat more precise estimates of round-to-round differences (than independently chosen samples), and (ii) may make it easier to recruit field interviewers since the same interviewers used in the first round could be hired for the second round.

The most statistically useful but also operationally difficult approach to sampling in a repeat round of GATS would be to return to *all* of the same *selected* sample of household locations (i.e., residential addresses) that were assigned to the field staff in the baseline round, thus making a complete cohort sample of selected households the statistical end-result for the two rounds. In a “complete cohort” approach, interviewers would return to *all selected* household locations from the baseline (or first) round, including those that: (i) led to completed respondent interviews, (ii) were found to be ineligible for the study, (iii) became eligible household nonrespondents, and (iv) where the household responded but did not lead to a respondent interview. Household residents found at the sampled locations at the time of data collection for the repeat round under this approach would be recruited to complete a GATS interview for that round, even if the set of eligible household residents had changed since the baseline or a completely new group of residents had moved there. Thus, only a portion of a completely overlapping sample of residential locations for the repeat sample would involve the same households and individuals as the baseline round.

Completely overlapping selected samples in a multi-round survey is statistically beneficial because they enhance the quality of estimated round-to-round differences. The feature of this multi-round design strategy that makes it beneficial is the planned (versus unplanned or random) overlap in the household samples between the two rounds¹.

Two samples with planned (as opposed to unplanned, or random) overlap will generally produce more precise estimates of differences between the two sample-specific estimates than between two independently chosen samples. Moreover; the higher the percent of planned overlap, the greater the precision gain in estimated differences. Specifically, if B and R are baseline and repeat round estimated rates, respectively, then the variance of their difference is, $Var(R - B) = Var(R) + Var(B) - 2\theta Cov(R, B)$, where $0 \leq \theta \leq 1$ is a measure of the effect of planned overlap for the two sample designs. Since $Cov(R, B)$ is usually positive and θ would perhaps begin to approach 1 in a complete overlap approach, if the baseline and repeat round designs differ only by the amount of planned overlap they have, $Var(R - B)$ will be less under this approach than for any other design option where there is a lower percent of planned overlap (i.e., the options, R1-R3, described next)².

Assuming response rates were similar for baseline and repeat rounds, this approach would result in roughly the same respondent sample size for the repeat round as the baseline round. If a larger response sample size were required for the repeat round, the baseline sample could be supplemented as needed. Similarly, if a smaller sample was needed, a portion of the sample could be randomly deleted (see OPTION R1, below).

¹ Unplanned overlap is that which occurs by chance alone (i.e., population members happen to be chosen in both samples). By comparison, planned overlap occurs by designating some portion of the baseline sample to be members of the repeat round sample. Planned overlap among members of the two samples may be complete (resulting in a cohort sample) or partial.

² θ will not equal 1 even in a complete cohort of selected household locations since household composition and/or the chosen resident for interview will differ in some of the overlapping locations.

Because of its reliance on the availability of records of locator information for selected households that were chosen for the baseline round (which, by the time that the repeat round is begun, could be several years previous), and the need to somehow deal with new residential construction and the shift in the composition of the target population of individuals 15 years of age or older, it seemed impractical to consider this approach for the sample in a repeat round. For these reasons, it was decided (in consultation with CDC staff) not to make a complete cohort sample of selected households one of the options to consider for repeat Phase I countries.

OPTION R1: Partially Overlapping Design

In this design option, the plan is to return to a *portion of* (the same) *selected* sample of household locations that were assigned to the field staff in the baseline round, thus making a partial cohort sample of selected households the statistical end-result. While stopping short of using independently chosen samples at each round, this is a seemingly more practical alternative to baseline and repeat round samples that are completely overlapping. Like completely overlapping samples, this option can produce more precise estimated differences over time than independently chosen samples. Its gains in precision over independently chosen samples are not as great as completely overlapping samples, however.

- *Partial overlap* of two samples occurs when a random subset of members of the first sample are designated to be members of the second sample.
- This design option could occur in any of the following situations: (i) the country developed its own stand-alone sample for the baseline and wants to save data collection costs by using the same PSU sample from the baseline, but wants/needs to select a new household sample with each PSU for the repeat round, (ii) the country wants to once again integrate the GATS sample with another survey and a portion of the master (donor) sample must be re-drawn, or (iii) the country wants a smaller respondent sample size and to use the same selected household sample as in the baseline.
- Overlap may apply to any or all sampling stages down to household selection, noting that applying it completely to all of these stages (i.e., 100% overlap in all sampling stages) produces a completely overlapping selected household sample. Thus, a partially overlapping design is one where <100% overlap occurs in at least one stage.
- Overlap in the later sampling stages (e.g., at the segment or household level) is more statistically effective than overlap at the PSU stage of sampling, since more of the variation in key study outcome variables tends to be found in sampling units used at the later stages of selection. Due to the limited improvement in precision in overlapping PSU designs and complexity in analysis overlap in PSUs will likely be ignored in actual analysis.
- Note that comparability of baseline and repeat round estimates *does not depend* on changes that may be made in the probability sampling approach for the repeat round, but *rather strongly depends on* consistency in key elements of the survey design proposed for the repeat round (e.g., questionnaire wording and context, source and quality of the household lists used to sampled households, sample recruitment protocol, quality control plan, field staff recruitment and supervision approach).

Advantages

- This option can produce more statistically precise baseline round — repeat round estimated differences; the higher the percent of planned overlap, the greater the precision gain in estimated differences. Specifically, $Var(R - B)$ for OPTION R1 would be less than $Var(R - B)$ for any design option involving independently selected samples (i.e., OPTIONS R2 or R3), since $\theta = 0$ for independently chosen samples and thus $Var(R - B) = Var(R) + Var(B)$ which misses the variance-reducing benefit of the covariance arising out of planned overlap. If sample sizes do not change much, the quality of cross-sectional estimates at the repeat round would be virtually the same as at baseline.

Disadvantages

- One limitation of this option is that, if a portion of sampled households overlap by design, interviewing staff must be able to locate and get individuals at the same address to respond. In most instances they will be dealing with the same households and residents as before, so the same individuals would need to be convinced to agree to participate in the GATS interview a second time.
- Field staff would need to be trained and supervised to use great care in verifying that each assigned residential address in the overlap portion of the repeat round sample is the same as the corresponding sample address from the baseline sample.
- Analysis to produce variance estimates of repeat round — baseline round difference estimates would be more complex than the other options, since a portion of the sample has covariance arising out of the planned overlap to accommodate while the rest of the samples are independently chosen. The alternative would be to ignore this covariance term but that would cause variances of differences to be overstated.
- If sample sizes do not change much, the quality of cross-sectional estimates at the repeat round would be virtually the same as at baseline.

OPTION R2: Identical Structure but Independent Selection Design

In this design option the structure (i.e., sampling units and strata) of the repeat sample is identical to that of the baseline sample, but samples in each of the selection stages are independently drawn for the repeat round. Depending on the country's wishes for the repeat round, the sample sizes at each stage of selection of the repeat round sample design may or may not differ from the baseline sample design.

- This design option could occur in any of the following situations: (i) the country wishes to use the same sample design as in the baseline but wishes/needs to select a completely new sample for the repeat round using the same stage-specific sampling frames as used to select the baseline sample, or (ii) the country wishes to use the same sample design as in the baseline but wishes to select a completely new sample for the repeat round using stage-specific sampling frames that have been updated from (and thus are presumably more current than) those used in the baseline.
- Depending on the country's wishes for the repeat round, the sample sizes at each stage of selection of the repeat round sample design may or may not differ from the baseline sample design.

- If the country produced a design summary table for the GATS sample design in the baseline round (i.e., a tabular summary of the sampling units, use of stratification, selection methods, and sample sizes for each sampling stage), this table for the repeat round would look precisely the same.
- Even though the sample design structure (i.e., the number of stages, sampling units, and use of stratification) is the same for both rounds under this option, the baseline and repeat round samples of respondents would be statistically independent since samples in each stage are independently chosen for each round.
- Note that comparability of baseline and repeat round estimates *does not depend* on changes that may be made in the probability sampling approach for the repeat round, but *rather strongly depends on* consistency in key elements of the survey design proposed for the repeat round (e.g., questionnaire wording and context, source and quality of the household lists used to sampled households, sample recruitment protocol, quality control plan, field staff recruitment and supervision approach).

Advantages

- Independently chosen samples have no covariance term to accommodate in estimating variances of estimated repeat round - baseline round differences, so analysis is less complicated (i.e., the variance of the estimated difference is simply the sum of the variance of the baseline round estimate plus the variance of the repeat round estimate, or $Var(R - B) = Var(R) + Var(B)$).
- If the number of sampling stages, the definitions of sampling units in each stage, and the sample sizes for each stage are the same as in the sample design for the baseline round, the quality of cross-sectional estimates at the repeat round will be virtually the same as at baseline.

Disadvantages

- If the number of sampling stages, the definitions of sampling units in each stage, and the sample sizes for each stage are the same as in the sample design for the baseline round, the quality of cross-sectional estimates at the repeat round will be virtually the same as at baseline.
- The statistical precision of estimated differences would not be as high as for those from OPTION R1. Specifically, $Var(R - B)$ for OPTION R1 would be less than $Var(R - B)$ for a repeat design option involving independently selected samples, since $\theta = 0$ for independently chosen samples and thus $Var(R - B) = Var(R) + Var(B) - 2\theta Cov(R, B) = Var(R) + Var(B)$ for this option, thus missing the variance-reducing benefit of the covariance arising out of planned overlap.

OPTION R3: Completely Independent Selection Design

In this design option the country relies on a completely different sample design in the repeat round than was used in the baseline round. This means that different sampling units may be used in earlier selection stages and another individual and/or statistical organization will provide the sample.

- This design option could occur in any of the following situations: (i) a new statistical organization will provide the sampling expertise to select the repeat round sample than was used for the baseline round, (ii) the sample from a different national survey is used to produce a sample for GATS that is integrated with another survey sample or master sample, or (iii) the GATS sample

for the baseline round was integrated with another survey or master sample, but the country decides it wishes to create a stand-alone design for the repeat and subsequent rounds.

- Depending on the country's wishes for the repeat round, the sample sizes at each stage of selection of the repeat round sample design may or may not differ from the baseline sample design.
- Note that comparability of baseline and repeat round estimates does not depend on changes that may be made in the probability sampling approach for the repeat round, but rather strongly depends on consistency in key elements of the survey design proposed for the repeat round (e.g., questionnaire wording and context, source and quality of the household lists used to sampled households, sample recruitment protocol, quality control plan, field staff recruitment and supervision approach).

Advantages

- Independently chosen samples have no covariance term to accommodate in estimating variances of estimated repeat round — baseline round differences, so analysis is less complicated (i.e., the variance of the estimated difference is simply the sum of the variance of the baseline round estimate plus the variance of the repeat round estimate, or $Var(R - B) = Var(R) + Var(B)$).
- If the number of sampling stages and the sample sizes for each stage are at least approximately the same as in the sample design for the baseline round, the quality of cross-sectional estimates at the repeat round will be about the same as at baseline.

Disadvantages

- If the number of sampling stages and the sample sizes for each stage are at least approximately the same as in the sample design for the baseline round, the quality of cross-sectional estimates at the repeat round will be about the same as at baseline.
- If different statistical organizations, with varying quality of sampling frames, sample recruitment strategies, and survey measurement, are involved in producing the samples for the two rounds, a possible downside to this option is that comparison of survey estimates may be confounded by these quality differentials.

The statistical precision of estimated differences would not be as high as for those from OPTION R1. Specifically, $Var(R - B)$ for OPTION R1 would be less than $Var(R - B)$ for a repeat design option involving independently selected samples, since $\theta = 0$ for independently chosen samples and thus $Var(R - B) = Var(R) + Var(B) - 2\theta Cov(R, B) = Var(R) + Var(B)$ for this option, thus missing the variance-reducing benefit of the covariance arising out of planned overlap.

Appendix D: Survey Design Checklist

The *GATS Sample Design Manual*, *GATS Mapping and Listing Manual*, and *GATS Sample Weights Manual* are designed to offer both requirements and recommendations, as well as suggested guidelines for countries to follow as they develop appropriate sample design, data collection and weighting procedures for their implementation of GATS. The first step in this survey design process is the preparation of the GATS Sampling Design Proposal which is submitted to the GATS Survey Review Committee (SRC) for review and approval. The Sampling Design Proposal should be completed by trained statisticians identified in each country who will have primary responsibility for GATS sample selection, collaborating with country public health representatives.

Country statisticians should read and understand each relevant GATS manual fully before developing their sample design proposal. The CDC Focal Point assigned to assist countries in the design and implementation of GATS will work closely with statisticians and other country representatives in the preparation of the proposal. CDC survey statisticians are also available to provide technical assistance on sample design, selection and weighting. The following checklist offers country survey statisticians and public health representatives with the primary elements of survey design that should be described in the Sampling Design Proposal.

Once completed and reviewed by CDC, the proposal will be submitted to the SRC for formal review and approval. SRC review is an iterative process in which SRC members comment on the survey elements described in the Sampling Design Proposal and make suggestions for modifications to improve survey design and better align design elements with the GATS survey design protocol. Every sample design proposal should include sections that address the following review criteria. The list below is not a complete list of all things to include in the proposal, but rather items that are often overlooked or underspecified thus causing delays in completing the proposal review by the SRC.

In parentheses, we have provided sections of the GATS manuals that will be useful in preparing the GATS Sampling Design Proposal.

Survey Objectives (refer to *GATS Sample Design Manual Chapter 2*)

- ☐ Planned estimation domains/key reporting subpopulations and intended comparisons (cross sectional and/or across time in repeat surveys) are fully described.

Target Population & Sample Frame (refer to *GATS Sample Design Manual Chapter 3*)

- ☐ Target population is fully defined, including a listing of the criteria for eligibility.
- ☐ Any geographic exclusions explicitly described and proportions of country population excluded provided.
- ☐ Household residence requirement is described with any deviation and justification from GATS household residence criteria provided.
- ☐ A description of the area sampling frame is provided including source, recency and the count of areal clusters on the frame. Provide an explanation if the frame count is <1,000.

- ☐ If a Master Sample will be used as the frame to select the GATS sample rather than last census provide:
 1. Reasons for using a Master Sample instead of creating a new sample just for GATS.
 2. Detailed information about the Master Sample design structure (i.e., # of selection stages, sampling units in each selection, and selection methods in each stage).
 3. Source and means of computing probabilities of selection for sample selection in each stage of the Master Sample.
 4. Detailed description of the approach to be followed in subsampling the Master Sample for GATS subsample are described to create the GATS sample.
- ☐ Count and percent of the country's population in each of the proposed explicit strata.
- ☐ Definition of household to be used in creating the sampling frames for the household stage of sampling.

Summary of Sample Design Features (refer to *GATS Sample Design Manual Chapter 4*)

- ☐ A detailed definition of explicit and implicit strata is provided with description of frame sorting for selection within implicit strata.
- ☐ Sample sizes selected and final respondent sample sizes by explicit strata are provided for each stage of sample selection. These counts are consistent across survey proposal.
- ☐ Proportionality/disproportionality in sampling by strata is discussed and justified. Reasons for equal, proportionate, or some compromise allocation between urban and rural respondents is presented.
- ☐ PSU sampling rate is provided. Justification is made if rate exceeds 10%.
- ☐ Sample Design Summary Table is provided that fully describes sample design.

Forming Primary Sampling Units (PSUs) (refer to *GATS Sample Design Manual Chapter 6*)

- ☐ The primary sampling unit (PSU), the sampling unit in the first stage of selection, is explicitly defined.
- ☐ Average number of expected participating households per PSU is presented.

First Stage of Sampling: Selecting PSUs using Probability Proportionate to Size (refer to *GATS Sample Design Manual Chapter 7*)

- ☐ Specific PPS selection algorithm (e.g., Sampford's PPSWOR, PPSWR, PPS Systematic, etc.) and program used to select the PSU sample is provided.
- ☐ If PPS systematic selection is proposed for PSU selection, implicit stratification is described along with a plan to form pseudo stratum pairs of PSU to calculate variances for use later in analysis.

Intermediate Stages & Selecting Households (refer to *GATS Sample Design Manual Chapter 8* and *GATS Mapping and Listing Manual Chapters 4-5*)

- ☐ Secondary sampling unit (SSU) formation and the sampling frames for intermediate stages of selection are fully described.

- ☐ The sample selection method to be used in each intermediate and household stage of sampling is fully described.
- ☐ Description of how household lists will be formed.
- ☐ Steps proposed to assure the quality, accuracy, and completeness of household lists are described.
- ☐ Field worker supervision and training for household listing is provided.
- ☐ Justification for gender randomization, if used, is provided.

Selecting Individuals within Screened Households (refer to *GATS Sample Design Manual Chapter 9*)

- ☐ Method of selection of individuals within households is provided with justification of the proposed approach differs from the GATS protocol.

Determining Sample Sizes at Each Stage of Selection & Reporting (refer to *GATS Sample Design Manual Chapter 10*)

- ☐ Where sample sizes deviate from GATS recommendations, respondent sample sizes based on precision requirements by key reporting populations are given with mathematical justification. These calculations should, if possible, be based on Design Effect (DEFF) values from recent similar country health surveys.
- ☐ Sample sizes for repeat surveys are computed using DEFF values available from prior GATS rounds and considering power to detect difference in rates between baseline and repeat rounds. 80% power is preferred for all key reported differences. (Note: Power calculations have already been done for many GATS countries, and are available for use by country statisticians by contacting the CDC focal point for the country.)

Computation plan for sample base weights (refer to *GATS Sample Design Manual Chapter 11* and *GATS Sample Weights Manual Section 3.1*)

- ☐ Detailed explanation of how stage-specific selection probabilities, and thereby sample base weights, will be calculated with a description.
- ☐ If the weights will be adjusted for differential response rates among sample subgroups, the calculation of these adjustments should be described in detail.
- ☐ Describe in detail how the sample weights will be calibrated to population census counts through post-stratification or iterative raking.
- ☐ Describe the prior experience of the country organization in computing sample weights. Verification that individual components of the sample base weights will be made available to the SRC for quality assurance review.

Global Adult Tobacco Survey (GATS)