



Global Adult Tobacco Survey (GATS)



Quality Assurance: Guidelines & Documentation

Global Adult Tobacco Survey (GATS) Quality Assurance: Guidelines and Documentation

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): Quality Assurance: Guidelines and Documentation*. 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 Quality Assurance for GATS	1-3
2. Quality Assurance Process Plan	2-1
3. Quality Assurance: Pre-Data Collection	3-1
3.1 Questionnaire Design and Review Process	3-1
3.2 Programming and Version Control	3-2
3.3 Pretest	3-4
3.4 Questionnaire Finalization	3-4
3.5 Data Transfer and Management	3-5
3.6 Sample Design and Review Process	3-5
3.7 Master Sample Selection File Preparation	3-6
3.8 Training Workshops	3-6
4. Quality Assurance: Data Collection and Management	4-1
4.1 Field Data Collection: Setup and Maintenance	4-1
4.2 National Data Center Data Management	4-5
5. Quality Assurance: Post-Data Collection	5-1
5.1 Cleaning and Preparing Data for Sample Weight Calculations	5-1
5.2 Measures of Quality: Sampling, Sampling Error, and Weights	5-5
5.3 Measures of Quality: Coverage, Nonresponse, and Other Non-Sampling Errors	5-10
5.4 Formal Review of Statistical Quality	5-15
5.5 Creation of Analytic Data File	5-17
5.6 Public Use Data File Creation	5-17
6. Data and Documentation Requirements	6-1
7. Bibliography	7-1
Appendix A: Glossary of Terms	A-1
Appendix B: Pre-Data Collection	B-1
B.1 GATS Questionnaire Adaptation and Review Process	B-1
B.2 GATS Questionnaire Programming Process	B-2
B.3 Process for Setting Up Handhelds with GSS Program	B-2

Appendix C: Post-Data Collection.....	C-1
C.1 Final Disposition Codes and Response Rate Calculations.....	C-1
C.2 Pattern of Post-Stratification Weights Calibration Adjustments Among Adjustment Cells	C-4
C.3 Effect of Variable Sample Weights on the Precision of Survey Estimates	C-6
C.4 Overall Design Effect on the Precision of Survey Estimates and the Within-PSU Intracluster Homogeneity of Corresponding Key Survey Estimates	C-7
C.5 Margin of Error for Key Survey Estimates.....	C-9
C.6 Estimates of Sampling Errors.....	C-11
C.7 Household Frame Coverage Rate	C-17

Exhibits

Number	Page
Exhibit 2-1. Quality Assurance Flow Diagram.....	2-1

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 provides guidance in assessing and ensuring the quality of data collected in GATS. It is a supplement to existing GATS manuals in that it will not include descriptions of procedures documented elsewhere. Instead, it describes the elements of the quality assurance process that should take place throughout the implementation of GATS. Adherence to these quality assurance guidelines is extremely important to the success of this survey. The manual is intended for use by all individuals involved in GATS implementation and quality assurance process including IT personnel and the statisticians responsible for sampling and weighting. A Glossary of Terms is provided in *Appendix A*.

In this document, the GATS quality assurance process is organized into three chronological stages: Pre-Data Collection, Data Collection and Management, and Post-Data Collection. Quality assurance activities vary for different stages of the quality assurance process, and guidelines are provided for every stage. *Chapter 2* begins by providing a process plan describing the overall scope of quality assurance activities.

Chapter 3: Pre-Data Collection includes sections describing quality assurance activities that should be carried out during the design and review of the country-specific questionnaire, programming of the questionnaire for the handheld used during data collection, sample design, and creation of the Master Sample Selection file.

Chapter 4: Data Collection and Management is divided into three sections. The first section addresses quality assurance activities performed in the field including the setup and maintenance of the case management system and questionnaire. It also includes sub-sections regarding data management and sets the standards for routine reporting during field work. The second section describes data management and quality assurance that occur once the interviews have been conducted and the data are uploaded to the central site. The third section refers to the aggregation and quality control issues that take place at the national level. This section also describes the preparation of the raw data file for cleaning.

Chapter 5: Post-Data Collection describes quality assurance activities to be carried out once the field work is completed and the raw data file has been prepared for cleaning. The first section outlines procedures for cleaning and preparing the data for sample weight calculations. The second section describes assessing the quality of the sampling, sampling error, and sample weights while the third

section focuses on quality assessment of coverage, nonresponse, and other non-sampling errors. Finally, the fourth section outlines the creation of the final analytic data file.

These chapters are intended to provide quality assurance guidelines and recommendations as best practices. The detailed methodologies/documentation for each section of the chapters are provided in the appendices, as applicable.

1.3 Quality Assurance for GATS

Quality assurance is a process consisting of systematic activities designed to ensure, assess, and confirm the quality of the data collected during a survey (Biemer and Lyberg, 2003). In the past, data quality was synonymous with data accuracy, but the term has evolved into a more general concept. High quality data is simply defined as data that are suitable for its intended purpose. This includes not only accuracy but also attributes such as timeliness, accessibility, and comparability, making quality a multi-dimensional concept. A dataset is accurate to the extent that it is free of errors. It is timely if it is available at the time it is needed. Accessibility is determined by the relative ease or difficulty of use. Quality data are comparable if they are the same from one unit to another, whether that unit comparison is between individuals, interviewers, PSUs, or international borders. Standardization increases the comparability of survey data, for example, the use of standardized questions ensure that different interviewers will all ask the same questions in the same ways. Similarly, the standardization of quality assurance processes ensures that data errors have been resolved in the same ways.

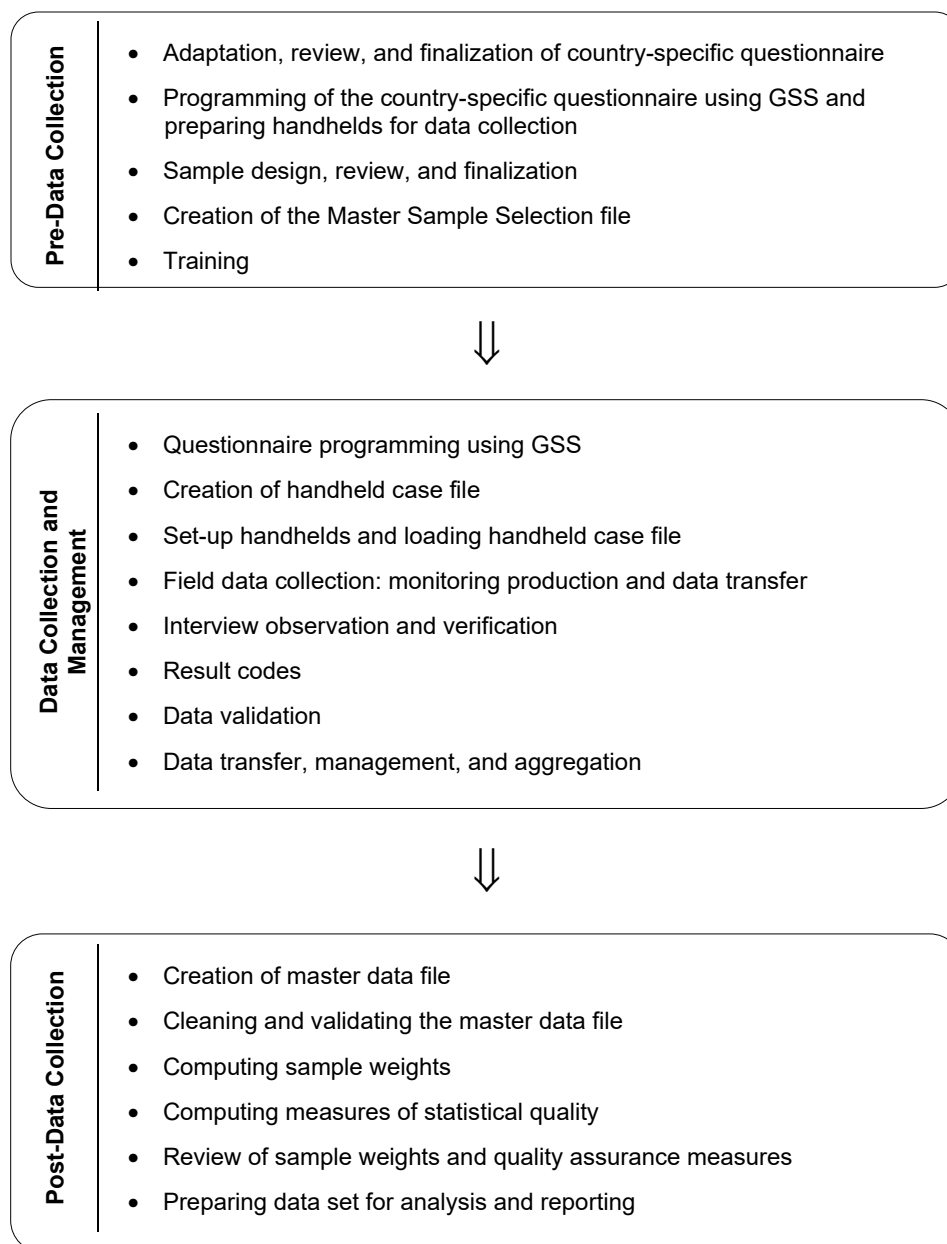
The multiple dimensions of quality often constrain one another, so maximizing quality is a matter of compromise and balance among all of them. For example, high quality data should be accurate but must also be timely. If it takes a long time to ensure a very high degree of accuracy, the data may become irrelevant.

The guidelines described in this document represent standardized procedures for the quality assurance of data collected by the implementation of GATS. Countries are encouraged to incorporate additional quality assurance activities beyond those described in the document.

2. Quality Assurance Process Plan

The following exhibit provides a diagram of quality assurance activities in terms of general timing and flow of non-overlapping steps. These steps are organized under three stages: pre-data collection, data collection, and post-data collection.

Exhibit 2-1. Quality Assurance Flow Diagram



3. Quality Assurance: Pre-Data Collection

The pre-data collection phase includes a series of tasks that must be completed in order to prepare for the full survey data collection. These tasks include questionnaire design, programming, conducting the pretest, sample design, master sample selection file preparation and training. This chapter describes the mechanism of quality assurance and standard guidelines that should be followed in these various areas. This chapter specifically addresses:

- 3.1 Questionnaire Design and Review Process
 - 3.1.1 Guidelines for Questionnaire Adaptation
- 3.2 Programming and Version Control
 - 3.2.1 Processes for Programming
 - 3.2.2 Checking the Questionnaire
 - 3.2.3 Handheld Setup
- 3.3 Pretest
- 3.4 Questionnaire Finalization
- 3.5 Data Transfer and Management
- 3.6 Sample Design and Review Process
- 3.7 Master Sample Selection File Preparation
- 3.8 Training Workshops

3.1 Questionnaire Design and Review Process

GATS maintains a standardized process for participating countries to design their country-specific questionnaires. The GATS Questionnaire Review Committee (QRC)—a group comprised of experts in tobacco control and questionnaire design from developed and developing countries — reviews and approves all GATS questionnaires. The QRC works closely with each country to adapt the GATS questionnaire for each country’s situation while maintaining the standard GATS core questions (refer to the *GATS Core Questionnaire with Optional Questions* for details) to ensure comparability across countries. Specific details of questionnaire adaptation and the review process can be found in *Appendix B.1*.

The GATS Questionnaire Review Committee (QRC) reviews and approves the questionnaire to ensure quality and comparability across countries.

3.1.1 Guidelines for Questionnaire Adaptation

The QRC recommends countries follow specific guidelines in adaptation of the *GATS Core Questionnaire* in order to facilitate the review process and ensure quality standards:

- Highlight all country adaptations to the *GATS Core Questionnaire* (for ease of reference). This includes question item lists, response categories, optional questions, and country added questions.

- Use strikethrough to indicate any core questions that the country wants to delete (for ease of reference).
- Maintain core and optional question numbering and ordering to maintain consistency for country comparison. (This is not always possible or optimal but should be fulfilled as much as possible.)
- For newly added country-specific questions, use double lettering depending on the section (e.g., AA10, BB17, EE4). This will not disrupt current numbering by adding in new country questions.
- Change skip instructions only if needed to accommodate either added optional questions or country added questions.
- Do not revise standard core questions (except for country specific categories) in order to maintain consistency for cross country comparison. (Justification should be provided to the QRC for any exceptions.)
- Limit the number of additional questions in order to keep the questionnaire at a reasonable length.
- Limit the complexity of additional questions, for ease of programming each country questionnaire.
- Provide justification for adaptations. Including justification helps the QRC in the process for reviewing and approving the country questionnaires.

3.2 Programming and Version Control

GATS uses a computer-assisted personal interviewing (CAPI) mode of data collection using handheld devices as the electronic instruments. The handheld devices are equipped with General Survey System (GSS) software¹. There are always unique features to each country as each country customizes their questionnaires, has different languages with special fonts, and may be using different hardware (e.g., different models of handhelds may be used). Given the complexities, there are a number of common steps and best practices that should be followed.

Standard processes and checks should be followed to ensure the programming of the GATS questionnaires is done in a high quality and efficient manner.

There is a sequence of steps (provided in *Appendix B.2*) that require cooperative and energetic input from the host country well in advance (preferably 6-8 weeks before the start of training for a pretest). These steps require that host country staff, such as IT staff/programmers, survey staff, and survey managers, be identified and available for work either during the orientation workshop or soon after.

These staff will be trained either in person or via webinars to get the preparatory work started and are expected to devote a significant number of work hours to getting ready for the GATS pretest and full survey implementation.

The subsections below provide specific issues with the questionnaire and survey preparation process that should be emphasized to increase quality.

¹ The General Survey System (GSS) is designed to run on an Android platform and has been tested and implemented using Android-based tablets.

3.2.1 Processes for Programming

The following processes should be followed in order to maintain quality and efficiency for GATS preparations:

- Survey/Field staff/survey implementation focal points, not IT staff, are responsible for verifying the accuracy of the country specific language (translations) based on their country questionnaire. The IT staff will work with the survey/field staff/survey implementation focal points for inserting the text using the GSS questionnaire design software.
- The implementing agency focal point must approve the *Household Questionnaire* (HQ) and *Individual Questionnaire* (IQ) (in writing) before field staff training (with version numbers).
- Text changes: A survey questionnaire wording change must be updated in all languages that will be administered.
- Changing the programming logic (ranges, validity checks, etc.) of an approved questionnaire requires re-review by the QRC.
- Changing the content or wording of a question or responses requires re-review by the QRC.
- Version control of DB3 files (survey questionnaire database files): The host country owns the DB3 files and country IT /Survey staff are responsible for maintaining audit trail of changes to questionnaires and GSS programming specifications.
- After pretest debriefing and full implementation starts, the country should send the final CMSDB, Survey0 and Survey1 DB3 files.

3.2.2 Checking the Questionnaire Program

There are several quality assurance steps for checking the questionnaire program that should be followed and documented:

- Programmers must update the questionnaire program version every time a revision is made. Also, this must be confirmed before starting to review the questionnaire program. An archive of older DB3 files for history and backup is kept by the GSS software.
- For every language, the full GSS programming specifications is reviewed by viewing each question on the handheld and each question checked to match the question text, responses, and special characters exactly with that on the *GATS Questionnaire Programming Specifications* document. If any problems are found, they are documented clearly on the specifications, after one iteration the problems are fixed, and then the complete checking iteration is performed. This iterative process continues until every question is approved on the handheld. There is absolutely no wavering from this process. This must be done with a local language survey expert. It should not be relegated to programmers and not to staff without the requisite language expertise. The country focal point or representative should sign off on the final version after reviewing the changes/updates made. The specs should be reviewed before signing off to check that each and every question has been checked by going back and forth on the questionnaire (refer to the *GATS Data Management Training Guide, Module 2.1* for details)
- Skip patterns and English text are always well tested ahead of the site visit. The questionnaire software should not be modified during field data collection unless it is absolutely critical.

- For every language, the count of opening and closing HTML tags, (,), [,], { , and } symbols must be the same. Otherwise, this needs to be re-checked and fixed.

3.2.3 Handheld Setup

There are formal quality assurance procedures for loading the handhelds with the GSS program that should be followed:

- Before loading, all handhelds should have completed initial setup and if applicable, previous versions of GATS software should have been uninstalled and relevant folders deleted.
- Labels are printed with quality control step checkboxes (for installing GATS on the handhelds) that are affixed to the handhelds. The person completing the quality control on the handheld load must confirm that every step has been completed.

Appendix B.3 provides further details about the step by step process for loading the handhelds (refer to *GATS Data Management Training Guide, Module 2.2* for details).

3.3 Pretest

There are six main objectives for conducting the in-country pretest, prior to full survey implementation. Achieving the following objectives will help to ensure that the GATS full survey data collection is carried out in a high-quality manner:

1. Provide in-country staff with assistance to program, test, and maintain a handheld-based survey, as needed.
2. Review the GATS pretest questionnaire including question and response wording, probing instructions, skip logic programming, and length and time of interview.
3. Provide in-country staff with assistance in the use of handheld hardware and software, as needed.
4. Observe pretest implementation in one rural and one urban environment, if feasible and as applicable.
5. Test the data collection, transfer and management plan during the pretest that will be adapted for full survey implementation.
6. Debrief GATS partners on strengths, weaknesses, and recommendations on pretest preparations and implementation.

The GATS pretest provides an opportunity to test the questionnaire and data collection and management procedures in order to ensure the full survey will be carried out in a high quality manner.

3.4 Questionnaire Finalization

In order to finalize the GATS questionnaire for full survey implementation, the following steps should be implemented:

- **Pretest debriefing.** All staff (Field Interviewers, Field Supervisors, project staff) should provide their input on how the questionnaire worked as a whole, as well as specific survey questions to target for revisions.

- **Review pretest data.** Analyzing pretest data can provide valuable input for making post-pretest revisions. Some specific issues to explore:
 - Item nonresponse: Identify questions with high rates of “Don’t know” or “Refused” responses.
 - Frequencies: Examine frequency distributions of key questions to determine if data look appropriate. Something out of the ordinary could signal a questionnaire design issue or some other problem (e.g., training issue).
 - Allowable ranges: Review appropriateness of the allowable ranges programmed into the handhelds and adjust for full survey implementation as needed.
 - Response categories: Determine if country-specific questionnaire adaptations (e.g., modifying question item lists and response categories) were appropriate. For example, measure the frequencies of cigarette brands in order to modify the list of brands for the full survey questionnaire.
- **Revise for full survey implementation.** The revised post-pretest version should be submitted for QRC review, including justification (based on results from pretest) for revisions. The same quality assurance guidelines outlined in *Sections 4.1* and *4.2* (as applicable) should be followed in order to prepare the questionnaire for full survey implementation (including revised programming, etc.).

3.5 Data Transfer and Management

The pretest is a complete dry-run of the full survey implementation for the data management process. Based on pretest experience and recommendations, the data transfer and management process will be finalized for full survey implementation. Therefore, each stage of the data management process should be conducted as per the data management plan considering the full survey. For example:

- The feedback mechanism of sharing reports should be practiced during the pretest.
- The data transmission method—such as WiFi-based, SIM-based, cloud-based, or FTP site—should be tested during the pretest and necessary adjustments, if any, should be made for the full survey based on lessons learned from the pretest.

3.6 Sample Design and Review Process

GATS maintains a standardized process for participating countries to develop their sample design. The GATS Sampling Review Committee (SRC)—a group comprised of experts in sampling methodology from developed and developing countries—reviews and finalizes the sample designs for participating countries. The SRC and CDC focal points work closely with each country to adapt the GATS sample design while maintaining GATS standards to ensure comparability for cross-country analysis. Specific details of the GATS sample design can be found in the *GATS Sample Design Manual*.

As part of the GATS sampling process, mapping and listing may need to be conducted in order to create an appropriate area probability sampling frame. For complete procedures and recommendations on quality assurance in mapping and listing, refer to the *GATS Mapping and Listing Manual*.

3.7 Master Sample Selection File Preparation

Once the GATS sample has been drawn, a Master Sample Selection file needs to be prepared. The Master Sample Selection file is a dataset that contains a CaseID for every household in the sampling list along with the information necessary to compute the sample weights and analyze the complex survey data.

The Master Sample Selection file includes sample identifiers and geographical identifiers. Sample identifiers such as stratum, primary sampling unit (PSU), secondary sampling unit (SSU/segment), type of place of residence (urban/rural), and the pre-selected gender assignment of the household, if gender randomization is used. Geographical identification such as address information including codes for country, region, province/state, county/district, and city/town/village.

The Master Sample Selection file includes a unique CaseID, sample identifiers, and geographic identifiers of the household.

Once the Master Sample Selection file is finalized, it should be organized and converted to text file using Excel and/or MS Access. This text file will be used to load cases onto the handheld devices for data collection, as applicable. Also, using the GSS software, the Case Grid and Case Info screens should be finalized to reflect the revised case file. The following chapter provides further details about this process.

3.8 Training Workshops

Training is an important quality assurance mechanism for GATS implementation. The training workshops provided for countries includes the GATS orientation, pretest (discussed in previous sections), and full survey training. The trainings are part of the GATS protocol and standards to ensure high-quality data are collected for each country.

GATS trainings are provided at all levels of the country survey team involved in the implementation of the survey, from interviewers to supervisors, IT personnel and the central team overseeing the whole process from initiation to completion. Country and international partners facilitate GATS training at all the levels. Training of all personnel ensures that everyone involved in the survey process are clear about their role in ensuring good quality data. Training is an on-going process conducted prior to survey implementation and during survey implementation with thorough debriefing at the end of each training.

The goals of GATS training are to:

- Ensure a standardized application of the survey protocol, standards, and materials.
- Ensure a clear understanding of the rationale of the survey and the survey protocol.
- Ensure standard application of GATS electronic data collection, management, transfer, and aggregation system.
- Engage and motivate interviewers.
- Provide practical suggestions.
- Improve the overall quality of the data produced.

3.8.1 GATS Orientation Workshop

The orientation workshop is the first quality assurance step in the GATS process. This involves an in-depth review of the standard GATS protocol including questionnaire, sample design, electronic data collection and management, and dissemination. Countries share their experiences in conducting surveys reflecting their technical capacity to implement GATS in the respective country but also to determine necessary technical assistance from partners to the agency in order to achieve the desired outcomes. Orientation workshops review guidelines for proposal development and allows for countries to adapt the protocol to their country situation. Finally, the orientation provides a step-by-step guide on proposal development, submission, review, and funding critical to initiation and implementation of the survey project.

3.8.2 Pretest

The objectives for conducting a pretest have been detailed in *Section 3.3*. The pretest training includes multiple components including training of IT staff, field interviewer/supervisor training, field practice, and debriefing. The pretest allows project staff to gain experience in training field staff. During pretest, IT and field staff are trained in software and hardware needed to support GATS implementation. Training ensures that staff are adequately prepared and understand the software and hardware needed to support GATS implementation.

3.8.3 Full Survey Training

Full survey training for interviewers and supervisors is a key quality assurance mechanism for GATS. The goal is to train interviewers and supervisors to produce useful and reliable data from the field survey. The training covers the rationale for the survey rooted in the background on tobacco control, and global tobacco surveillance system, interviewing techniques, identification of sampled households, consents, questionnaire content, and practical components which, includes role-playing, and observation in the field. Trainers review the paper questionnaires and the questionnaire on electronic handheld tablets with participants, in order to enhance their understanding of the survey. In addition, through role plays and field practice, trainees get familiar with the questionnaire and the nature of anticipated problems during an interview with difficult respondents and how to handle them. Training materials are also provided to the trainees as reference materials during the training and during field work. Materials include questionnaire, question-by-question specifications, and interviewers manual.

4. Quality Assurance: Data Collection and Management

This chapter describes data quality processes and procedures for GATS data collection and management. Data collection and management encompasses GATS GSS software setup, use of the software during field data collection, transfer of the data collected, aggregation and creation of a master aggregated database file. It also includes monitoring, reporting and technical support during field data collection.

For any given country, data collection, aggregation, and quality assurance should be performed at two levels: 1) in the field, 2) at the country's national data center (NDC). This chapter describes the data quality processes and procedures applicable to each of the two levels, specifically addressing:

4.1 Field Data Collection: Setup and Maintenance

4.1.1 Case File and Case Management

4.1.2 Handheld Electronic Data Collection Devices and Survey Questionnaire Software

4.1.3 Field Data Collection Quality Assurance Procedures

4.2 National Data Center Data Management

4.2.1 Aggregation

4.2.2 Quality Control of Collected Data

4.2.3 Communication and Technical Support

4.1 Field Data Collection: Setup and Maintenance

GATS standard protocol recommends certain systematic procedures be established and followed by IT and field staff to maintain accurate and complete data throughout the period of data collection. These procedures are described below.

4.1.1 Case file and Case Management

1. It is recommended that the case file be prepared and maintained at the NDC using the GSS software, as applicable, based on the country's data management implementation plan. (Note that the case file is a subset of the Master Sample Selection File.)
2. During case file creation and prior to handheld setup:
 - All Field Interviewers (FIs) should be identified and assigned a handheld. CaseID assignments should be unique and specific to each FI using the FIID variable in the case file. At the time of setup, the same CaseID should not be assigned to more than one FIID. This will ensure duplicates do not occur, the FI case assignment load is reasonable, and cases are readily identifiable.
 - All FIIDs should follow the GATS convention of an assigned 6-digit numeric ID.
 - All CaseIDs should follow the GATS convention of an assigned 6-digit numeric ID.

3. The case file should include an assignment for every handheld/FI that will collect data.
 - Ensure all FIs listed in the FIID spreadsheet have an assignment in the case file. The file should be reviewed prior to setting up the handheld devices. When loaded on the handhelds, the expected number of cases should appear in the CMS. If not, check the FIID assigned in the handheld software and check the FIID in the case file. It should be noted that for each CaseID, there will be two records created that are loaded into the CMS (HQ=00 and an IQ=01).
4. Each handheld case grid should contain only the cases the FI is assigned and intended to work. Typically, no more than 300 CaseIDs should be assigned per FI and loaded into the handheld. Performance, such as the application load time and the ability to easily identify a case during field work, could be affected if more than the recommended number of cases is loaded.
5. Prior to handheld case assignment, perform a frequency review of the case file fields while still on the NDC computer. Refer to the *GATS Programmer's Guide to General Survey System CMS* case file document for field requirements:
 - The CASEID field should be checked for duplicates and should follow the GATS standard of 6-digit numeric ID.
 - Check that the PROJECTNAME field indicates the country name to ensure a unique country identifier field exists in the case file (the update projectname query can be used in the MS Access case file template).
 - Check that the CreateDate variable value is the date the file was created.
 - Compare the Master Sample Selection file with the handheld case file to ensure completeness and correctness. (Note that the handheld case file is a subset of the Master Sample Selection file.)

The *GATS Programmer's Guide to General Survey System* contains detailed information regarding the case file creation, field descriptions and the handheld FI inventory spreadsheet.

4.1.2 Handheld Electronic Data Collection Devices and Survey Questionnaire Software

1. **Field Interviewer Settings:** Each handheld should be labeled with a unique FIID following the GATS convention and with the same FIID set and displaying within the CMS Settings menu. This information should be logged at the NDC in a spreadsheet along with the FI's name and handheld serial number.
 - Cross reference the FIID settings: check that the spreadsheet assignment, FIID field in the Settings menu and handheld label all match prior to data collection.
2. **Software Versions: CMS, Household, Individual Questionnaires:** Handhelds should be loaded with the final approved versions of the questionnaires.
 - For each handheld, verify the version for each questionnaire is loaded and functional by checking the Information option at the top of the screen and verifying the versions in the Survey Information section are correct.
 - When possible, all handhelds should be setup at the same time to ensure consistency in version, configuration, and QC process.

3. **Loading Cases:** At the start of data collection/field work, only cases listed in the case file with a matching FIID should appear in the CMS of each handheld. Training cases should be erased before field work begins to avoid the possibility of working a training case and having these cases exported. Analysts should be prepared to filter out training cases from live data for situations where FIs break this rule.

The *GATS Programmer's Guide to General Survey System* contains detailed instructions regarding handheld settings and creation and installation of the handheld related files.

4.1.3 Field Data Collection Quality Assurance Procedures

This section provides an overview of recommended quality assurance procedures that may be implemented during field data collection. Both the *GATS Field Interviewer Manual* and the *GATS Field Supervisor Manual* should be referred to for more detailed information.

Role of the Field Supervisor

Field Supervisors (FSs) are the critical link between FIs and the GATS survey management staff. They are expected to monitor production and performance, and communicate any field issues that may have an impact on the timely completion of GATS. They are entrusted with ensuring that their team's data are collected according to the GATS data collection protocol and are of the highest quality. To help FIs understand the value that the GATS team places on data quality, the FS must make quality control an integral part of their weekly activities.

The first step of the FS in ensuring data quality is communicating quality expectations to their FIs. Supervisors are expected to provide continuous feedback to FIs concerning quality issues, both positive and negative, and will emphasize the importance the GATS team places on *quality*.

Monitoring Production

Production includes all activities required to successfully achieve the GATS response rate goals. These goals include initiating activities for each assigned household; carrying out contacting, locating, and refusal conversion efforts; and completing these activities successfully, in compliance with all GATS specifications. The FS and country survey staff should work together to monitor the time taken to conduct an interview on a regular basis. In addition, monitoring production during data collection is essential to ensuring GATS is completed on schedule and within the allotted budget.

The following types of reports can be generated using the GSS software, and should be used by the country survey staff or IT staff to monitor various aspects of production and ideally should be shared with the FSs on a regular basis. When and where these reports are available will depend on the data management model in place for each country.

- **HQ Case Status and IQ Case Status:** These reports list the most recent result codes for the household and individual cases by interviewer. This report is viewed at the national level and can be shared with the FSs and/or FIs.
- **Exceptions Report:** This report would assist the GATS staff to monitor the number of revisits for any cases that are not complete and help understand the reasons for revisits.

- *IQ Frequency Distributions Report*: This report would help the GATS survey staff and IT staff to monitor questionnaire responses by interviewer or PSU, in order to assess the quality of the survey interviews by looking at measures of quality such as:
 - Potential problems with interview routing or “missing” data.
 - Higher than expected rates of “don’t know” or “refused” responses.
 - Higher than expected rates of “other” responses, rather than one of the pre-coded answer choices.
 - “Other, specify” responses that are unclear and cannot be easily coded or that are duplicates of one of the pre-coded options.

Monitoring Data Transfer

Regular transmission of data from FIs to the NDC is vital for the management of GATS and for early detection and resolution of potential problems. FSs are responsible for ensuring that each FI transmits data on a regular schedule (typically, daily). FSs can work with country IT staff to determine if there are any issues with transmissions.

Observation

Observing FIs is an important part of the FS’s quality control efforts and can serve to help FIs become better at their job. Observation of FIs’ work will also help avert any temptation on their part to take shortcuts in administering the questionnaire or otherwise not follow the GATS interviewing protocol. At a minimum, FSs are expected to observe each of their FIs during the first few days of the field period and then more infrequently, based on the implementing agencies’ guidelines. On these observations, the FS will accompany the FI to sampled households to verify that the FI is recording the visit outcome correctly in the GSS Case Management System and conducting the household screening and subsequent interview properly. These early visits will go a long way toward making sure that errors in questionnaire administration or use of the result codes are identified quickly and that the FI can be given additional training or guidance as needed.

In addition, FSs may decide to accompany FIs on particularly difficult visits (for example, households that have refused previously).

Verification

One way to check the quality of the data collected by interviewing staff is to conduct short verification interviews with households that have already been screened and interviewed. Conducting a short verification interview will allow FSs to confirm that the FI has done the following:

- Identified and screened the correct household. Occasionally, a FI may select a household different from the one that was sampled. If this happens, the FS must direct the FI to go to the correct household and conduct the screening and interviewing with that household.
- Recorded the correct age, gender, and smoking status of household members. If the age is not properly recorded (for example, the FI includes residents younger than age 15 on the roster in addition to residents aged 15 or older), this error will impact the identification of eligible individuals and could cause the FI to select an incorrect household member to interview.
- Administered the *Individual Questionnaire* to the selected household member.

The exact number of verification interviews will be determined in conjunction with the implementing agency management staff; however, it is generally expected that FSs will conduct randomly selected verification interviews for about 10% of each FI's assignment. Verification interviews may consist of either (1) a short verification that asks the selected respondent a small number of questions to verify that the respondent recently completed a survey on smoking-related topics and to assess the performance of the interview or (2) if a country desires, a full re-interview of the *Household Questionnaire* and, if feasible, the *Individual Questionnaire*. Recognizing that respondents may perceive the re-interview process as a burden, it is not necessary to administer the entire *Individual Questionnaire*. Furthermore, verification interviews may be conducted via paper-and-pencil administration or via a handheld loaded with the households to be verified. The FI's handheld should never be used to conduct the re-interviews, rather a separate handheld should be used.

Later, the FS will compare the responses obtained from the re-interview with those recorded by the FI. If differences exist and it appears that the FI conducted an interview with the incorrect person in the household, the FS will send the FI back to the household to interview the correct individual, and the responses should subsequently be compared again. In all cases, whether an error was discovered in the verification process or not, the information from the verification interview should be transmitted to the NDC along with all other materials.

Assigning Non-interview Final Result Codes

FSs should approve the final non-interview result code before a questionnaire is finalized. (The *GATS Field Supervisor Manual* provides further information for attempting to complete pending non-interview cases.) If no further action should be taken on a case, the appropriate final non-interview result code should be entered into the Case Management System by the FI (see *GATS Field Interviewer Manual* for list and description of field result codes).

As data collection is completed in a certain area (e.g., PSU, region, state, etc.), all cases that were worked should have a final result code assigned to them. The completed *Household* and *Individual Questionnaires* will automatically be assigned a final completed result code by the program. For *Household* and *Individual Questionnaire* cases that did not result in a completed interview (e.g., refusals, nobody home), non-interview final result codes should be assigned by the FIs with approval from their supervisors.

As data collection is completed, all cases that were worked should have a final result code assigned to them.

4.2 National Data Center Data Management

This section describes data collection and management procedures at the NDC level: combining/aggregating the files, quality control aspects including reporting and monitoring, communication and technical support, and creation of raw data files.

4.2.1 Aggregation

As data is received at the data aggregation point (cloud server, FTP server, etc.), files should be reviewed, processed and quality control activities performed. Some documentation and archiving of data files should take place.

During processing, the aggregation software sorts and processes the files in the selected folder (typically C:\GATSAndroid2\IncommingDB3s\...) in alphabetical order regardless of any date. It is critical for files to be aggregated in the order exported. When aggregating all exported files, the exported filename convention of FI#_YYYY_MM_DD.db3 provides for automatic alphabetic sorting and correctly processes the data. Export files are cumulative and hence the most recent file contains all the data, including updates or edits that were in earlier files.

The aggregation processing does not allow for duplicate CaseIDs.

Aggregation Processing

- As a starting point at the beginning of data collection, a master aggregation file exists in the C:\GATSAndroid2\DATA folder named master.db3zero. This is required at the start of data collection so that aggregation starts with a master file not containing any data. The master.db3zero should be copied, and the subsequent file should be renamed to master.db3 when re-initialization is necessary.
- A Node.id file must reside in the C:\GATSAndroid2\ folder.
 - Before aggregation is run the first time on the machine, the node file should be edited to contain a unique and meaningful name.
 - Once aggregation has run, the "Node.ID" file should not be modified.
- For countries implementing data encryption, the latest db3 files collected should be decrypted using the Data Aggregation > Database Decryptor option in the GSS software. This must be performed before aggregation if data encryption is being used.
- Aggregate the latest db3 collected files to those previously aggregated for efficiency as long as files are received and processed in order (by date exported). While it is acceptable to aggregate to an empty master file each time aggregation is run, it should be noted this will take longer as all exported db3 files would need to be processed.
- Save all exported files received by archiving them into a subfolder within the C:\GATSAndroid2\ folder. They do not need to be aggregated again.

During and after aggregation, several points should be checked:

- All files received contain filenames that included the date the file was exported thus ensuring processing is not performed with files out of order. If files are received out of order, (i.e., earlier data files are received subsequently), the latest data for a particular CaseID will be overwritten with earlier perhaps incomplete data. To ensure when aggregating to a non-empty master.db3 that you have processed the latest data, refer to the db3 files archived. (Always aggregate the files in date order or just use the most recent files.)
- Correctness and completeness of data aggregation: were all the expected db3 files accounted for and included in the aggregation run?
- Was the most recent db3 file for each FI aggregated?
- Data review: Checking and validation should be performed. The analyst should review the event data as well as the completed questionnaires to identify potential problems, patterns, or

inconsistent data. It will also generate frequencies, by question, for review. Information about viewing DB3 files can be found in the *GATS Data Management Implementation Plan*.

Review the aggregation summary report and the aggregation status report generated from aggregation to confirm counts.

The Data Aggregation Summary Report provides information about the results of the aggregation. Each row shows the results for an aggregation unit (typically a FI) and a specific database table. The table is denoted in the AggTable column, the number of rows in the input table is shown in the AggRowsTotal column, and the number of rows inserted from that table is shown in the last column AggRowsInserted.

The Household Screening Status report is a summary of the status of all the cases in the aggregated data file. For each aggregation unit (typically a FI) the total number of cases is reported for that unit in the Cases column. Then the total is broken out into the various status categories. The status columns (Completed Ints, Unworked, Pending, and Final) depend on the most recent result code for the case.

- Using the aggregated data or any db3 file, flat file formatted data or CSV files can be created using Data Aggregation utilities in the Tool Kit.
- Review of the Event data (contained in the DU and DuEvt files) for content and understanding of field operations is recommended.
- The DU file contains only one record for each CaseID (HQ and IQ have distinct CaseIDs -00 and -01 respectively) and the most recent result code for that CaseID. The DuEvt file contains the history of the events, for a given CaseID, entered by the FI(s) and may contain more than one record per CaseID.

Generating Data Files for Review

- The questionnaire data (Responses table) are stored in one table for both the HQ and IQ forms. There is one row per question (QID) so these data need to be restructured prior to analysis. The responses data should undergo transposition so it can be easily reviewed and readily used in statistical software. This will create a CaseID based record for both the HQ and IQ data. The GSS software transposes the responses data based on the Survey db3 files specified with separate CaseID level files, flat files, for HQ and IQ.
- GATS events data are stored in the DUEVT table and are stored with one row for each result for each CaseID. GATS DU (Dwelling unit) data are stored in the DU table and are stored as one row per CaseID. The DU data also contains a copy of the most recent result code for each case.

The table below summarizes these data sources from the handheld system.

Table Data Source	Table	Rows
Questionnaire data	Responses	One row per question for every CaseID and form
Event data	DUEVT	One row for per result for every CaseID and form
Dwelling Unit	DU	One row per dwelling unit
Address edits or changes	Address Log	One row per address edit
FI Notes	Notes	One row per note type (either at the CaseID level case or per question (QID))

4.2.2 Quality Control of Collected Data

At the NDC, quality control procedures should be in place to routinely review, monitor, report, and backup the data as they come in and are processed. The NDC should establish a daily processing schedule at minimum for reporting and monitoring and one that also includes a routine backup process.

The GATS standard protocol consists of a list of GATS indicator variables and suggested reporting at the country level. The key indicator variables with details and information can be found in the *GATS Indicator Definitions*. Countries may add/adapt to the indicator variables as applicable to the final approved questionnaire.

It is recommended that key indicator variables be reviewed at the raw data level during data collection.

Country Developed Reports: It is recommended the country develop reporting for case level event status. Survey monitoring (e.g., field progress, data collection, and transmission) is a key coordinated, ongoing activity.

- During the data collection period, the NDC should provide a weekly response rate report generated from the GSS software. The information should indicate the completed interviews and overall response rate. Examples of fieldwork and response rate status reports can be found in the *GATS Programmer's Guide to General Survey System*.
- The NDC is encouraged to use the reports generated from the GSS software and as required to write reports in Access^{1,2}, SPSS^{1,3}, SAS^{1,4}, STATA^{1,5} or other software that will aid in the review of the data for inconsistencies, anomalies and incompleteness. The transposed aggregated data should be used for generating reports.

Inconsistencies in the data may be identified by performing validation checks across variables within and across the HQ and the IQ. Please refer to *Chapter 5* on post-data collection quality assurance for information on validation checks.

¹ Use of trade names is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services or the Centers for Disease Control and Prevention.

² Microsoft Office Access© (Microsoft Corporation, Redmond, Washington).

³ SPSS© (SPSS Inc., Chicago, Illinois).

⁴ SAS© (SAS Institute Inc., Cary, North Carolina).

⁵ STATA© (Stata Corp., College Station, Texas).

Backup procedures: Data should be backed up on a routine basis, such as weekly depending on the frequency of aggregation.

- Individual db3 files (FI level) should be archived.
- Data archived should also be backed up to a secondary point such as a network or flash drive and a copy should be maintained at two separate physical locations.
- After data collection has ended, reconciliation of data received should take place with respect to the master sample selection file as well as exported files from the field.

4.2.3 Communication and Technical Support

Communication from the NDC during data collection should entail weekly status reports being sent to the WHO and CDC and the reporting of any and/or all technical issues. All handheld and software related technical issues whether resolved or unresolved should be reported by email as soon as possible from when they occur. This will help facilitate timely technical support, issue resolution and information sharing.

5. Quality Assurance: Post-Data Collection

The post-data collection phase includes a series of tasks that need to be completed in order to prepare an analytic data file for conducting data analysis and refers to a stage once all the survey data have been collected and aggregated. This encompasses preparing the data for sample weight calculations; assessing the quality of the sampling, sampling error, and weights; and measuring the quality of frame coverage, nonresponse, and other non-sampling errors. This also includes preparation of an analytic data file. This chapter describes quality assurance guidelines and procedures applicable and recommended to each of these activities, specifically addressing:

- 5.1 Cleaning and Preparing Data for Sample Weight Calculations
 - 5.1.1 Creating the Master Database File
 - 5.1.2 Removing Confidential Identifier Variables
 - 5.1.3 Cleaning and Validating the Merged Data File
 - 5.1.4 Assigning Final Disposition Codes
- 5.2 Measures of Quality: Sampling, Sampling Error, and Weights
 - 5.2.1 Pattern of Post-stratification Weights Calibration Adjustments Among Adjustment Cells
 - 5.2.2 Multiplicative Effect of Variable Sample Weights on the Precision of Survey Estimates
 - 5.2.3 Overall Design Effect on the Precision of Survey Estimates and the Within-PSU Intracluster Homogeneity of Corresponding Key Survey Estimates
 - 5.2.4 Margin of Error for Key Survey Estimates
- 5.3 Measures of Quality: Coverage, Nonresponse, and Other Non-Sampling Errors
 - 5.3.1 Household Frame Coverage Rate
 - 5.3.2 Patterns of Respondent Cutoff Rates
 - 5.3.3 Patterns of Household Response Rates by First Stage Sampling Strata
 - 5.3.4 Patterns of Person-level Response Rates Among Variables Used for Nonresponse Adjustments
 - 5.3.5 Patterns of Person-level Refusal Rates Among Variables Used for Nonresponse Adjustments
 - 5.3.6 Item Nonresponse Rates for Fact Sheet Indicator Variables
- 5.4 Formal Review of Statistical Quality
 - 5.4.1 Sample Weights Documentation
- 5.5 Creation of Analytic Data File
- 5.6 Public Use Data File Creation

5.1 Cleaning and Preparing Data for Sample Weight Calculations

This section provides guidelines for the merging of data files, the validation of variables and skip patterns, as well as the creation of the disposition codes using a statistical software package¹.

¹ The GATS Data Coordinating Center (DCC) at CDC provides technical assistance for use of the following statistical software packages: SAS, SPSS, and STATA.

5.1.1 Creating the Master Database File

The Generate Analysis Files option combines survey data into a single, comma delimited file and creates SAS, SPSS, and STATA input programs capable of reading the newly merged master database file. This can be created using the GSS software.

The files used in the Generate Analysis Files utility include the final aggregated Master db3 file created by aggregating the db3 files from all interviewers, and the two questionnaire databases (Survey0.db3 and Survey1.db3).

Ensure using the updated file locations of the aggregated file, and the survey questionnaire databases when using the Generate Analysis Files utility. To change, update the file location or name in the text box or select the “...” button and the system will allow you to select an existing file. The master database file will be a comma delimited file in Unicode (UTF-8) format. The utility will also generate input code for reading the data in SAS, SPSS, and STATA.

To view the CSV file, a Unicode-enabled text editor such as WordPad is required. The SAS, SPSS, and STATA files are also in Unicode to ensure that labels and formats are displayed correctly. The SAS, SPSS, and STATA programs may need to be updated prior to running them to ensure that country specific options are set. The process would create a master dataset. This master database file created should be used to remove any confidential identifier variables and should be shared with the country's IT focal point.

5.1.2 Removing Confidential Identifier Variables

From the master dataset created, remove the confidential identifier fields from Master Sample Selection fields section.

5.1.3 Cleaning and Validating the Merged Data File

Verify that variables have appropriate values, verify skip patterns worked correctly, and check on blank fields. While many of these data quality checks are implemented during interview administration by the handheld software, it is important to check again in case there were undetected errors in the handheld software programming.

1. Check that the data skip patterns have worked as specified in the questionnaire. The original data skip patterns derived from the final approved country questionnaire should be verified for all core variables and it is recommended to check the patterns for all other variables as well.
 - A respondent who did not answer an item due to a skip pattern would have a blank field for that skipped item.
 - For example, if a respondent refused to answer either question A02a (the month the respondent was born) or A02b (the year the respondent was born) then the respondent should respond to question A03 (the age of the respondent) but if both responses are provided for A02a and A02b then A03 should be skipped.
 - For example, in question B01 if a respondent answers that he/she smokes on a 'daily' basis (a value of 1) then he/she would then proceed onto question B04 and should not have a response for questions B02 and B03.

- In SAS, a blank field for a numeric variable is equal to '.' (for character variables it is ' '). In SAS, this skip pattern could be verified using the code below. Any records that do not follow the skip pattern would be output to a SAS dataset.

```
IF B01 = 1 & (B02 NE . OR B03 NE .) THEN FLAG=1;
ELSE IF B01 = 2 & (B02 = . OR B03 = .) THEN FLAG =1;
ELSE IF B01 = 3 & (B02 NE . OR B03 = .) THEN FLAG=1;
ELSE FLAG=0;
IF FLAG=1 THEN OUTPUT; *A value of 1 for FLAG indicates the skip pattern did not work
and a value of 0 indicates the pattern did work;
```
2. Check that the only blank fields are those that result from a skip pattern (e.g., they are not applicable for the respondent and should be blank). Any other blank values are invalid.
- The fields that are blank due to the skip patterns will need to be converted to a new value. This will differentiate them from blank fields that are not attributable to skip patterns (invalid blank fields). Since original variable values should not be written over, create a new dataset to complete this step.
 - In SAS use the code '.S' for a missing value when the question was not applicable (validly skipped). The SAS code for the skip pattern used above could be modified to read:

```
IF B01 = 1 & (B02 = . OR B03 = .) THEN B02 = .S & B03 = .S;
ELSE IF B01 = 3 & B02 = . THEN B02 = .S;
```
 - After all skip pattern values have been replaced there should be no other blank fields. If other blank fields do exist, then output those records to the error file.
3. Check each variable to ensure that invalid values are not present. Use the country's data dictionary (codebook) to see the valid values for each variable in the dataset.
- Check that the values for the responses to each question are valid.
 - For example: The only valid responses for B01 are: 1 ('daily'), 2 ('less than daily'), 3 ('not at all'), -9 (refused), or the missing value code specific to the software utilized (e.g., .S when using SAS).
 - For example: Some variables have ranges associated with them so that a value less than or greater than the range should not exist. So, for variable B04, the lowest value that is possible should be 0 and the highest value possible should be 98. There is also an option for a value of -9 if the response was "Don't know" or "Refused". The missing value code specific to the software utilized (e.g., .S when using SAS) would also be a valid value.
 - *In SAS, records with invalid responses for B01 could be selected and output to the error file using the following code:*

```
IF B01 NOT IN(1,2,3,-7,-9) THEN OUTPUT;
```
 - *If the analyst prefers, simple frequencies could instead be run on each of the variables to see if values are out of range. Any records with invalid values could then be identified and output to the error file.*

- Verify the respondent's age
 - Before verifying that the respondent is the correct age, the age must first be calculated and included on the dataset. In order to do the calculation, use the date of the IQ interview (IQ_EVENTDATE) and the month (A02a) and year (A02b) of the respondents' births, with the day assumed to be the 15th of the month. If respondents provide responses of 'Don't know' or 'Refused' to either A02a or A02b then the survey moves on to question A03 and they are asked to provide their age (A03 should only be answered, however, if either A02a or A02b are not reported). If the month or year of birth is not reported then use A03 as the respondent's age. If the month and year are reported then calculate the age by subtracting the birth date from the date of the interview. Verify that the age of the respondent is valid (the valid range is from 15 to 125).
 - *In SAS, the age could be calculated and then verified using the following code:*

```
IF A02A <=12 AND A02B NE -7 AND A02B NE -9 THEN DO;
  BIRTHDATE = MDY(A02A, 15, A02B);
  AGE = FLOOR(YRDIF(BIRTHDATE, IQ_EVENTDATE, 'ACTUAL'));
  END;
  ELSE AGE = A03;
```

4. Other data validation checks:

- Check to make sure that if any of the following variables were not skipped, their values are less than the respondent's age: B04, B05, B08, B09, B11, B12, B13, C04, C05, C08, C09, C11, C12, and C13 (depending on the metric of the variable, these variable values might need to be converted into years first).

5.1.4 Assigning Final Disposition Codes

The merged data file will include the final result codes that were assigned by FIs to all of the sample cases worked during data collection (see *GATS Field Interview Manual* for description of field result codes). Each household interview case and each individual interview case should have one final result code (see *Section 4.1.3* for assigning non-interview final result codes during data collection).

Final disposition codes should be assigned to each case. The final disposition codes are used for calculating and reporting response rates and quality assurance measures.

Using the final result codes, final disposition codes should be assigned. The final disposition codes are then used for calculating and reporting response rates and quality assurance measures. Refer to *Appendix C.1* in order to assign the final disposition codes from the final result codes as well as calculations for response rates and a template for reporting these results from data collection.

Below are important guidelines that should be followed for this process:

1. Each case should be assigned one final disposition code for the HQ based on the final HQ field result code. If the HQ interview was completed and one person was selected for the IQ interview (final disposition code 1), a final disposition code should be assigned for the IQ as well. If the final HQ disposition code is something other than 1 (e.g., no one selected, refusal, incomplete), then there will be no final IQ disposition code since the IQ interview was never worked.

2. There are key questions in the IQ that determine the legitimacy of the interview. If the recorded answer to any of these key questions is “Don’t Know” or “Refused,” the interview is not considered legitimate. Thus, for all IQ cases with a final result code of 400 (Completed *Individual Questionnaire*), the rules below should be followed in order to determine which disposition code should be assigned.
 - If B01, B02, or B03 (key smoking prevalence questions) were recorded as “Don’t Know” or “Refused,” a final disposition code of 16 (Selected Respondent Incompetent) should be assigned. This should not be a common occurrence. (This rule may also be applied for C01, C02, and C03 (key smokeless prevalence questions) only if smokeless tobacco use is common in the country.)
 - For all other IQ cases with a final result code of 400 (Completed *Individual Questionnaire*), a final disposition code of 11 (Completed *Individual Questionnaire*) should be assigned. This will occur for most of the cases.
3. For all IQ cases with a final result code of 402 (Completed Part of *Individual Questionnaire*), the rules below should be followed in order to determine which disposition code should be assigned.
 - If the IQ was completed through at least question E01 AND none of the key questions (described in point #2 above) were answered “Don’t Know” or “Refused,” a final disposition code of 11 (Completed *Individual Questionnaire*) should be assigned.
 - If the IQ was broken off before question E01, a final disposition code of 12 (Incomplete) should be assigned. (The cases with a final disposition code of 12 (Incomplete) are considered nonrespondents for GATS and the data are not included for analysis.)
 - If the IQ was completed through at least question E01 but any of the key questions (described in point #2 above) were answered “Don’t Know” or “Refused,” a final disposition code of 16 (Selected Respondent Incompetent) should be assigned. (As described in point #2 above, these respondents are determined to be incompetent if they cannot provide legitimate answers to the key questions and thus, their interview data are not included for analysis.)
4. Only cases with a disposition code of 11 should be included in the final analytic data set. (Thus, it is important to assign the disposition codes properly.)
5. Use cross tabs to check all final result codes against the disposition codes for misclassification. If the two codes do not match up as they should then this would indicate a problem with the software code used to create the disposition codes.

5.2 Measures of Quality: Sampling, Sampling Error, and Weights

This section describes calculations to directly assess the quality of estimates from GATS samples and to indicate the effects of sampling clusters and unequal weights on these estimates. Guidelines to check the accuracy of the calculated weights are also included.

5.2.1 Pattern of Post-Stratification Weights Calibration Adjustments Among Adjustment Cells

The last step in producing sample weights involves calibrating the weights to population counts by known correlates of key study outcome measures, called calibration variables (e.g., gender, education, age, urban/rural and region, as suggested in the *GATS Sample Weights Manual*).

General background and instructions for computing the post stratification adjustments are given below in this section. More detail on these topics is given in *Appendix C.2*.

1. **Background** — Calibration adjusts for sample imbalance not accommodated by the nonresponse adjustment. Separate adjustments are applied to all members of strategically formed adjustment cells. One wishes to increase the weights for those population subgroups that remain underrepresented and decrease the weights for those population subgroups that remain overrepresented. The further that the values of calibration adjustments depart from 1.00 (either on the high or low side) the greater the potential impact of sample imbalance (beyond that accommodated by nonresponse adjustment) on the bias of survey estimates.
2. **Producing post-stratification adjustments** — Procedurally, calibration in the form of post-stratification involves forming “adjustment cells” by the cross-classification of the correlate measures. The “post-stratification adjustment” (PSA) in each of these adjustment cells is <1 if the subsample in that category were overrepresented in the sample after accounting for sample selection and nonresponse, and the $PSA > 1$ in those categories where the sample was underrepresented.
3. **Reporting post-stratification adjustments** — Create a table that lists all of the adjustment cells, indicating for each how the cells are defined by the categorical variables used for calibration. For each cell record the computed value of the post-stratification adjustment and observe its size relative to 1.00. It is optimal if all PSAs are close to 1, with some a little greater than 1, and the rest a little less than 1.00.

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

The *GATS Sample Design Manual* calls for a design where selection probabilities (and thus sample weights) will vary somewhat due to the use of estimated measures of cluster sizes, adjustments to sample weights, and equal allocation of sample sizes among regions when regional estimates meeting GATS precision standards are required. The *GATS Sample Weights Manual* describes how these weights are to be computed. Once the questionnaire data have been cleaned, and the final sample weights have been attached, GATS sample data are ready for analysis and this task.

General background and computational instructions for computing this effect are given below in this section. More detail on these topics is given in *Appendix C.3*.

1. **Background** — Variation in sample weights can increase the amount of sampling error in survey estimates and thus lead to larger variances and standard errors of these estimates. More specifically, the multiplicative increase in the variance of survey estimates, referred to as $Meff_w$, depends on how variable the weights are for the set of sample observations that are used to produce the estimate.

The more variable weights are, the larger the value of $Meff_w$.

2. **Estimating $Meff_w$** — The simple mean and variance of the weights for the set of sample observations used to produce survey estimates are needed to compute $Meff_w$.
 - The value of $Meff_w$ for estimates derived from this set of sample observations is computed by first computing the ratio of the variance and the square of the mean, and then adding one to this ratio. The specific computational formula is equation (1) in *Appendix C.3*.
3. **Reporting $Meff_w$** — Since $Meff_w$ applies to all estimates derived from the specific set of sample observations that define a reporting domain for survey estimates (e.g., the over $Meff_w$ all population, the rural population, education, age, gender, etc.),
 - should be computed and reported for all the main population subgroups for which GATS estimates will be computed and reported. This could be done in a table with a list of reporting subgroups and the associated values of $Meff_w$.
4. **Interpreting $Meff_w$** — A proper interpretation of the value of $Meff_w$ computed for a reporting domain is the following:

It is preferable for $Meff_w < 2.00$.

 - “Variation in sample weights increased the variance of all estimates (from the reporting domain) by a factor of ($Meff_w$).”
 - If $Meff_w < 2.00$, efforts to reduce the variation of weights are unnecessary.
5. **Example** — Suppose that for rural estimates $Meff_w = 1.62$

“Variation in sample weights increased the variance of all estimates from rural respondents by a factor of 1.62.”

 - $Meff_w = 1.62$ is considered to be acceptably low for these estimates.

5.2.3 Overall Design Effect on the Precision of Survey Estimates and the Within-PSU Intracluster Homogeneity of Corresponding Key Survey Estimates

The *GATS Sample Design Manual* calls for a design where clusters of survey respondents are selected and the *GATS Sample Weights Manual* describes how sample weights are to be computed. Once the questionnaire data have been cleaned, and the final sample weights have been calculated, GATS sample data are ready to be reviewed before analysis findings are reported.

General background and instructions for computing the overall design effect on a survey estimate and the within-PSU intracluster homogeneity are given below in this section. More detail on these topics is provided in *Appendix C.4*.

1. **Background** — The variance of a survey estimate ($\hat{\theta}$) from a complex sample design divided by the variance of a comparable estimate based on a simple random sample of the same size is called the overall “design effect” for $\hat{\theta}$, or $Deff(\hat{\theta})$. There are two multiplicative components to $Deff(\hat{\theta})$: the design effect due to cluster sampling, $Deff_{CS}(\hat{\theta})$, and the multiplicative effect of variable sample weights, $Meff_w$ (see *Section 5.2.2*). Furthermore, $Deff_{CS}(\hat{\theta})$ is directly related to the amount of within-PSU homogeneity, ρ , and the average sample cluster size.

2. **Estimating $Deff(\hat{\theta})$ and R_b** — An estimate of $Deff(\hat{\theta})$ can be requested directly with some survey analysis software packages (e.g., SUDAAN), or it can be computed from the estimate and its variance when the estimate is a proportion or rate. Since $Deff(\hat{\theta})$ depends on $Deff_{CS}(\hat{\theta})$, R_b , and the average sample cluster size, one can obtain an estimate of R_b by solving for it and using the provided estimates of $Deff(\hat{\theta})$ and these other measures.
 - Detailed formulae to estimate, $Deff(\hat{\theta})$ when survey analysis software do not provide it, and to estimate R_b under any circumstances, are found in equations (1) and (2) in *Appendix C.4*. The resulting estimates are denoted by $deff(\hat{\theta})$ and \hat{R}_b , respectively. We use $deff_{CS}(\hat{\theta})$ to denote the estimate of $Deff_{CS}(\hat{\theta})$.
 - Since there will be many values of $deff(\hat{\theta})$ and \hat{R}_b , summarize them by reporting the median, minimum, and maximum values.
3. **Reporting $Deff(\hat{\theta})$ and R_b** — Estimates of $Deff(\hat{\theta})$ and R_b should be reported for all key study measures (see *Appendix C.4* for definitions).
4. **Interpreting $Deff(\hat{\theta})$ and R_b** — A proper interpretation of the estimated values, $deff(\hat{\theta})$ and R_b , is the following:
 - “The variance of the survey estimate (of the population characteristic, θ), given the GATS sample design, is $(deff(\hat{\theta}))$ times greater than if simple random sampling had produced the same number of respondents.”
 - It is preferred that the overall design effect for an estimate, or $Deff(\hat{\theta})$, is less than 2.50.
 - The amount of total variation in the study measurement (associated with θ) left unaccounted for within PSUs is \hat{R}_b .
 - Rarely is $\hat{\rho}_\theta > 0.15$ or negative; more typically $\hat{\rho}_\theta < 0.05$ or even $\hat{\rho}_\theta < 0.01$.
5. **Example** — Analysis to estimate the current smoking prevalence rate, p , produces the following from a sample where the average sample cluster size was 20:
 $deff(\hat{\theta}) = 1.92$, $Meff_w = 1.36$.
 “The variance of the survey estimate of the current smoking prevalence rate, given the GATS sample design, is 1.92 times greater than if simple random sampling had produced the same number of respondents.”
 - The computed value of $\hat{\rho}_\theta = 0.022$ is typical.

Rarely is $\hat{\rho}_\theta > 0.15$;
 more typically $\hat{\rho}_\theta < 0.05$
 or even $\hat{\rho}_\theta < 0.01$.

5.2.4 Margin of Error for Key Survey Estimates

There are several ways to report the statistical precision of survey estimates. An estimate's margin of error (MOE) is one of them. GATS recommends reporting the estimated margin of error (MOE) along with estimates for key survey measures. The *GATS Sample Design Manual* describes the three main

features of GATS samples that will influence the statistical quality of estimates and findings from the data. These features are the use of stratification, selection of clusters of target population members, and the selection of population members with unequal probabilities (thus requiring the use of sample weights in analysis). Various survey analysis software packages can be used to accommodate these features in analysis of GATS data.

General background and computational instructions are provided in this section. More detail on these topics is given in *Appendix C.5*.

1. **Margin of Error is Measure-Specific** — Each estimate obtained from GATS data has its own margin of error. MOE is by definition the expected half-width of a confidence interval of an estimate of a key survey measure. Another way to interpret MOE is how close the estimate is likely to be to the actual survey measure in the population.
2. **Key Survey Measures** — Estimates of change from one round to another as well as round-specific estimates in GATS will be produced for various behavioral/use rates of tobacco. These estimates are likely to be reported for the population as a whole, as well as by various population reporting subgroups such as by gender, urban/rural, and age categories. Regional estimates may be reported when overall regional sample sizes are at least 8,000.

The following three rates were considered to be most important for planning purposes:

- *Tobacco Use Prevalence Rate* —The percentage of persons currently using tobacco products;
 - *Secondhand Smoking Rate* —The percentage of persons 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.
3. **Estimating MOE** — An estimate of MOE is not usually computed by survey analysis software, but the information needed to compute it is generally available. Three things are needed to compute and interpret MOE: (i) the estimate of the survey measure, (ii) its estimated standard error (or variance), and (iii) a specified measure associated with the desired statistical confidence in the value of MOE.

The level of confidence is usually based on a value (Z) of the standard normal distribution. For example, for a 95% level of confidence, we can use $Z=1.96$.

MOE is computed as the product of the desired confidence measure and the standard error of the estimate. The actual computational formula is equation (1) in *Appendix C.5*.

4. **Reporting MOE** — It is recommended that MOE be computed for all estimates associated with key survey measures. This includes overall national estimates of these measures as well as estimates of these measures for all important reporting subgroups (e.g., by gender, urban/rural, region, and age).

An estimate of MOE should be computed for all estimates associated with key survey measures.

- Key survey estimates and their associated values of MOE should be presented together. Information on reporting estimates of sampling errors and example table shells with GATS key indicators and recommended reporting values are provided in *Appendix C.6*.
5. **Interpretation** — When taken with the value of a survey estimate, MOE indicates how close the estimate is likely to be to the actual value in the population.

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

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

6. **Example** — Suppose that the reported value of a GATS survey estimate is 43.4%, with a standard error of 1.5611% that was computed in accordance with the actual sample design in that country.

The value of MOE for this estimate will be $3.1\% = (1.96)(1.5611\%)$

Interpretation:

“We are 95% confident that the estimate, 43.4%, is within 3.1% of the corresponding population value.”

5.3 Measures of Quality: Coverage, Nonresponse, and Other Non-Sampling Errors

5.3.1 Household Frame Coverage Rate

As described in the *GATS Sample Design Manual*, the penultimate stage of GATS sampling involves selecting a sample of households among those enumerated in PSUs or segments. The quality of the household sample and ultimately of GATS estimates thus depends on the completeness of these household lists, which are prepared in a variety of ways.

General background and instructions for computing a household frame coverage rate (FCR) are given below in this section. More detail on these topics is given in *Appendix C.7*.

1. **Background** — A key part of the sampling process in a household survey is the enumeration of households in the sampling stage where household addresses are chosen. Different household enumeration approaches may be used (e.g., traditional household listing by field enumerators, modified voter lists). The usual end-result is that there is a net under-enumeration of households and thus incomplete coverage in the final sample of households. Since the biasing effect of undercoverage depends both on: (i) the aggregate difference between key study measures for households covered by the frame and those missed by the frame, and (ii) the extent of noncoverage (or 100 percent minus the household coverage rate) in the sample, it is useful to estimate the frame coverage rate for households.
 - A frame coverage rate (FCR) is defined as the number of households accounted for in the household frame divided by the number of households in the population.
 - While useful in describing sample household frame coverage, often data are not available from a recent census on the number of total households in the population. If available, FCR

should be computed but it is not a critical data element for the GATS official review of statistical quality measures.

2. **Estimating the FCR** — The numerator of the FCR can be estimated using the sum of pre-adjustment (base) weights over all selected households. The denominator must be obtained from some current and highly reputable source (e.g., a recent census or some other recent survey of superior sample size to GATS).
3. **Reporting the FCR** — The estimated FCR should be reported in the *GATS Country Report* with the description of the sample design or with a separate assessment of the data quality.
4. **Interpreting the FCR** — The FCR indicates the extent of frame coverage of households but not the size of coverage bias, since the latter depends on the former plus the aggregate difference (in key survey measures) between covered and non-covered households. A household coverage rate above 95% would be considered acceptable.

5. **Example** — The best available external estimate of the number of households in the target population of a GATS country is 36,436,103, while the weighted sum of the 13,000 selected sample households for GATS is

A household coverage rate above 95% would be considered acceptable.

34,788,543. The resulting coverage rate of $0.955 = \frac{34,788,543}{36,436,103}$ would be considered acceptable.

5.3.2 Patterns of Respondent Cutoff Rates

Some survey interviews are not fully completed. The respondent may find participation in the interview to be burdensome and thus lose patience with the interviewer, or the respondent may need to end the interview early due to legitimate scheduling limitations and not be able or willing to reschedule a time to finish the interview. In either scenario, a partially completed interview is an indicator of respondent disengagement, which may be seen as a reflection on the quality of interviewer efforts or the respondents' attitudes towards the survey topic and thus ultimately the quality of the data.

- **Data Source(s):** The data file with final household/person disposition codes should be used for these calculations (see *Appendix C.1*). An IQ disposition code of 11 indicates that the respondent completed at least Sections A through D of the GATS individual interview. An IQ disposition code of 12 indicates the interview was discontinued before Sections A through D could be completed.
- **Method/Computation:** Define a survey "respondent" to be any selected individual who is assigned a final IQ disposition code of 11 or 12. Also define a cut-off rate (*COR*) to be:

$$COR = \frac{R_c}{R} = \frac{\# \text{ of GATS respondents who only partially completed the interview}}{\# \text{ of GATS respondents}}$$

or

$$= \frac{\text{Disposition code 12}}{\text{Disposition codes (11 + 12)}}$$

- **Uses:** Values of COR should be computed directly for the sample as a final stage of quality assurance after data collection is completed. Moreover, ongoing efforts to identify and take remedial action on ineffective FIs should be routinely done by FSs throughout data collection (using corresponding field result codes for computation). This would entail computation of COR by individual interviewer or interviewing team, and identifying those interviewers/interviewer teams that have exceptionally high rates (e.g., rates that are greater than the mean plus two standard deviations of rates among all interviewers/teams). Values of COR might also be computed by: (1) the week of data collection in which the interview took place, (2) respondent age, (3) respondent gender, and (4) respondent urban/rural location.
- **Interpretation:** Generally, the lower the value of COR the better. A drop in COR over the span of data collection would indicate that respondents may have been less motivated to provide useful data to GATS towards the end of data collection. While cutoff rates are useful measures of data collection performance and overall survey quality, computation for the official GATS review of statistical quality measures is not critical.

5.3.3 Patterns of Household Response Rates by First Stage Sampling Strata

Household Response Rate (HRR) is computed as the ratio of the weighted number of households for which complete roster data have been provided to the weighted number of in-scope households (see formula 19, *GATS Sample Weights Manual*). Households providing only partial roster data are treated as refusals or break-offs for the household response rate and not included in the numerator of the household response rate calculation. The *Household Questionnaire* is treated as complete if it has a household final disposition code of 1 (see *Appendix C.1*). Using only in-scope households, the household-level response rate is computed as follows:

$$\text{Household Level RR} = \frac{[1]}{[1] + [3] + [4] + [5] + [6] + [9]} \times 100$$

where

- 1 = Completed Household Questionnaire, One Person Selected
- 2 = Completed Household Questionnaire, No One Selected
- 3 = Completed Part of Household Questionnaire, Could Not Finish Roster (Incomplete Interview)
- 4 = Household Questionnaire Not Complete, Could Not Identify Screening Respondent
- 5 = Nobody Home
- 6 = Household Refusal
- 9 = Other Household Nonresponse.

Note that the HQ final disposition code 2 is excluded from both numerator and denominator of the household-level response rate since these households are considered ineligible. This response rate conforms to RR1 as defined by the American Association of Public Opinion Research (AAPOR) (<https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx>).

Full documentation of how household response rates are calculated should be provided by countries. Household response rates should be presented for the entire country by first stage sampling strata (i.e.,

urban/rural, and region) if sample sizes are sufficient for regional estimation. Further break down, if appropriate, should be made using cross classification of urban/rural and region.

5.3.4 Patterns of Person-level Response Rates Among Variables Used for Nonresponse Adjustments

A person-level response rate is computed as the ratio of the weighted number of partially or fully completed GATS interviews to the weighted number of eligible sampled individuals chosen from the roster of all household members. Specifically, person-level final disposition codes are as follows:

- 11 = Completed Individual Questionnaire
- 12 = Incomplete Interview
- 13 = Selected Individual Was Later Determined to Be Ineligible for GATS
- 14 = Selected Respondent Not Home
- 15 = Selected Respondent Refusal
- 16 = Selected Respondent Incompetent
- 17 = Other Individual Nonresponse.

One way to compute a person-level response rate is as follows:

$$Person - Level RR = \frac{[11]}{[11] + [12] + [14] + [15] + [16] + [17]} \times 100$$

The person-level response rate above assumes that all of the selected individuals with unknown eligibility (i.e., final disposition code 14) are eligible to participate in the GATS. This may lead to underestimates of person-level response rates if interviewers often select respondents who are found to be ineligible for the survey once the interview begins. It is recommended that countries estimate the proportion of those respondents selected from the roster who are truly eligible to respond to the GATS survey (*e*) using weighted dispositional code frequencies as:

$$e^{person-level} = \frac{[11] + [12] + [15] + [16] + [17]}{[11] + [12] + [13] + [15] + [16] + [17]}$$

If this proportion for the sample as a whole is less than 0.90, countries should adjust the unknown component of the person-level response rate by multiplying unknowns (final disposition code 14) by this proportion (*e*). The following formula conforms to the AAPOR (2004) response rate RR3:

$$Person - Level RR = \frac{[11]}{[11] + [12] + [e^{person-level}[14]] + [15] + [16] + [17]} \times 100$$

This response rate conforms to RR3 as defined by AAPOR and documented above.

Full documentation of how response rates are calculated should be provided by countries.

Person-level response rates should be presented for the entire country and by regional subareas when sample sizes are sufficient for regional estimation. Further breakdown within region (if appropriate) should

be made for those variables used to define nonresponse adjustment. These may have included urban/rural and the roster-reported age (15-24, 25-34, 35-44, 45-54, 55+), gender (male, female) and current smoking status (smoker, non-smoker).

The *GATS Field Supervisors Manual* describes FS responsibilities for review and remediation of low interviewer team response rates.

5.3.5 Patterns of Person-level Refusal Rates Among Variables Used for Nonresponse Adjustments

Refusal rates are computed as the ratio of the number of weighted eligible, selected respondents who refused participation for the GATS *Individual Questionnaire* (*Individual Questionnaire* final disposition code 15) to the total weighted number of completed (11) and incomplete (12) individual GATS questionnaires, plus non-respondents of unknown eligibility (14), those who refused (15), incompetent respondents (16) and other individual nonrespondents (17). Using individual level disposition codes, the computation is:

$$\text{Person – Level Refusal Rate} = \frac{[15]}{[11] + [12] + [14] + [15] + [16] + [17]} \times 100$$

This refusal rate conforms to REF1 as defined by the American Association of Public Opinion Research (AAPOR).

Person-level refusal rates should be computed directly for the sample as a final stage of quality assurance after data collection is completed. Moreover, ongoing efforts to identify and take remedial action on ineffective FIs should be routinely done by FSs throughout data collection. This would entail computation of refusal rates by individual interviewer or interviewing team, and identifying those interviewers/interviewer teams that have exceptionally high rates (e.g., rates that are greater than the mean plus two standard deviations of rates among all interviewers/teams). These refusal rates might also be computed by: (1) the week of data collection in which the interview took place, (2) respondent age, (3) respondent gender, and (4) respondent urban/rural location.

Full documentation of how refusal rates are calculated should be provided by countries. Refusal rates should be presented for the entire country and by regional subareas when sample sizes are sufficient for regional estimation. Further breakdown within region (if appropriate) should be made for those variables used to define nonresponse adjustment. These may have included urban/rural and the roster-reported age (15-24, 25-34, 35-44, 45-54, 55+), gender (male, female) and current smoking status (smoker, non-smoker).

5.3.6 Item Nonresponse Rates for Fact Sheet Indicator Variables

A GATS item nonresponse rate (i.e., the percentage rate of all respondents who fail to provide useful data for a specific interview questionnaire item among all respondents who should have answered the question) should be computed and documented after processing of the survey data is completed for the questionnaire items corresponding to all indicators that will be included in the country-specific *GATS Fact Sheet* (see *GATS Indicator Definitions* for more information). Item nonresponse rates (NRR) are computed as the ratio of the number of respondents for whom an in-scope valid response was not

obtained (*MISSING*^x for item x), to the total number of unit level respondents (I) minus the number of respondents with a valid skip for item x (*V*^x):

$$NRRI^x = \frac{MISSING^x}{(I - V^x)} \times 100$$

The total number of unit level nonrespondents of x (*MISSING*^x) will be obtained from an unweighted frequency of respondents with missing data for item x after appropriate cleaning to ensure proper skip patterns were followed. The total number of unit level respondents will be obtained from the total unweighted frequency of responding male or female to question A01 (a variable with no anticipated blank fields). The total number of respondents with a valid skip for item x can be obtained as the frequency of item x with a response of .S (see *Section 5.1.3*). Many of the *GATS Fact Sheet* indicators are composite measures obtained from responses to several questions. Nonresponse to any item used to obtain a composite measure will result in item nonresponse of the composite measure. Item nonresponse rates below 5% are considered low. Similar to quality measures 5.3.2 and 5.3.5 above, ongoing efforts by FSs should target excessive item nonresponse by individual interviewers or interviewer teams as a means to improving the quality of data collection as the data are being collected in the field.

5.4 Formal Review of Statistical Quality

A scientific protocol has been developed to conduct an official review of statistical quality measures in each GATS country by SRC after the survey data have been processed and the sample weights have been computed, but before the data analysis is conducted and the results are produced. The purposes of this review are to verify that sample weights were computed in accordance with specifications in the *GATS Sample Weights Manual*, to examine various measures indicating high quality survey data, and to assure that key aspects of sampling and analysis are fully documented. Technical assistance in preparation for this review will be available.

Described next are the specific quality measures and documentation that will be needed to conduct these reviews.

5.4.1 Sample Weights Documentation

Each step in the data weighting and adjustment process should be fully described in a document prepared by country statisticians directly responsible for computing and adjusting statistical weights for the GATS sample. The document should include a narrative that describes in detail the weighting and adjustment process, and including illustrative examples of precisely how all components of the final weights were computed for 5-10 actual sample respondents. More specifically, each of the following should be prepared for examination by those conducting the review¹:

- Summarize the survey sample design by stage and describe how the selection probabilities for each stage and the overall base weight were computed.
- Describe in detail how nonresponse and calibration adjustment strategies were completed, indicating specifically: (i) how weighting classes were defined for household- and person-level

¹ Note that since many of the items on this list should also be included with the documentation for each *GATS Country Report*, drafts of text of this documentation (but containing the final versions of the required numerical results) can be sent to those conducting the official review of statistical quality for the country.

nonresponse adjustments, (ii) what adjustment cells were formed to calibrate the weights, and (iii) to which population counts the weights were calibrated. Refer to the *GATS Sample Weights Manual* as a guide for this description.

- Choose 5-10 actual respondents to demonstrate the weights computation process that was followed for them. Respondents should be selected over different strata, PSUs, and gender groups. If weighting components based upon half-open intervals were used, a respondent from one of these geographical clusters should be included as one of the 5-10 examples. At the conclusion of this step, computation of the final adjusted weight should be displayed for each example. More specifically, for each of these respondents:
 - Identify the components of each conditional and unconditional probability (from each sampling stage) that was used to produce the BASE WEIGHT.
 - Present and fully describe each component in the computation of household- and person-level NONRESPONSE ADJUSTMENT. At the conclusion of this step, base weights should be adjusted for household- and person-level nonresponse.
 - Present and fully describe each component of the CALIBRATION WEIGHT ADJUSTMENT method to force the final weighted distribution of the sample to match a chosen calibration population.
- Present the following computational results based on formulas given in *Sections 5.2 and 5.3* of this manual:
 - Post-stratification calibration adjustments for each adjustment cell used in calibration (5.2.1)
 - Multiplicative effect of variable sample weights (*Meff*) by all population subgroups for which survey estimates will be reported in the *GATS Country Report* (5.2.2)
 - Individual and summary measures (i.e., minimum, maximum, and median) of the overall design effect and measure of within-PSU homogeneity for all survey estimates presented in the *GATS Fact Sheet* (5.2.3)
 - Individual and summary measures (i.e., minimum, maximum, and median) of the Margin of Error (*MOE*), or other comparable measure of the precision, for all survey estimates presented in the *GATS Fact Sheet* (5.2.4)
 - Individual and summary measures (i.e., minimum, maximum, and median) of household response rates by the adjustment cells used to produce the household-level nonresponse adjustment (5.3.3)
 - Person-level response rates among the adjustment cells used for person-level nonresponse adjustments (5.3.4)
 - Person-level refusal rates among the adjustment cells used for the person-level nonresponse adjustment (5.3.5)
 - Individual and summary measures (i.e., minimum, maximum, and median) of item nonresponse rates for all survey items used to produce the tobacco use rates in the *GATS Fact Sheet* (5.3.6)

The *GATS Sample Weights Manual (Chapter 4)* presents several suggested methods to better ensure the quality of sample weights. The following checks will be completed, and thus the data to complete them will be required, by those conducting the statistical quality review:

- The average size of the base weights, divided by the average size of the nonresponse-adjusted weights, is roughly equal to the final overall person-level response rate.
- The weighted distribution of the sample (using the final adjusted weights) among all calibration adjustment cells should precisely match the corresponding distribution of external population counts to which the weights were calibrated.
- The sum of the final adjusted weights over the full sample should be the total size of the population count (of those 15 years of age or older) for the source used for calibration. If that source was the last census, then this sum of weights should be the target population count from the last census.
- Most of the post-stratification adjustments (by adjustment cell) should be either slightly greater or less than 1.
- The *Meff* is ≤ 2.00 for each reporting subgroup for the *GATS Country Report*.

5.5 Creation of Analytic Data File

After the sample weighting and all quality assurance checks have been completed, a new file should be created containing only cases with the individual level final disposition code of 11 (Completed *Individual Questionnaire*). Only cases with an individual level final disposition code of 11 will be considered “respondents” to the *GATS Individual Questionnaire*.

Note that it is not necessary to create the final analytic dataset before conducting quality assurance checks using MOE, design effect or key survey estimates calculated by SAS or SPSS procedures for complex samples, because cases with IQ final disposition codes other than 11 will not have values for the weights variable (i.e., they will have missing values) and will be omitted automatically by the software as a default unless otherwise specified.

This new file is called the analytic data file that should be used when conducting data analyses to produce survey estimates (i.e., indicators), including the estimates that go into the *GATS Fact Sheet* and *GATS Country Report*.

5.6 Public Use Data File Creation

The Public Use Data Set (PUDS) is the data file released to the general public by CDC. It includes completed interview cases only with all survey response variables and standard GATS core indicator variables. Geographical variables and survey sampling variables such as stratum, cluster and final survey weights can also be found in the data file.

The PUDS is publicly released to provide external researchers access to GATS data. Based on the *GATS Data Release Policy*, it will be released one year after the country releases their *GATS Fact Sheet*.

6. Data and Documentation Requirements

This chapter describes the data files and documentation that will be sent from the countries and archived at the GATS Data Coordinating Center.

Data files and documentation requested from country:

1. Final programmed DB3 files: CMSDB.db3, Survey0.db3 and Survey1.db3.
2. Master DB3 file created by the Aggregation software from CMSDB.db3, Survey0.db3 and Survey1.db3.
3. Master database file (using English language variable) generated using the Generate Analysis Files utility containing transposed flat files for Responses table with valid responses for both the *Household* and *Individual Questionnaires* and Master Sample Selection file (Case ID and sample identifiers).
4. SAS, SPSS, and STATA syntax files.
5. Sample Weights components and documentation (refer to *Section 5.4*).
6. Weighted data set: the final, clean flat file(s) with weights, weighting specifications.

7. Bibliography

- Biemer, Paul P. and Lars E. Lyberg. 2003. Introduction to Survey Quality. John Wiley and Sons, Inc., Hoboken, NJ.
- Encyclopedia of Survey Research Methods*, (2007). Lavrakas, P.J. (Ed.), Sage Publishing, Inc. Los Angeles, CA.
- Gabler, S., Haeder, S., and Lahiri, P. (1999). "A model based justification of Kish's formula for design effects for weighting and clustering," *Survey Methodology*, 25 (1): 105-106.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Core Questionnaire Programming Specifications*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Core Questionnaire with Optional Questions*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Data Management Implementation Plan*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Data Management Training Guide*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Field Interviewer Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Field Supervisor Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Mapping and Listing Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Programmer's Guide to General Survey System*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Question by Question Specifications*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Sample Design Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Global Adult Tobacco Survey Collaborative Group. *Global Adult Tobacco Survey (GATS): Sample Weights Manual*. Atlanta, GA: Centers for Disease Control and Prevention, 2019.
- Internal Reference for SAS Special Missing Values:
<http://support.sas.com/documentation/cdl/en/lrcon/61722/HTML/default/a000992455.htm>
- Kish, L. (1965) *Survey Sampling*, Wiley and Sons, New York.
- Kalton, G. and Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics*, 19(2), 81-97.
- Potter, F. J. (1988). Survey of Procedures to Control Extreme Sampling Weights. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 453-458.

Appendix A: Glossary of Terms

Case File: The sample list and information that is loaded onto the handhelds for data collection.

Core Questionnaire: The standard base questions used in GATS. Participating countries adapt the core questionnaire to their specific needs by following specific guidelines.

DB3 files: The programming of the GSS survey software is done using DB3 files (SQLite).

General Survey System (GSS): The survey software program that is loaded on to the handheld devices to conduct the GATS survey.

GSS Programming Specifications: In addition to the paper questionnaire that a country uses to create their country-adapted version, GSS programming specifications are developed for each country-adapted version which are used for programming the questionnaire for the handheld.

Handheld Setup: The process of copying the build software to the handheld, typically performed on all handhelds at once before the start of Field Interviewer training.

Master Sample Selection File: The dataset that contains a CaseID for every household in the sampling list along with the information necessary to compute the sample weights and analyze the complex survey data.

Quality Assurance: A process consisting of systematic activities designed to ensure, assess, and confirm the quality of the data collected during a survey.

Questionnaire Review Committee (QRC): A group comprised of international experts in tobacco control and questionnaire design. The QRC reviews and approves all GATS country-adapted questionnaires to ensure quality, standardization, and comparability.

Sampling Review Committee (SRC): A group comprised of international experts in sampling. The SRC reviews and approves the sample designs, sample weights, and data quality assurance measures for participating countries.

Appendix B: Pre-Data Collection

B.1 GATS Questionnaire Adaptation and Review Process

Questionnaire Adaptation: GATS countries will work with CDC to develop adapted country specific GATS questionnaire which includes the following:

- **Adapting the GATS Core Questionnaire.** Countries will adapt the core questionnaire for their specific situation. This includes modifying question item lists (e.g., types of tobacco) and response options (e.g., list of cigarette brands) that are appropriate for the country.
- **Incorporating optional questions.** The GATS maintains a suggested list of optional survey questions that countries can choose to add throughout the questionnaire.
- **Including additional questions developed by the country.** Countries can include their own survey questions that are not included in the core questionnaire and optional questions.
- **Deleting irrelevant questions.** Countries can remove core questions that are not relevant to their situation (e.g., no prevalence of smokeless tobacco). Proper justification should be provided.

The GATS Questionnaire Review consists of the following 10-Step Process:

1. **Adaptation of country-specific draft questionnaire.** Countries will work with CDC to create a country-specific draft questionnaire to submit to the QRC.
2. **Formatting.** CDC will review the submitted draft questionnaire and will format as needed (e.g., edit skip instructions, highlighting for QRC review).
3. **Expert review by Questionnaire Review Committee.** Members of the QRC review the questionnaire and provide comments.
4. **Compilation of reviewers' comments.** CDC compiles the QRC's comments and sends them back to the country for review.
5. **Communication (resolution).** The QRC and country work together to resolve any outstanding issues (via email and/or conference calls).
6. **Finalize adapted questionnaire.** Once all issues are resolved, the country-adapted questionnaire is finalized and official approval is provided by the QRC.
7. **Notification of acceptance by country.** The country grants their final acceptance of the QRC approved questionnaire.
8. **Translation.** Countries will translate the approved GATS questionnaire into appropriate language(s) and also provide a back-translation of the questionnaire into English. (Note: This might occur simultaneously with step 1 where countries first translate the core questionnaire and adapt in their own languages, then back-translate into English to submit to the QRC.)
9. **Verification of final translated and back translated versions.** The QRC will review the translated and back translated versions to verify the translation is appropriate (i.e., meaning of questions has not changed). (Note: This is part of the QRC review process if the translation is done simultaneously with step 1 and a back-translated version is initially submitted for QRC review.)
10. **Programming.** Programming specifications for the country-adapted questionnaire are created and the country-specific handheld survey is programmed.

B.2 GATS Questionnaire Programming Process

The major steps and timeline guidelines for GATS questionnaire programming are as follows:

1. GATS country questionnaire and pretest proposal approved 6-8 weeks prior to pretest training.
2. Questionnaire programming in English begins 6 weeks prior to pretest training.
3. Handheld hardware delivered and operational 4 weeks before pretest training.
4. Translations of GSS system menus and messages, if needed, completed 1 week before pretest training.
5. Translation of questionnaire texts (*Household* and *Individual Questionnaires*) in all host country languages completed and imported into survey0 and survey1 databases 2 weeks before pretest training.
6. Host country signoff on *Household* and *Individual Questionnaires* (HQ and IQ) 1 week before pretest training.
7. Start version control of questionnaires with the signed off version above and maintain strict control after this point.

B.3 Process for Setting Up Handhelds with GSS Program

This section describes the quality measures that need to be implemented at the different stages of the process to ensure that each handheld going to the field has the correct and same software and hardware specifications.

1. **Creating the master build:** The Master build should be finalized on a single machine that is free from any virus and that has proper antivirus software installed. After creation of the master build, this should be loaded on two different handhelds and a complete iteration of testing should be conducted on both the handhelds answering each question on the questionnaire. If any errors are found on any handhelds, the programming—as mentioned in *Section 3.3.2*—should be corrected and this process should be repeated until 100% correctness is achieved. This process should be conducted in the presence of the coordinator/focal person, along with the IT lead.
2. **Paste QC check stickers (template):** All handhelds should have a sticker with the FIID, serial number, and a blank QC checklist—placed on the back of the handheld. See example below:

701010		SN A10
	<u>QC1</u>	<u>QC2</u>
<input type="checkbox"/> Initial Config	<input type="checkbox"/> Date	<input type="checkbox"/> Date
<input type="checkbox"/> APK Loaded	<input type="checkbox"/> FIID	<input type="checkbox"/> FIID
<input type="checkbox"/> Transmission	HQ ____	<input type="checkbox"/> Delete Training Cases
	IQ ____	<input type="checkbox"/> Power ____
	<input type="checkbox"/> Power ____	Transmission
Ini ____	Transmission	Cases ____
	QC1 ____	QC2 ____

3. **Prepare the handhelds:** After the final build is copied, all the handhelds should be setup using the proper step-by-step process (refer to the process on preparing handhelds described in the *GATS Programmer's Guide to General Survey System*). The process of separating the handhelds in different boxes (before setup and after setup) should again be undertaken.
4. **Conduct Quality Control Stage 1 (QC1):** The handhelds should be equally divided between the number of quality control staff who should be from within the team who have setup the handhelds or have supervised the process. Each step of QC1 should be conducted on every handheld. QC1 should be a rigorous quality check process and the following items should be checked for correctness on each handheld and then marked on the sticker with the checklist:
 - 1) Date and time
 - 2) FIID (Verify sticker with HQ ID)
 - 3) Household Qx version
 - 4) Individual Qx version
 - 5) Power – battery Life

Other checks could be added based on the country setting such as number of cases, handhelds for male/female interviewers, etc. For any identified problems, the handheld should go to a different box and should be reset again, and the entire process should be repeated on these handhelds starting from step 1.
5. **Conduct Quality Control Stage 2 (QC2):** QC2 is conducted after all the handhelds are cleared through QC1 and signed off on the sticker. The handhelds should be reallocated in such a way that the person who does Quality Control Stage 1 is different from the person who does Quality Control Stage 2. The same process as QC1 above should be conducted during QC2; the additional step in QC2 is to check if all the stickers are marked and checked off as well. It is important to conduct QC2 since it has been observed that a small proportion of errors are picked up during QC2. Any handhelds with problems identified during QC2 should be put through the iteration again starting with a hard reset in step 1.
6. **Sign off on completion sheet by IT lead and Country Focal person:** After each and every handheld has passed through both QC1 and QC2, a quick summary report outlining the number of handhelds, HQ and IQ version numbers, date, time and members of the handheld production and quality control team should be documented and signed off by the IT lead and the country GATS focal person. The sign off should state that the handhelds are finalized for full survey training/field work and no further programming changes will occur, and they will not be used for any other purpose until GATS is completed in the country.

Appendix C: Post-Data Collection

C.1 Final Disposition Codes and Response Rate Calculations

After a careful review of the literature on tobacco and non-tobacco related national household surveys using either telephone or face to face mode of data collection in the household, a list of final disposition codes and response rate calculations were finalized for GATS. Several surveys were reviewed for this purpose including the Brazilian Institute for Geography and Statistics (IBGE) surveys, 2004 national Survey of Demography and Reproductive and Sexual Health (ENDSSR-2004), Georgia Reproductive Health Survey (RHS), Demographic and Health Surveys (DHS), National Family Health Survey (NFHS), Behavioral Risk Factor Surveillance System (BRFSS), Youth Risk Behavior Surveillance System (YRBSS), and the American Association for Public Opinion Research (AAPOR).

The final disposition codes and response rate reporting are described in the following tables in this section. *Tables C-1 and C-2* provide the final HQ and IQ disposition codes that should be assigned for various final result codes from the HQ and IQ questionnaires. (Further details on result codes are found in the *GATS Field Interviewer Manual*.) *Table C-3* provides an example template for how to calculate and report response rates and disposition codes in the *GATS Country Report*.

Table C-1: Household Questionnaire Final Disposition and Result Codes

Final Disposition Codes	Final Result Codes	Name and Description
1	200	Completed Household Questionnaire, One Individual Selected -Household Questionnaire is considered complete if the roster is complete
2	201	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	202	Completed Part of Household Questionnaire, Could Not Finish Roster -Incomplete Interview
4	203	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	209	Nobody Home
6	204	Household Refusal
7	205	Unoccupied House
8	206	Selected Address is Not a Household
9	208	Other Household Nonresponse

Table C-2: Individual Questionnaire Final Disposition and Result Codes

Final Disposition Codes	Final Result Codes	Name and Description
11	400, 402*	Completed Individual Questionnaire -Questionnaire completed through question E01 -None of the key questions (e.g., B01/B02/B03 and C01/C02/C03 where applicable) = Don't Know or Refused
12	402*	Incomplete -Questionnaire broken-off before question E01
13	403	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	409	Selected Respondent Not Home
15	404	Selected Respondent Refusal
16	407	Selected Respondent Incapacitated
17	408	Other Individual Nonresponse

* Cases with a final result code 402 (Completed Part of Individual Questionnaire) will either be assigned a disposition code of 11 (Completed IQ) or 12 (Incomplete). See the rules outlined in *Section 5.1.7*.

Table C-3: GATS Country Report Table 3.1—Template for Reporting Response Rates

Table 3.1: Number and percent of households and persons interviewed and response rates by residence (unweighted) – GATS [Country], [Year].						
	Residence				Total	
	Urban		Rural		Number	Percent
	Number	Percent	Number	Percent		
<i>Selected Household</i>						
Completed (HC)						
Completed – No one eligible (HCNE)						
Incomplete (HINC)						
No screening respondent (HNS)						
Nobody home (HNH)						
Refused (HR)						
Unoccupied (HUO)						
Address not a dwelling (HAND)						
Other ¹ (HO)						
Total Households Selected		100		100		100
<i>Household Response Rate (HRR) (%)²</i>						
<i>Selected Person</i>						
Completed (PC)						
Incomplete (PINC)						
Not eligible (PNE)						
Not at home (PNH)						
Refused (PR)						
Incapacitated (PI)						
Other ¹ (PO)						
Total Number of Sampled Persons		100		100		100
<i>Person-level Response Rate (PRR) (%)³</i>						
<i>Total Response Rate (TRR) (%)⁴</i>						
¹ Other includes any other result not listed.		³ The Person-level Response Rate (PRR) is calculated as:				
² The Household Response Rate (HRR) is calculated as:		$\frac{PC * 100}{PC + PINC + PNH + PR + PI + PO}$				
$\frac{HC * 100}{HC + HINC + HNS + HNH + HR + HO}$		⁴ The Total Response Rate (TRR) is calculated as: (HRR x PRR) / 100				
Notes:						
— An incomplete household interview (i.e., roster could not be finished) was considered a nonrespondent to the GATS. Thus, these cases (HINC) were not included in the numerator of the household response rate.						
— The Total Number of Sampled Persons should be equal to the number of Completed [HC] household interviews.						
— A completed person interview [PC] includes respondents who had completed at least question E01 and who provided valid answers to questions B01/B02/B03 (and C01/C02/C03 where applicable). Respondents who did not meet these criteria were considered as incomplete (PINC) nonrespondents to GATS and thus, were not included in the numerator of the person-level response rate.						

C.2 Pattern of Post-Stratification Weights Calibration Adjustments Among Adjustment Cells

Background

The last step in producing sample weights involves calibrating the weights to population counts by known correlates of key study outcome measures, called calibration variables (e.g., gender, education, age, urban/rural and region, as suggested in the *GATS Sample Weights Manual*). Procedurally, calibration in the form of post-stratification involves forming “adjustment cells” by the cross-classification of the correlate measures. The “post-stratification adjustment” (PSA) in each of these adjustment cell is <1.00 if the sample in that category were overrepresented in the sample after accounting for sample selection and nonresponse, and the $PSA > 1.00$ in those categories where the sample was underrepresented. Examining the pattern of PSAs among adjustment cells helps to understand where in the sample imbalance remained after sample imbalance due to nonresponse was corrected.

Data Source(s)

A data file containing the final set of sample weights adjusted for nonresponse, as well as all computational components of each respondent’s final weight should be used for these calculations.

Computational Software

None in particular since the computations required can easily be programmed in any software environment.

Computation

Procedural steps for computing post-stratification adjustments are described in *Section 3.3* of the *GATS Sample Weights Manual*. Using the final computational results of these calculations, arrange the set of PSAs as a multi-way table by the correlate variables that defined the adjustment cells (i.e., the calibration variables). One hopes to find that all PSAs are close to 1.00, some a little greater than 1.00, the rest a little less than 1.00.

Interpretation

Large departures from 1.00 in these PSA suggests that there were serious problems with sample imbalance due to differential coverage of the household sample, undetected problems with sample weights calculations, or major demographic shifts in the target population between the date of the population counts to which the sample was calibrated and the date of sampling and data collection in the GATS survey.

Table C-4: Hypothetical example table of post-stratification weights calibration adjustments, among region, urban/rural, gender, and age.

Region	Urban/Rural	Gender	Age	Adjustment
1	Urban	Male	15 – 24	
			25 – 44	
			45 – 59	
			60+	
	Rural	Female	15 – 24	
			25 – 44	
			45 – 59	
			60+	
2	Urban	Male	15 – 24	
			25 – 44	
			45 – 59	
			60+	
	Rural	Female	15 – 24	
			25 – 44	
			45 – 59	
			60+	
3	Urban	Male	15 – 24	
			25 – 44	
			45 – 59	
			60+	
	Rural	Female	15 – 24	
			25 – 44	
			45 – 59	
			60+	

C.3 Effect of Variable Sample Weights on the Precision of Survey Estimates

Background

Variation in sample weights often creates an increase in the variance (i.e., reduction in precision) of survey estimates. This kind of variation will be an important issue in country sample designs that have oversampled by urban/rural or by region. Kish (1965, *Section 11.7*) proposed a simple but long-standing model to measure the multiplicative increase in variance when the size of weights is uncorrelated with the survey measurement used to estimate a simple population characteristic (θ). We will use $Meff_w$ to denote this multiplicative effect.

$Meff_w$ is not estimate-specific but it is specific to reporting domain for estimates. It will thus apply to all GATS estimates in domain. For instance, $Meff_w$ computed for overall population estimates should be computed from the entire sample, but $Meff_w$ calculated for estimates reported for persons living in rural areas should be computed from rural respondents only.

Data Source(s)

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

Computational Software

No special software is needed. The computation of $Meff_w$ can easily be programmed using output from the survey analysis software that is used.

Computation

Focusing on the final weight (w_i) computed for each of n respondents to whom $Meff_w$ applies, compute the simple average of the weights to obtain \bar{w} , and then use the formula,

$s_w^2 = \sum_{i=1}^n (w_i - \bar{w})^2 / (n - 1)$ to calculate the variance of the sample weights. We can calculate $Meff_w$

from the final set of sample weights for GATS respondents as,

$$Meff_w = 1 + \frac{s_w^2}{\bar{w}^2}, \quad (1)$$

Note that $Meff_w \geq 1$ since \bar{w}^2 and s_w^2 will always both be positive.

Interpretation

While one clearly prefers for $Meff_w$ to be as close to one as possible, $Meff_w > 2.00$ might be viewed as substantial and thus requiring some form of remedial action. Weight trimming or censoring strategies are the most common remedy for excessively variable weights (Potter, 1988). The final decision as to whether one should trim the weights depends on finding a balance between the benefit on precision resulting from trimming and the effect on survey estimates resulting from changes in the weights. If weight trimming reduces $Meff_w$ but does not appreciably change weighted estimates for key study outcome measures, the trimming step may be justifiable given the required time and effort to do so.

Example

For a GATS sample the mean of the final sample weights among all X,XXX survey respondents is $\bar{w} = YYY.YYYY$ and their simple variance is $s_w^2 = ZZZZ.ZZ$. The multiplicative effect of variable

weights on overall national estimates in a country is $Meff_w = 1 + \frac{YYY.YYYY}{XXX.XX} = U.UU$

A sample with $Meff_w = U.UU < 2.00$ can be considered as one where variation in weights is sufficiently modest to not require any efforts to reduce the effect of variable weights on estimates, such as weight trimming (see Potter, 1988).

C.4 Overall Design Effect on the Precision of Survey Estimates and the Within-PSU Intracluster Homogeneity of Corresponding Key Survey Estimates

Background

Kish (1965) originally defined a “design effect” denoted by ‘deff’ to measure the multiplicative change in the variance of a survey estimate due to the use of cluster sampling for a given sample design, as compared to the variance of a comparable estimate from a simple random sample of the same size. Each of the GATS key survey outcome measures is a proportion (p), or a rate expressible as a percentage or an average (mean). Kish’s original model for the design effect (due to cluster sampling) of an estimate (\hat{p}) of p is simply,

$$Deff_{CS} = 1 + \rho_{\theta}(\bar{m} - 1) ,$$

where ρ_{θ} is a respondent-within-PSU measure of relative intra-cluster homogeneity for the survey measure used to estimate p (measuring the degree to which members of the same sample cluster are more similar than the members at large in the population), $\bar{m} = n/a$ is the average number of respondents per sample PSU, n is the overall respondent sample size, and a is the number of sample PSUs.

When a survey sample involves both sampling clusters and unequal sample weights, the overall design effect on a survey estimate, denoted by $Deff(\hat{\theta})$, has been shown by Gabler, et al. (1999) to be the product of due to both cluster sampling (see above) and the multiplicative effect of variable sample weights (see discussion of $Meff_w$); i.e.,

$$\begin{aligned} Deff(\hat{\theta}) &\equiv \frac{V(\hat{p}; design, n)}{V(\hat{p}_{srs}; srs, n)} \\ &= [Meff_w][Deff_{CS}(\hat{\theta})] = \left[1 + \frac{s_w^2}{\bar{w}^2}\right][1 + \rho_{\theta}(\bar{m} - 1)] . \end{aligned} \quad (1)$$

Data Source(s)

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

Computational Software

The same survey analysis software packages as seen in the task for margin of error (refer to *Appendix C.5*).

Computation

The estimated variance or standard error (i.e., $v(\hat{p}; design, n)$ or $\sqrt{v(\hat{p}; design, n)}$ in the Equation 1, respectively) is returned by survey analysis software. To estimate ρ_θ we can simply solve for ρ_θ in Equation 1 and we get,

$$\rho_\theta = \frac{\frac{Deff(\hat{\theta})}{Meff_w} - 1}{\bar{m} - 1} \quad (2)$$

To estimate ρ_θ for each survey estimate we can compute $Meff_w$ from the sample, we know \bar{m} by knowing the number of sample PSUs (a) and the overall respondent sample size (n) and for many types of survey analysis software (e.g., SUDAAN) we can request an estimate of $Deff(\hat{\theta})$ along with the estimate ($\hat{\theta}$). If an estimate of $Deff(\hat{\theta})$ is not available directly from the software output for use in determining $Deff(\hat{\theta})$ as defined in Equation 1 or ρ_θ as seen in Equation 2, then for estimated proportions we can estimate the overall design effect as,

$$deff(\hat{\theta}) = \frac{v(\hat{p}; design, n)}{v(\hat{p}_{srs}; srs, n)} = \frac{v(\hat{p}; design, n)}{\hat{p}(1-\hat{p})/n},$$

since $v(\hat{p}_{srs}; srs, n) = \hat{p}(1-\hat{p})/n$. The estimate of ρ_θ can then be obtained directly from Equation 2.

Interpretation

One usually hopes for the overall design effect for an estimate, or $Deff(\hat{\theta})$, to be less than 2.50. Estimates of ρ_θ (for within-PSU homogeneity) are typically small positive numbers between 0 and 1 (most typically between 0 and 0.05). In some instances, estimates of ρ_θ will be small negative numbers.

To summarize levels of $deff(\hat{\theta})$ among all survey estimates in GATS, it is recommended to estimate $Deff(\hat{\theta})$ for several key survey outcome measures (e.g., current smoking prevalence, ever smoking prevalence, quit rates, etc.) by various reporting domains (e.g., by gender, age, urban/rural, region, and various cross-classifications of these variables). Then for each key outcome measure summarize the estimated values of $Deff(\hat{\theta})$ among all reporting domains by reporting the following summary statistics for the estimate values of $Deff(\hat{\theta})$: median, minimum, and maximum.

Summarize values of $\hat{\rho}_\theta$ (median, minimum, and maximum) among all survey estimates in GATS in the same way you did for $deff(\hat{\theta})$.

Some survey systems or organizations prefer reporting design effect denoted by 'deft' in place of 'deff'. Deft is calculated for each estimate as the ratio between the standard error using the given sample design and the standard error that would result if a simple random sample had been used.

C.5 Margin of Error for Key Survey Estimates

Background

Some basic definitions of measures of statistical precision from survey samples are needed. First, note that all estimates from samples are likely to differ somewhat from what is being estimated in the population. This difference is generally referred to as **sampling error**. Probability samples like those used in GATS enable us to produce summary measures of sampling error that indicate the precision of estimates. There are several different summary measures of precision for survey estimates. To define these measures, let us use the symbol, θ , to denote the population characteristic we are estimating (e.g., the prevalence rate of persons who currently smoke cigarettes). We will use $\hat{\theta}$ to represent the estimate of θ based on the sample.

The first common summary measure of precision is the **variance of the estimate**, written as $V(\hat{\theta})$. As we shall see, all other measures are somehow related to $V(\hat{\theta})$. The variance of survey estimates, and all other related measures, are quantitative indicators of how much sample estimates would vary among all of the possible samples that the sample design used in the survey could produce. A second measure of precision is the **standard error of the estimate**, defined as $SE(\hat{\theta}) = \sqrt{V(\hat{\theta})}$. The standard error of an estimate is thus simply the square root of the variance of that estimate, and is computed in the same units as the estimate $\hat{\theta}$. A third measure of the statistical precision is the **relative standard error** of the estimate, defined as $RSE(\hat{\theta}) = SE(\hat{\theta}) / \theta = \sqrt{V(\hat{\theta})} / \theta$. Since $RSE(\hat{\theta})$ measures precision relative to the size of what is being estimated, it is unit-free and thus a more comparable indicator of precision among estimates than the $SE(\hat{\theta})$.

The recommended measure of precision for countries to report for key estimates in GATS (e.g., tobacco use prevalence rate, etc.) is the **margin of error**, defined as $MOE(\hat{\theta}) = [Z][SE(\hat{\theta})]$, where Z is a measure of the level of confidence for the measure and $SE(\hat{\theta})$ is the standard error of $\hat{\theta}$. For most national estimates in GATS we could use $Z = 1.96$ for a 95% level of confidence. The estimated value ($moe(\hat{\theta})$) reported for $\hat{\theta}$ is interpreted as follows:

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

GATS analysts are urged to report the value of $moe(\hat{\theta})$ for all key estimates, including all national estimates of the main indicators of tobacco use. $moe(\hat{\theta})$ for corresponding regional estimates should also be reported if the sample design calls for increased regional sample sizes to accommodate the production of regional estimates that will meet GATS precision standards.

Since measures of precision can easily be obtained in statistical software, GATS countries are strongly urged to routinely request precision estimates for all of its generated estimates, even though they only need to be reported for major national (and possibly regional) estimates. Estimates whose $MOE(\hat{\theta})$ approaches or exceeds the size of the estimate itself should be reported with caution, if at all. As a further word of caution, it is important to note that just as with $\hat{\theta}$ itself, estimates of $V(\hat{\theta})$, $SE(\hat{\theta})$, $MOE(\hat{\theta})$, or $RSE(\hat{\theta})$ are subject to sampling error. This is important since all of these estimates may be imprecise if they are based on small samples, particularly those involving less than 10-20 sample PSUs.

Data Source(s)

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

Computational Software

Estimates of population characteristics ($\hat{\theta}$) and their corresponding values of $moe(\hat{\theta})$ must take into account statistically important features of the sample design. Estimated characteristics must be weighted and associated precision measures must accommodate the use of stratification, cluster sampling, without-replacement sampling, and sample weights. It is well known that failure to do so typically contributes to biased estimates (particularly of precision) and thereby inappropriate interval estimates and tests of significance. Therefore, GATS country analysts are strongly urged to use analysis software that allows one to fully account for the sample design used to produce the survey data. This means using sample weights to produce all estimates (of $\hat{\theta}$ and $moe(\hat{\theta})$) for descriptive analysis. It also means using software that follows a widely accepted approach to estimate variances and standard errors of survey estimates. The three main approaches to variance estimation are the Taylor Series Linearization (TSL) approach, the Balanced Repeated Replication (BRR) approach, and the Jackknife Repeated Replication (JRR) approach (Wolter, 1985). Most survey analysis packages use a Taylor Series Linearization (TSL) approach to variance estimation, and so for reasons of comparability among GATS country findings, software that follows the TSL approach is recommended.

Several statistical organizations around the world have developed computer software to analyze data from complex samples like GATS¹. These software programs will not only produce estimates of tobacco use (i.e., $\hat{\theta}$) but they can also produce estimates of precision (i.e., usually either $V(\hat{\theta})$ or $SE(\hat{\theta})$) that appropriately account for key design features in GATS, namely cluster sampling, the use of stratification, and varying selection probabilities (i.e., sample weights):

More Information

For more detail on the definitions given above, see the *Encyclopedia of Survey Research Methods*, (2007). For the online information to all of the survey analysis software packages, visit the website sponsored by the Survey Research Methods Section of the American Statistical Association (<http://www.hcp.med.harvard.edu/statistics/survey-soft/>). In addition to these links to analysis software, the ASA-SRMS website also includes several other useful links and resources on survey design and implementation (see <http://www.amstat.org/sections/SRMS/index.html>).

¹ The GATS Data Coordinating Center (DCC) at CDC provides technical assistance for use of the following statistical software packages: SAS, SPSS, and STATA.

Computation

Output from the software listed above will report a value of $\hat{\theta}$, as well as its estimated variance, denoted by $v(\hat{\theta})$, or its standard error, written as $se(\hat{\theta})$. From these reported values, one can compute the estimate margin of error for $\hat{\theta}$, as

$$moe(\hat{\theta}) = [t][se(\hat{\theta})] = [t]\sqrt{v(\hat{\theta})} \quad (1)$$

Interpretation

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

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

C.6 Estimates of Sampling Errors

The estimates from a sample survey are affected by two types of error: (1) non-sampling errors, and (2) sampling errors. *Non-sampling errors* are the result of errors or mistakes that cannot be attributable to sampling and were made in implementing data collection and data processing, such as errors in coverage, response errors, non-response errors, faulty questionnaires, interviewer recording errors, data processing errors, etc. Although numerous efforts were made during the implementation of GATS in [country] to minimize those errors, non-sampling errors are impossible to avoid and difficult to evaluate statistically.

The sample of respondents selected in the GATS [country] was only one of the samples that could have been selected from the same population, using the same design and sample size. Each of these samples would yield results that differed somewhat from the results of the actual sample selected. *Sampling errors* are a measure of the variability between all possible samples. The extent of variability is not known exactly, but can be estimated statistically from the survey results.

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

- **Value (R):** Weighted prevalence estimate of the indicator
- **Standard error (SE):** Sampling errors are usually measured in terms of standard errors for particular estimate or indicator (R). Standard error of an estimate is thus simply the square root of the variance of that estimate, and is computed in the same units as the estimate.
- **Sample size (n):** Total number of observations used to calculate the prevalence estimate (R).
- **Design effect:** *Design effect* denoted by 'deff' is the ratio of the actual variance of an indicator, under the sampling method used in the survey, to the variance calculated under the assumption of simple random sampling. The square root of the design effect denoted by 'deft' is used to show the efficiency of the sample design and is calculated for each estimate as the ratio between the standard error using the given sample design and the standard error that would result if a simple random sample had been used. A DEFT value of 1.0 indicates that the sample design is as efficient as a simple random sample, while a DEFT value above 1.0 indicates the increase in the standard error due to the use of a more complex sample design. In general, for a well-designed

survey, DEFT usually ranges from 1 to 3. It is common, however, for DEFT to be much larger, up to 7 or 8.

- **Relative standard error (RSE):** Relative standard error also known as coefficient of variation (CV) is the ratio of the standard error to the value of the indicator.
- **Margin of Error (MOE):** Margin of error is computed as the product of the desired confidence measure and the standard error of the estimate. The level of confidence is usually based on a value (Z) of the standard normal distribution. For example, for a 95% level of confidence, we can use $Z=1.96$.
- **Confidence limits** ($R \pm 1.96SE$) are calculated to show the interval within which the true value for the population can be reasonably assumed to fall. For any given statistic calculated from the survey, the value of that statistics will fall within a range of plus or minus two times the standard error of the statistic in 95 percent of all possible samples of identical size and design.

Calculation of Standard Error

If the sample of respondents had been selected as a simple random sample, it would have been possible to use straightforward formulas for calculating sampling errors. However, the GATS [year] sample is the result of a multi-stage stratified design, and, consequently, it was necessary to use more complex formulae. For the calculation of sampling errors from GATS [country] data, [statistical software version] was used. The Taylor linearization method of variance estimation was used for survey estimates that are means or proportions.

The Taylor linearization method treats any percentage or average as a ratio estimate, $r = y/x$, where y represents the total sample value for variable y , and x represents the total number of cases in the group or subgroup under consideration. The variance of r is computed using the formula given below:

$$SE^2(r) = \text{var}(r) = \frac{1-f}{x^2} \sum_{h=1}^2 \left[\frac{m_h}{m_h - 1} \left(\sum_{i=1}^{m_h} Z_{hi}^2 - \frac{Z_h^2}{m_h} \right) \right]$$

in which, $Z_{hi} = y_{hi} - rx_{hi}$, and $Z_h = y_h - rx_h$

where h (=1 or 2) represents the stratum which is urban or rural,

m_h is the total number of PSUs selected in the h th stratum,

y_{hi} is the sum of the weighted values of variable y in the i th PSU in the h th stratum,

x_{hi} is the sum of the weighted number of cases in the i th PSU in the h th stratum, and

f is the overall sampling fraction, which is so small that it is ignored.

The results are presented in this appendix for the country as a whole, for urban and rural areas, and for gender. For each variable or indicator, the type of statistic (mean, proportion, or rate) and the base population are given in *Table C-5*. In addition to the standard error (SE) described above, *Table C-6* includes the value of the estimate (R), the sample size, the design effect (DEFF or DEFT), the relative standard error (SE/R), Margin of error (MOE) and the 95 percent confidence limits ($R \pm 1.96SE$), for each variable or indicator.

Table C-5: List of Indicators for Sampling Errors, GATS [country] [year]

Indicator	Estimate	Base Population
Current Tobacco Users	Proportion	Adults \geq 15 years old
Current Tobacco Smokers	Proportion	Adults \geq 15 years old
Current Cigarette Smokers	Proportion	Adults \geq 15 years old
Current Users of Smokeless Tobacco	Proportion	Adults \geq 15 years old
Daily Tobacco Users	Proportion	Adults \geq 15 years old
Daily Tobacco Smoker	Proportion	Adults \geq 15 years old
Daily Cigarette Smokers	Proportion	Adults \geq 15 years old
Daily Users of Smokeless Tobacco	Proportion	Adults \geq 15 years old
Former Daily Tobacco Users Among All Adults	Proportion	Adults \geq 15 years old
Former Tobacco Users Among Ever Daily Tobacco Users	Proportion	Ever daily tobacco users \geq 15 years old
Former Daily Tobacco Smokers Among All Adults	Proportion	Adults \geq 15 years old
Former Tobacco Smokers Among Ever Daily Smokers	Proportion	Ever daily tobacco smokers \geq 15 years old
Former Daily Users of Smokeless Tobacco Among All Adults	Proportion	Adults \geq 15 years old
Former Daily Users of Smokeless Tobacco Among Ever Daily Users of Smokeless Tobacco	Proportion	Ever daily users of smokeless tobacco \geq 15 years old
Time to First Tobacco use within 5 minutes of waking	Proportion	Daily tobacco users \geq 15 years old
Time to First Tobacco use within 6-30 minutes of waking	Proportion	Daily tobacco users \geq 15 years old
Smoking Quit Attempt in the Past 12 Months	Proportion	Current smokers and former smokers who have been abstinent for less than 12 months
Smokeless tobacco Quit Attempt in the Past 12 Months	Proportion	Current smokeless tobacco users and former smokeless users who have been abstinent for less than 12 months
Health Care Provider Asked about Smoking	Proportion	Current smokers and former smokers who have been abstinent for less than 12 months and who visited a HCP during the past 12 months
Health Care Provider Advised Quitting Smoking	Proportion	Current smokers and former smokers who have been abstinent for less than 12 months and who visited a HCP during the past 12 months
Use of Pharmacotherapy for Smoking Cessation	Proportion	Current smokers and former smokers who have been abstinent for less than 12 months
Use of Counseling/Advice or Quit Lines for Smoking Cessation	Proportion	Current smokers and former smokers who have been abstinent for less than 12 months
Planning to quit, thinking about quitting, or will quit smoking	Proportion	Current smokers \geq 15 years old
Exposure to SHS at Home	Proportion	Adults \geq 15 years old
Exposure to SHS at Workplace	Proportion	Adults who work indoors
Exposure to SHS in Government Buildings/Offices	Proportion	Adults \geq 15 years old
Exposure to SHS in Health Care Facilities	Proportion	Adults \geq 15 years old
Exposure to SHS in Restaurants	Proportion	Adults \geq 15 years old
Exposure to SHS in Public Transportation	Proportion	Adults \geq 15 years old
Last cigarette purchase in store	Proportion	Current manufactured cigarette smokers \geq 15 years old
Last cigarette purchase at street vendor	Proportion	Current manufactured cigarette smokers \geq 15 years old

Indicator	Estimate	Base Population
Last cigarette purchase at kiosk	Proportion	Current manufactured cigarette smokers \geq 15 years old
Noticed Anti-tobacco Information on radio or television	Proportion	Adults \geq 15 years old
Noticed Health Warning Labels on Cigarette Packages	Proportion	Current smokers \geq 15 years old
Thinking of Quitting Because of Health Warning Labels on Cigarette Package	Proportion	Current smokers \geq 15 years old
Noticed Any Cigarette Advertisement or Promotion	Proportion	Adults \geq 15 years old
Believes that Tobacco Smoking Causes Serious Illness	Proportion	Adults \geq 15 years old
Believes that Tobacco Smoking Causes Strokes	Proportion	Adults \geq 15 years old
Believes that Tobacco Smoking Causes Heart Attacks	Proportion	Adults \geq 15 years old
Believes that Tobacco Smoking Causes Lung Cancer	Proportion	Adults \geq 15 years old
Believes that Using Smokeless Tobacco Causes Serious Illness	Proportion	Adults \geq 15 years old
Believes that SHS Causes Serious Illness in Non-Smokers	Proportion	Adults \geq 15 years old
Number of Cigarettes Smoked per Day (by daily smokers)	Mean	Current daily cigarette smokers \geq 15 years old
Time since Quitting Smoking (in years)	Mean	Former smokers \geq 15 years old
Monthly Expenditures on Manufactured Cigarettes	Mean	Current manufactured cigarette smokers \geq 15 years old
Age at Daily Smoking Initiation	Mean	Ever daily smokers \geq 15 years old

Table C-6: Example Table for Reporting Sampling Errors

Indicator	Estimate (R)	Standard Error (SE)	Sample size (n)	Design Effect (DEFF or DEFT)	Relative Error (SE/R)	Margin of Error (MOE)	Confidence limits	
							Lower Limit (R-1.96SE)	Upper Limit (R+1.96SE)
Current Tobacco Users								
Current Tobacco Smokers								
Current Cigarette Smokers								
Current Users of Smokeless Tobacco								
Daily Tobacco Users								
Daily Tobacco Smoker								
Daily Cigarette Smokers								
Daily Users of Smokeless Tobacco								
Former Daily Tobacco Users Among All Adults								
Former Tobacco Users Among Ever Daily Tobacco Users								
Former Daily Tobacco Smokers Among All Adults								
Former Tobacco Smokers Among Ever Daily Smokers								
Former Daily Users of Smokeless Tobacco Among All Adults								
Former Daily Users of Smokeless Tobacco Among Ever Daily Users of Smokeless Tobacco								
Time to First Tobacco use within 5 minutes of waking								
Time to First Tobacco use within 6-30 minutes of waking								
Smoking Quit Attempt in the Past 12 Months								
Smokeless tobacco Quit Attempt in the Past 12 Months								
Health Care Provider Asked about Smoking								
Health Care Provider Advised Quitting Smoking								
Use of Pharmacotherapy for Smoking Cessation								
Use of Counseling/Advice or Quit Lines for Smoking Cessation								
Planning to quit, thinking about quitting, or will quit smoking								
Exposure to SHS at Home								
Exposure to SHS at Workplace								
Exposure to SHS in Government Buildings/Offices								
Exposure to SHS in Health Care Facilities								

Indicator	Estimate (R)	Standard Error (SE)	Sample size (n)	Design Effect (DEFF or DEFT)	Relative Error (SE/R)	Margin of Error (MOE)	Confidence limits	
							Lower Limit (R-1.96SE)	Upper Limit (R+1.96SE)
Exposure to SHS in Restaurants								
Exposure to SHS on Public Transportation								
Last cigarette purchase in store								
Last cigarette purchase at street vendor								
Last cigarette purchase in kiosk								
Noticed Anti-tobacco Information on radio or television								
Noticed Health Warning Labels on Cigarette Packages								
Thinking of Quitting Because of Health Warning Labels on Cigarette Package								
Noticed Any Cigarette Advertisement or Promotion								
Believes that Tobacco Smoking Causes Serious Illness								
Believes that Tobacco Smoking Causes Strokes								
Believes that Tobacco Smoking Causes Heart Attacks								
Believes that Tobacco Smoking Causes Lung Cancer								
Believes that Using Smokeless Tobacco Causes Serious Illness								
Believes that SHS Causes Serious Illness in Non-Smokers								
Number of Cigarettes Smoked per Day (by daily smokers)								
Time since Quitting Smoking (in years)								
Monthly Expenditure on Manufactured Cigarettes								
Age at Daily Smoking Initiation								

Notes:

1. Tables should be developed for subgroups (e.g., gender, urban/rural, and region2. Refer to the *GATS Indicator Definitions* for specific definitions and calculations of the following indicators).
2. The indicators provided in this table are produced from the core questionnaire. These can be adapted for each country as appropriate. In addition, countries may wish to produce the estimates for a subset of the indicators, rather than for all, as applicable.
3. Countries may wish to produce either 'deff' or 'deft' based on the standard practice applicable in country/organization.
4. The above tables may be included in the country report appendix as estimates of sampling errors to describe the data quality.

C.7 Household Frame Coverage Rate

Background

A key part of the sampling process in a household survey is the enumeration of households in the sampling stage where household addresses are chosen. Different household enumeration approaches may be used, and the usual end-result is that there is a net under-enumeration of households and thus incomplete coverage in the final sample of households. Since the biasing effect of undercoverage depends both on: (i) the aggregate difference between key study measures for households covered by the frame and those missed by the frame, and (ii) the extent of noncoverage (or 100 percent minus the household coverage rate) in the sample, it is useful to estimate the frame coverage rate for households. A frame coverage rate (FCR) is defined as,

$$FCR = \frac{N_F}{N} = \frac{\# \text{ of households accounted for by the household frame}}{\# \text{ of households in the population}} . \quad (1)$$

Data Source(s)

Two sources of data will be needed for these calculations. One is the final weighted data file used for analysis, which is used to estimate N_F in Equation 1. The other is the best available source of the country-wide household count, which is used as the measure of N in Equation 1. The latter source may be the most recent census, more recent updates to the last census using demographic methods, or a recent national survey who sample size and quality are at least as great as the survey done for GATS.

Computational Software

None in particular since the computations required can easily be programmed in any software environment.

Computation

The number of households accounted for by the household frame (N_F) can be estimated in either of two ways. Both require some of the stage-specific selection probabilities that were used to compute each respondent's base weight. One is a weighted estimate of N_F using base weights for sampling "segments" in the sample design (the *weighted sum approach*); the other is a sum of base weights for all selected households in the GATS sample, which is also an estimate of N_F (the *sum of household weights approach*).

Weighted Sum Approach -- In most GATS designs, segments will be the secondary sampling units (SSUs), but in general they are the area sampling units within which household enumeration is conducted to produce the sampling frames from which households are chosen. Assuming that the segment is the SSU, define W_j as the base weight for the j -th sample segment, computed as the one divided by the product of the selection probabilities for the PSU in which the segment is located times the segment-within-PSU selection probability. If M_j denotes the number of enumerated households in the j -th sample segment, then the estimate of N_F under this approach is then computed as,

$$\hat{N}_F^{(i)} = \sum_j^{All \text{ sample segments}} W_j M_j \quad (2)$$

Sum of Household Weights Approach – In this approach one uses the base weight for each selected household. The set of selected households includes those that are assigned to the field for household recruitment during GATS data collection. Let W_{jk} denote the base weight for the k-th selected household in the j-th sample segment. We compute W_{jk} as W_j divided by the k-th household's within-segment selection probability. The estimate of N_F under this approach is then computed as,

$$\hat{N}_F^{(ii)} = \sum_{j,k}^{All \text{ selected households}} W_{jk} \quad (3)$$

The best sources of a suitable value of the actual number of households in the population (N) are: (1) the most recent population census (provided it is generally viewed as of high quality), or (2) an estimated household count for N_F (obtained using Equations 2 or 3) from a recent and statistically superior sample survey (e.g., the master sample from which the GATS sample was randomly subselected).

Interpretation

An estimated (*FCR*) that is 95% or higher would indicate that the quality of household enumeration in the GATS sample was of acceptable quality.

Global Adult Tobacco Survey (GATS)